LUND UNIVERSITY

# The Almost People

## A Framework Proposal for the Balancing of Legal Interests in the Age of Social Robots

Öztürk, Anil

2023

Link to publication

*Citation for published version (APA):*
Öztürk, A. (2023). *The Almost People: A Framework Proposal for the Balancing of Legal Interests in the Age of Social Robots*. [Doctoral Thesis (monograph), Maynooth University].

*Total number of authors:*
1

*Creative Commons License:*
CC BY-NC-SA

**General rights**
Unless other specific re-use rights are stated the following general rights apply:
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.
• Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the public portal

Read more about Creative commons licenses: https://creativecommons.org/licenses/

**Take down policy**
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

LUND UNIVERSITY

PO Box 117
221 00 Lund
+46 46-222 00 00

Download date: 26. Aug. 2025

# THE ALMOST PEOPLE:

## A FRAMEWORK PROPOSAL FOR THE BALANCING OF LEGAL INTERESTS IN THE AGE OF SOCIAL ROBOTS

A Thesis Submitted in Fulfilment of the Requirements for
the Degree of Doctor of Philosophy in Law

**BY:**
Anil Ozturk

Maynooth University
School of Law and Criminology

4 October 2022

**Head of School:** Dr Fergus Ryan
**Supervisor:** Dr Brian Flanagan

# Table of Contents

**Acknowledgements** ..... **IV**

**Dedication** ..... **V**

**Declaration** ..... **VI**

**Abstract** ..... **VII**

**List of Abbreviations** ..... **VIII**

**Table of Cases** ..... **X**

**Table of Legislation** ..... **XII**

**PART A - BACKGROUND CONCEPTS AND MOTIVATION** ..... **14**

**CHAPTER I: GENERAL INTRODUCTION** ..... **15**

I.1 Background of the Study ..... 15
I.2 Scope of the Study ..... 25
*I.2.1 Legal Issues* ..... *25*
*I.2.2 Everyday Contexts* ..... *26*
*I.2.3 Social Interactions* ..... *29*
*I.2.4 Information Security and Privacy* ..... *32*
I.3 Objectives of the Study ..... 33
I.4 Significance of the Study ..... 34
I.5 Organisation of the Study ..... 36

**CHAPTER II: METHODOLOGY OF THE STUDY** ..... **39**

Introduction ..... 39
II.1 Function of the Law ..... 40
II.2 Function of Comparative Law ..... 42
II.3 Design of the Study ..... 46
*II.3.1 Conceptualisation of Legal Issues* ..... *46*
*II.3.2 Identification of Legal Values* ..... *47*
*II.3.3 Construction of New Legal Framework* ..... *48*
II.4 Choice of Legal Systems ..... 49
*II.4.1 EU Law* ..... *49*
*II.4.2 National Legal Systems* ..... *59*
Summary ..... 64

**CHAPTER III: CONCEPTUALISATION OF SOCIAL ROBOTS** ..... **65**

Introduction ..... 65
III.1 Evolution of the Robot Concept ..... 66
*III.1.1 Intellectual Origins: Mythological and Religious Accounts* ..... *66*
*III.1.2 Technological Origins: Automata of the Pre-Industrial Era* ..... *69*
*III.1.3 Formal Origins: Contributions of Science-Fiction* ..... *71*
III.2 Existing Definitions ..... 75
III.3 Distinctive Characteristics of Social Robots ..... 78
*III.3.1 Characteristic of Relative Autonomy* ..... *79*
*III.3.2 Characteristic of Social Agency* ..... *87*
Summary ..... 91

**PART B - HUMANESQUE: IMPLICATIONS OF ROBOT ANTHROPOMORPHISM ..... 93**

**CHAPTER IV: CHALLENGES OF ROBOT ANTHROPOMORPHISM ..... 94**

INTRODUCTION ..... 94
IV.1 ROBOT ANTHROPOMORPHISM ..... 96
*IV.1.1 Recorded Experiences ..... 96*
*IV.1.2 Psychological Experiments ..... 104*
IV.2 UNDESIRABLE OUTCOMES ..... 108
*IV.2.1 Loss of Genuine Human Connections ..... 108*
*IV.2.2 Diminution of Privacy and Solitude ..... 110*
*IV.2.3 Exploitation of Emotional Attachments ..... 115*
*IV.2.4 Robot Abuse ..... 117*
IV.3 LEGAL PROTECTION OF SOCIAL ROBOTS ..... 118
*IV.3.1 Moral Wrongness of Robot Abuse ..... 118*
*IV.3.2 Moral Wrongs and Criminal Law ..... 122*
*IV.3.3 Animal Protection Legislation ..... 124*
SUMMARY ..... 128

**PART C - FREEISH: IMPLICATIONS OF ROBOT AUTONOMY ..... 130**

**CHAPTER V: CHALLENGES OF CRIMINAL LIABILITY ..... 131**

INTRODUCTION ..... 131
V.1 ROBOT CRIME ..... 134
V.2 IRREDUCIBILITY OF ROBOT CRIMES ..... 140
V.3 JUSTIFICATION OF CRIMINAL LIABILITY ..... 142
*V.3.1 Affirmative Justifications ..... 144*
*V.3.2 Negative Limitations ..... 152*
*V.3.3 Possible Alternatives ..... 154*
V.4 INTENTIONAL STANCE ..... 157
*V.4.1 Robotic Mens Rea ..... 159*
*V.4.2 Robotic Criminal Punishment ..... 161*
SUMMARY ..... 164

**CHAPTER VI: CHALLENGES OF CIVIL LIABILITY ..... 166**

INTRODUCTION ..... 166
VI.1 CONTRACT LAW ..... 167
*VI.1.1 Core Principles of Contract Law ..... 169*
*VI.1.2 Shortcomings of Contract Law ..... 183*
VI.2 TORT LAW ..... 186
*VI.2.1 General Rule: Fault-Based Liability ..... 188*
*VI.2.2 Strict Liability ..... 193*
SUMMARY ..... 202

**PART D - PAX ROBOTA: RECOMMENDATIONS AND CONCLUSIONS ..... 204**

**CHAPTER VII: FEASIBILITY OF THE NEW LEGAL FRAMEWORK ..... 205**

INTRODUCTION ..... 205
VII.1 CONCEPTUAL POSSIBILITY OF NEW LEGAL STATUS ..... 207
VII.2 PRAGMATIC DESIRABILITY OF NEW LEGAL STATUS ..... 213
*VII.2.1 Arguments for New Legal Status ..... 215*
*VII.2.2 Arguments Against New Legal Status ..... 217*
SUMMARY ..... 220

**CHAPTER VIII: FUTURE OF ROBOT LAW ..... 222**

INTRODUCTION ..... 222
VIII.1 SUMMARY OF THE FINDINGS ..... 224
*VIII.1.1 Causes of Legal Issues ..... 224*
*VIII.1.2 Legal Issues ..... 227*
*VIII.1.3 Existing Proposals ..... 234*

VIII.2 Unified Framework for Robot Law ..... 239
*VIII.2.1 Modifications in Criminal Law ..... 240*
*VIII.2.2 Modifications in Private Law ..... 244*
Concluding Remarks ..... 259

**BIBLIOGRAPHY ..... CCLXI**

Books ..... CCLXI
Electronic Books (E-Books) ..... CCLXIII
Book Chapters ..... CCLXIV
Dictionaries & Encyclopaedias ..... CCLXVII
Conference Proceedings ..... CCLXVIII
Reports and Other Documents ..... CCLXIX
Journal Articles ..... CCLXX
Website and Blogs ..... CCLXXX

# Acknowledgements

First, I am deeply indebted to my supervisor, *Dr Brian Flanagan*, for his invaluable guidance and encouragement throughout the doctoral research process. The expert knowledge and insightful feedback he selflessly offered have shaped my research and improved its quality.

I am incredibly appreciative of my family for the unwavering support and motivation they offered at every step of my academic journey. I am extremely grateful to my father, *Muammer Ozturk*; whose commitment to the principles of justice during his forty years of judicial career is a continuous source of inspiration for me; and to my mother, *Nilgun Ozturk*, whose uncompromising confidence in the strength of my work (though she does not speak a word of English) keeps my momentum up. I would also like to thank my dear brother, *Omur Ozturk*, whose brief residence in Ireland and deep affection for this country and its people motivated me to come here for my Master's degree six years ago; I have only infrequently regretted this decision.

I would like to express my sincere gratitude to the beloved *Andy Morris,* who stood with me through the thick and thin of many adventures, including this one, helping me to stay focused and motivated during moments of doubt and uncertainty. Although he is no longer here with us, Andy's memory and love remain a constant source of strength and courage. In times of grief, it is not easy to focus on anything. On that note, words cannot express the depth of my appreciation for the wisdom and loving support so selflessly extended by *Marie Morris*, *Martin Morris*, and *Delphi Morris* in the past few months.

I would also like to give my thanks to the Faculty and staff members at Maynooth University School of Law and Criminology for providing a nurturing and enjoyable research environment over the years.

Finally, I thank my close friends for their patience throughout my seemingly endless academic endeavours.

# Dedication

This thesis is gratefully dedicated to *Andy Morris;* my best friend, my companion, and my partner-in-crime.

> Out beyond ideas of right and wrong,
> There is a field.
> I'll meet you there.
>
> —Rumi, *Selected Poems*[Ω]

[Ω] Rumi, *Selected Poems* (Coleman Barks tr, Penguin 2004) 36.

# Declaration

I hereby certify that the following work, which I now submit for assessment on the programme of study leading to the award of PhD, is composed solely by myself. It contains no material that has been previously submitted, in whole or in part, in any form, for another degree or diploma at any university or other institution of tertiary education. Except where stated otherwise by reference or acknowledgment, the work presented is entirely my own.

Student Name: **ANIL ÖZTÜRK**

Student Number: **17250686**

Date: **4th October 2022**

Student Signature:

# Abstract

Robots, which were seen as gimmicks in science fiction stories until not so long ago, have already crossed into reality. Thanks to the ever-growing autonomy of robots and ever-expanding variety of roles assigned to them, they are becoming more integrated into the ordinary course of everyday life. With the advent of social robots that can engage human beings on personal levels, for the first time, non-human entities are emerging as social interaction partners. In that regard, from the legal perspective, it is no longer possible to treat them as mere tools.

The autonomy of robots is expected to have significant impacts on various interests recognised by the legal principles that underlie existing legal instruments. However, almost none of the existing legal instruments were developed in consideration of the implications of robots' emerging roles as independent social actors. On explaining the inadequacy of existing legal instruments, I outline the prospect of a paradigm shift in the law's approach to human-robot social interactions.

A comparative analysis of German, Italian, and Irish legal systems -selected to represent the EU's diverse legal families- demonstrates that robots' autonomous behaviours and emerging roles as social interaction partners are likely to undermine the legal principles expressed most notably in the domains of private law (contract law and tort law) and criminal law. The conceptual deconstruction of existing legal instruments offered by these domains reveals that legal systems overlook the characteristics of social robots that set them apart from other artefacts, namely, their relative autonomy and social agency. These distinctive characteristics allow robots to perform unpredictable behaviours and to prompt human beings they interact with to anthropomorphise them. Overlooking these characteristics diminishes the adequacy of existing legal instruments

Ultimately, I conclude that the shortcomings of contemporary legal systems can be overcome by creating a new, unified legal framework that would enable the law to respond to the legal implications of robot autonomy and the phenomenon of robot anthropomorphism.

# List of Abbreviations

| | |
|---|---|
| **A 2d** | Atlantic Reporter, 2nd Series |
| **AC** | Appeals Cases |
| **AI** | Artificial Intelligence |
| **All ER** | All England Reports |
| **BGB** | Bürgerliches Gesetzbuch<br>*Civil Code* |
| **BGBl** | Bundesgesetzblatt<br>*Federal Law Gazette* |
| **BGH** | Bundesgerichtshof<br>*Federal Court of Justice* |
| **BGHZ** | Entscheidungen des Bundesgerichthofes in Zivilsachen<br>*Decisions of the Federal Court of Justice in Civil Matters* |
| **BVerfG** | Bundesverfassungsgericht<br>*Federal Constitutional Court* |
| **BVerfGE** | Entscheidungen des Bundesverfassungsgerichts<br>*Decisions of the Federal Constitutional Court* |
| **CA** | Court of Appeal |
| **Ch** | Chancery |
| **Ch D** | Chancery Division of the High Court |
| **CC** | Codice Civile<br>*Civil Code* |
| **Corte Cost** | Corte Costituzionale<br>*Constitutional Court* |
| **Cost** | Costituzione della Repubblica Italiana<br>*Constitution of the Italian Republic* |
| **CP** | Codice Penale<br>*Penal Code* |
| **DCFR** | Draft Common Frame of Reference |
| **EC** | European Commission |
| **ECHR** | European Convention of Human Rights |
| **ECJ** | European Court of Justice |
| **ECR** | European Court Reports |
| **ECtHR** | European Court of Human Rights |
| **ER** | English Reports |
| **EU** | European Union |
| **Exch Ch** | Court of Exchequer Chamber |
| **EP** | European Parliament |
| **EUCo** | European Council |
| **F 3d** | Federal Reporter, 3rd Series |
| **GG** | Grundgesetz<br>*Basic Law [for the Federal Republic of Germany]* |
| **HC** | High Court |
| **HL** | House of Lords |
| **ILRM** | Irish Law Reports Monthly |
| **IR** | Irish Reports |
| **Ir Jur Rep** | Irish Jurist Reports |
| **LR** | Law Reports |
| **Md Ct Spec App** | Maryland Court of Special Appeals |
| **Mo App** | Missouri Court of Appeals |
| **NJW** | Neue Juristische Wochenschrift<br>*New Legal Weekly Journal* |

| | |
|---|---|
| **OLG** | Oberlandesgericht<br>*Higher Regional Court* |
| **PC** | Privy Council |
| **QB/KB** | Queen's Bench/King's Bench Division of the High Court |
| **RG** | Reichsgericht<br>*Imperial Court of Justice* |
| **RGBl** | Reichsgesetzblatt<br>*Reich Law Gazette* |
| **RGZ** | Entscheidungen des Reichsgerichts in Zivilsachen<br>*Decisions of the Imperial Court of Justice in Civil Matters* |
| **SC** | Supreme Court |
| **SW 2d** | South Western Reporter, 2nd series |
| **TEU** | Treaty on the European Union |
| **TFEU** | Treaty on the Functioning of the European Union |
| **WLR** | Weekly Law Reports |

# Table of Cases

## International Cases

### ECtHR Cases

Roman Zakharov v Russia [2015] ECHR 1065
Szabo and Vissy v Hungary App No 37138/14 (ECHR, 12 January 2016)
Antovic and Mirkovic v Montenegro [2017] ECHR 1068

### CJEU Cases

Flaminio Costa v ENEL (Case 6/64) [1964] ECR 585
Van Gend en Loos v Nederlandse Administratie der Belastingen (Case 26/62) [1963] ECR 1

## Domestic Cases

### Germany

BGH March 31, 1982, 83 BGHZ 313
BVerfG October 19, 1993, 89 BVerfGE 214
BVerfG February 7, 1990, 81 BVerfGE 242
BVerfG February 26, 1968, 25 BVerfGE 269, 285
BVerfG October 25, 1966, 20 BVerfGE 323, 331
BVerfG January 16, 1957, 6 BVerfGE 32
OLG Hamm December 2, 1985, NJW 781 (1986)
RG January 8, 1926, 112 RGZ 290

### Italy

Corte Cost, sentenza 17 marzo 2000, n 70
Corte Cost, sentenza 30 giugno 1994, n 268
Corte Cost, sentenza 15 maggio 1990, n 241
Corte Cost, ordinanza 22-30 novembre 1988, n 1055
Corte Cost, sentenza 23-24 marzo 1988, n 364
Corte Cost, sentenza 13 marzo1969, n 37
Corte Cost, sentenza 21 marzo 1968, n 16
Corte Cost, sentenza 8 aprile 1965, n 30
Corte Cost, sentenza 20 febbraio 1962, n 7
Cass, sez III Penale, sentenza, 29 aprile 2019, n 17691

## Ireland

B v DPP [1997] 3 IR 140 (SC)
CC v Ireland [2006] 4 IR 66
Conole v Redbank Oyster Co [1976] IR 191 (SC)
Conway v Irish National Teachers Organisation [1991] 2 IR 305 (SC)
Donnelly v Timber Factors Ltd [1991] WJSC-SC 372, [1991] 1 IR 553 (SC)
DPP v Byrne [1994] 2 IR 236 (SC)
John Orr Ltd v Orr [1987] ILRM 702 (HC)
Lord Mayor, Alderman & Burgess of Dublin v Hayes [1876] IR 10 CL 226 (Exch Ch)
Maguire v Ardagh [2002] IESC 21, [2002] 1 IR 385
People (DPP) v D'OT [2003] IR 286
Rohan Construction Ltd v Insurance Corporation of Ireland Plc , [1986] ILRM 419
Stanhope v Hospitals Trust Ltd [1936] 2 Ir Jur Rep 25 (SC)

## United Kingdom

Allen v Flood [1898] 1 AC 1 (HL)
Attorney General v Corke [1933] Ch 89 (ChD)
Cambridge Water Co Ltd v Eastern Counties Leather plc [1994] 1 All ER 53 (HL)
Carslogie Steamship Co Ltd v Royal Norwegian Government [1952] AC 292 (HL)
Case of Sutton's Hospital (1612) 77 ER 937, 10 Co Rep 1
Halsey v Esso Petroleum Co Ltd [1961] 2 All ER 145 (QB)
Haynes v Harwood [1935] 1 KB 146 (CA)
Home Office v Dorset Yacht Co Ltd [1970] AC 1004 (HL)
Knightley v Johns [1982] 1 All ER 851 (CA)
McKew v Holland [1969] 3 All ER 1621 (HL)
Noble v Harrison [1926] 2 KB 332 (KB)
Overseas Tankship (UK) Ltd v Morts Dock and Engineering Co Ltd [1961] AC 388(PC)
Petrofina (Great Britain) Ltd v Martin [1966] 1 Ch 146 (CA)
Printing and Numerical Registering Co v Sampson (1874-75) LR 19 Eq 462
Read v J Lyons & Co Ltd [1947] AC 156 (HL)
Rylands v Fletcher (1868) LR 3 HL 330 (HL)
Transco plc v Stockport MBC [2003] UKHL 61, [2004] 2 AC 1

## United States

ACLU v NSA, 493 F 3d 644 (6th Cir 2007)
Comptroller v Family Entertainment Centers, 519 A 2d 1337 (Md Ct Spec App 1987)
Frye v Baskin, 231 SW 2d 630 (Mo App 1950)
Hale v Henkel, 201 US 43 (1906)
International Shoe Co v Washington, 326 US 310 (1945)
Lochner v New York, 198 US 45 (1905)
Trustees of Dartmouth College v Woodward, 17 US 518 (1819)
West Coast Hotel Co v Parrish, 300 US 379 (1937)
People v Ingram, 492 NE2d 1220, 1220-21 (NY 1986)
Malaske v State, 89 P3d 1116, 1117 n 1 (Okla Crim App 2004)

# Table of Legislation

## International Legislation

### International Agreements, Charters and Conventions

UN Convention on the Use of Electronic Communications in International Contracts (adopted 23 November 2005, entered into force 1 March 2013) 2898 UNTS 3

International Covenant on Civil and Political Rights (adopted 16 December 1966, entered into force 23 March 1976) 999 UNTS 171 [ICCPR]

Convention for the Protection of Human Rights and Fundamental Freedoms (European Convention on Human Rights, as amended) [ECHR]

### EU Legislation

Consolidated Version of the Treaty on European Union [2016] OJ C326/13

Consolidated Version of the Treaty on the Functioning of the European Union [2016] OJ C326/47

Charter of Fundamental Rights of the European Union [2012] OJ C326/95

Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation) OJ [2016] L119/1

Directive (EU) 2017/1371 of the European Parliament and of the Council of 5 July 2017 on the fight against fraud to the Union's financial interests by means of criminal law [2017] OJ L198/29.

Directive [2014] 2014/57/EU of the European Parliament and of the Council of 16 April 2014 on Criminal Sanctions for Market Abuse (Market Abuse Directive) OJ L 173/179

Council Directive 93/13/EEC of 5 April 1993 on Unfair Terms in Consumer Contracts [1993] OJ L95/29

Council Directive 89/391/EEC of 12 June 1989 on the Introduction of Measures to Encourage Improvements in the Safety and Health of Workers at Work [1989] OJ L183/1

European Parliament Resolution 2020/2014(INL) of 20 October 2020 with Recommendations to the Commission on a Civil Liability Regime for Artificial Intelligence [2021] OJ C404/10

European Parliament Resolution 2020/2012(INL) of 20 October 2020 with Recommendations to the Commission on a Framework of Ethical Aspects of Artificial Intelligence, Robotics and Related Technologies [2021] OJ C404/63

European Parliament Resolution 2015/2103(INL) of 17 February 2017 with Recommendations to the Commission on Civil Law Rules on Robotics [2018] OJ C252/239

European Commission, 'Proposal for a Regulation of the European Parliament and of the Council Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act) and Amending Certain Union Acts' COM (2021) 206 final

European Commission, 'Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions, "Towards an EU Criminal Policy: Ensuring the effective implementation of EU policies through criminal law"' COM (2011) 573 final

# National Legislation

## Germany

Bundesdatenschutzgesetz [BDSG] [Federal Data Protection Act], June 30, 2017, BGBl I at 2097, last amended by Art 10 of the Law of June 23, 2021, BGBl I at 1858
Bürgerliches Gesetzbuch [BGB] [Civil Code], August 18, 1896, RGBl at 195, revised January 2, 2002, BGBl I at 42, last amended by Art 4 of the Law of July 15, 2021, BGBl I at 1146
Straßenverkehrsgesetz [StVG] [Road Traffic Act], May 3, 1909, RGBl at 437, revised March 5, 2003, BGBl I at 310, last amended by Art 1 of the Law of July 12, 2021, BGBl I at 3108
Luftverkehrsgesetz [LuftVG][Aviation Act], August 1, 1922, RGBl at 681, revised May 10, 2007, BGBl I at 698, last amended by Art 131 of the Law of August 10, 2021, BGBl I at 3436
Tierschutzgesetz [TierSchG][Animal Welfare Act], 24 July, 1972, BGB1 I at 1277 , revised May 18, 2006 , BGB1 I at 1206, last amended by Art 105 of the Law of June 19, 2020, BGBl I at 3436,
Strafvollzugsgesetz [StVollzG][Prison Act], March 16, 1976, BGB1 at 581, last amended by Art 27 of the Law of October 5, 202, BGBl I at 4607
Strafgesetzbuch [StGB][Criminal Code], revised November 13, 1998, last amended by Article 1 of the law of July 11, 2022, BGBl at 1082)
Reichtierschutzgesetz,[Reich Animal Protection Act], November 24, 1933, RGB1 at 987

### Italy

Costituzione della Repubblica Italiana [Cost] [Constitution of the Italian Republic], GU 27 dicembre 1947, n 288
Codice Civile [CC] [Civil Code], Regio Decreto 16 marzo 1942, n 262, GU 4 aprile 1942, n 79
Codice Penale [CP][Penal Code], Regio Decreto 19 ottobre 1930, n 1398, GU 26 ottobre 1930, n 251
Legge, 12 febbraio 2019 n 12, GU 12 febbraio 2019, n 36
Legge, 20 luglio 2004, n 18, GU 31 luglio 2004, n 178

### Ireland

Constitution of Ireland 1937
Family Home Protection Act 1976
Fifteenth Amendment of the Constitution Act 1995
Animal Health and Welfare Act 2013
The Criminal Law (Insanity) Act 2006
National Minimum Wage Act 2000
Organisation of Working Time Act 1997

European Communities (Unfair Terms in Consumer Contracts) Regulations 1995, SI 1995/27

# PART A

# Background Concepts and Motivation

> For suppose that every tool we had could perform its task, either at our bidding or itself perceiving the need, and if -like the statues made by Daedalus or the tripods of Hephaestus, of which the poet says that 'self-moved they enter the assembly of the gods' -shuttles in a loom could fly to and fro and a plucker play a lyre of their own accord, then master craftsmen would have no need of servants nor masters of slaves.
>
> —Aristotle, *The Politics*[*]

[*] Aristotle, *The Politics* (Trevor J Saunders ed and T A Sinclair tr, Penguin 1981) 65.

# Chapter I: General Introduction

> God now said, 'Let us make human beings in our image, after our likeness; and let them hold sway over the fish of the sea and the birds of the sky, over the beasts, over all earth, over all that creeps upon the earth.' So, God created human beings in the [divine] image, creating them in the image of God, creating them male and female.
>
> —*Genesis* 1:26-27[1]

## I.1 Background of the Study

Just as God created human beings in the divine image -*in God's likeness*- according to the creation account in the book of Genesis, present-day robotics engineers are working on creations of their own-*robots*- that are *after our likeness*[2] in their abilities, behaviours, and sometimes other physical and cognitive attributes. Work in the field of robotics over the last six decades, itself culminated through thousands of years of intellectual and technological development, has been remarkably fruitful. Robots, once only existing as figures belonging to the imaginary worlds of mythology and legends[3] before being gradually conceptualised through the contributions of science fiction,[4] have been present in the real world in some shape or form since the mid-20th century.

Going back to the 1960s, the first so-called robots were nothing like the artificial, quasi-living entities portrayed in the popular imagination. On the contrary, they were immobile, bulky machines with no decision-making powers; they could only perform action sequences that were pre-programmed by human controllers.[5] Nonetheless, these proto-robots were still revolutionary inventions for the time, as the high deployment rates they reached facilitated the automation of multiple industries.

---

[1] In the present study, the version of the Bible used is *The Contemporary Torah: A Gender-Sensitive Adaptation of the Original JPS Translation* (David E Stein ed, JPS 2006).

[2] i.e., similar to human beings in specific ways.

[3] e.g., the automata of Hephaistos in Ancient Greek mythology or the Golem in Jewish folk tales.
See Ludwig Blau, Joseph Jacobs, and Judah David Eisenstein, 'Golem', *The Jewish Encyclopaedia*, (1904) vol 6, 36-37; Homer, *The Iliad* (Bernard Knox ed and Robert Fagles tr, Penguin 1990) 479.
See also Chapter III, Section III.2.1, 'Intellectual Origins: Mythological and Religious Accounts'.

[4] From Mary Shelley's *Frankenstein* and Karel Capek's *R.U.R.* and to numerous stories and novels by Isaac Asimov, human-made autonomous entities have been one of the popular elements of the science-fiction genre.
See Mary Shelley, *Frankenstein: The Modern Prometheus* (first published 1818, Simon & Schuster 2004); Karel Čapek, *R.U.R.: Rossum's Universal Robots* (David Wyllie tr, originally published 1920, eBooks@Adelaide 2014); Isaac Asimov, *Robot Visions* (Penguin 1991) 17-18.
See also Chapter III, Section III.2.3, 'Formal Origins: Contributions of Science Fiction'.

[5] The first so-called robot, Unimate, was an electronically controlled hydraulic heavy-lifting arm that could repeat pre-programmed motion sequences. Unimate was initially deployed in a die-casting factory and was used to remove and stack hot metal parts from the die-casting machine. Neil G Hockstein and others, 'A History of Robots: From Science Fiction to Surgical Robotics' (2007) 1 Journal of Robotic Surgery 113, 114.

Technological advances since then have made the development of progressively autonomous and increasingly sophisticated robots possible. Nevertheless, during the first few decades of robotics, the focus remained on designing robots that could undertake roles that were hindered or made impossible by human beings' sensitivities and weaknesses.[6] Consequently, most robots introduced until the 1990s were industrial manipulators. Though these early robots were set to replace human beings in the performance of dangerous, dirty, and dull tasks, they were by no means intended to be visible parts of everyday life.[7] Instead, they often operated behind the scenes, much like industrial robots of the present day, roaming in environments and conditions different from what we, as human beings, would consider habitable. Patrick Lin gives the following examples for the functions fulfilled by industrial robots:

> Military unmanned aerial vehicles (UAVs) surveil from the skies for far more hours than a human pilot can endure at a time. Robots crawl around in dark sewers, inspecting pipes and leaks for cracks. (…) Not afraid of danger, they also explore volcanoes and clean up contaminated sites.[8]

In addition to industrial robots that single-handedly dominated the first few decades of robotics, robots of another type -*service robots*- also have become increasingly available and gained lasting popularity since the late 1990s.

Unlike industrial robots, service robots are devised to work with human beings in shared spaces such as homes, schools, workplaces, and even urban streets. Naturally, the increasing use of service robots boosts the visibility of robots across the board. That boosted visibility, in turn, increases the demand for robots in the market and reveals robotics as one of the more profitable sectors. Ultimately, with the lasting popularity of service robots, they are expected also attract significant investments for developing new applications.[9] Nonetheless, most service robots are still devised to replace human labour in the performance of dull or monotonous assignments, such as basic housekeeping tasks. Hence, although they are closer to the popular image of robots compared to industrial robots, service robots are not categorically expected to show any humanlike attributes except for the characteristic of relative autonomy.

---

[6] Hans Peter Moravec, 'Robot', *Encyclopaedia Britannica* (4 January 2018).

[7] Marco Nørskov, 'Technological Dangers and the Potential of Human-Robot Interaction: A Philosophical Investigation of Fundamental Epistemological Mechanisms of Discrimination' in Marco Nørskov (ed), *Social Robots: Boundaries, Potential, Challenges* (Taylor & Francis 2018) 99.

[8] Patrick Lin, 'Introduction to Robot Ethics' in Patrick Lin, Keith Abney and George A Bekey (eds), *Robot Ethics: The Ethical and Social Implications of Robotics* (MIT 2014) 4.

[9] C W M Naastepad and Jesse M Mulder, 'Robots and Us: Towards an Economics of the Good Life" (2018) 76 Review of Social Economy 302, 305.

For the purposes of the present study, 'relative autonomy' denotes a robot's ability to function according to the sense-think-act paradigm. In contrast, 'artificial intelligence' refers solely to thinking capabilities – just one of the capabilities that a robot is required to possess to be identified as such.[10] Whereas artificial intelligence is essentially the brain of a robot and enables it to interpret its environment and learn from past experiences, it does not allow the robot to perceive or act on its environment.[11]

Since the early 2000s, roboticists have also produced numerous specialised service robots that can personally communicate, build rapport, and even negotiate with human beings. Rodney Brooks describes the first examples of these robots as follows:

> A number of robots that people have built, including Kismet and My Real Baby, are able to express emotions in humanlike ways. They use facial expressions, body posture, and prosody in their voices to express the state of their internal emotions. Their internal emotions are a complex interplay of many subsystems.[12]

These specialised service robots have additional capabilities enabling them to recognise, react to, and sometimes even mimic human beings' emotive expressions and social actions, and they are intended to create more interactive and personalized experiences for human beings.[13] Because of the additional capabilities allowing them to interact with human beings on personal levels, these specialised service robots are distinguished from other robots, and are described with adjectives such as 'social',[14] 'sociable',[15] and 'socially interactive'[16] across various resources.

---

[10] Arto Laitinen and Otto Sahlgren, 'AI Systems and Respect for Human Autonomy' (2021) 4 Frontiers in Artificial Intelligence.

[11] Teresa Zielinska, 'Professional and Personal Service Robots' (2016) 4 International Journal of Robotics Applications and Technologies 63, 65; Juan Angel Gonzalez-Aguirre and others, 'Service Robots: Trends and Technology' (2021) 11 Applied Sciences 1, 11-13.
See also Chapter III, Section III.3.1.a, 'Sense-Think-Act Paradigm'.

[12] Rodney A Brooks, *Flesh and Machines: How Robots Will Change Us* (Pantheon 2002) 155.

[13] Kirsten Weir, 'The Dawn of Social Robots: Roboticists Must Dip Heavily into Psychological Science' (2019) 49 Monitor on Psychology 50.

[14] See, e.g., Simon Couglan and others, 'Could Social Robots Make Us Kinder or Crueller to Humans and Animals?' (2019) 11 International Journal of Social Robotics 741, 741; Riccardo Campa, 'Social Robots: A Bridge Between the Two Cultures' (2021) 5 KnE Social Sciences; Anna Henschel, Guy Laban, and Emily S Cross, 'What Makes a Robot Social? A Review of Social Robots from Science Fiction to a Home or Hospital Near You' (2021) 2 Current Robotics Reports 9, 11; John P Ulhøi and Sladjana Nørskov, 'The Emergence of Social Robots: Adding Physicality and Agency to Technology' (2022) 65 Journal of Engineering and Technology Management.

[15] See, e.g., Cynthia Breazeal, *Designing Sociable Robots* (MIT 2002) 1; Cory D Kidd, Will Taggart, and Sherry Turkle, 'A Sociable Robot to Encourage Social Interaction Among the Elderly' (IEEE International Conference on Robotics and Automation, Orlando, May 2006); Oscar Déniz and others, 'An Engineering Approach to Sociable Robots' (2007) 19 Journal of Experimental & Theoretical Artificial Intelligence 285, 287; Glenda Shaw-Garlock, 'Looking Forward to Sociable Robots' (2009) 1 International Journal of Social Robotics 249, 250.

[16] See, e.g., Kerstin Dautenhahn, Terrence Fong and Illah Nourbakhsh, 'A Survey of Socially Interactive Robots' (2003) 42 Robotics and Autonomous Systems 143, 144; Eric Deng, Bilge Mutlu, and Maja J Mataric, 'Embodiment in Socially Interactive Robots' (2019) 7 Foundations and Trends in Robotics 251, 253.

The present study refers to these specialised service robots as *social robots.* The reason is that the adjective *social* refers to specific additional capabilities of these robots without focusing too much on the roles they undertake. It is assessed that the term *social robot* includes robots whose functions are not merely limited to socialising while still having the capabilities required to engage human beings in socially meaningful ways. Notwithstanding what can be described as social robots' 'people skills', they are neither human beings, nor even living entities. The fact that they are designed to emulate human social behaviours does not change the nature of social robots as functional artefacts. As human-made, inanimate contraptions, social robots do not suffer from various health conditions that affect human beings, nor can they go through mood swings - they are also incapable of getting angry, bored, confused, disgruntled, or even tired.

Because of the functional combination of mechanical efficiency with humanlike sociability, social robots are gradually emerging as chosen partners for numerous social interactions. In effect, they are slowly entering into social interaction spheres as independent social actors.[17] The following quote from a human participant who interacted with a social robot in the context of a psychological experiment illustrates the influence of the specific capabilities of these robots on the human psyche:

> Although he's a robot, but you begin to like him. You know, you begin to have a bit of friendship, affection. Because, you know, if he can respond to you then you feel that he is alive, so you don't want to have any aggressiveness with him at all.[18]

With the increasing numbers of social robots and in light of the constantly evolving abilities they display, it is no longer 'futuristic' to talk about robots as companions, confidants, and even lovers to human beings. Suppose the shift in the general perception of robots in society from functional artefacts operated by human beings to independent social actors is taken to represent 'the robotic moment',[19] as termed by prominent psychologist Sherry Turkle. In that case, it can be defended that humankind has already entered that moment.[20] The advent of social robots as independent social actors is poised to challenge the law's function in helping preserve the social order since the law does not acknowledge the legal relevance of social robots' distinctive characteristics.

---

[17] Mark Ingebretsen, 'Where's My Personal Robot?' (2009) 24 IEEE Intelligent Systems 90, 91.
[18] Katie A Riddoch and Emily S Cross, '"Hit the Robot on the Head with This Mallet" - Making a Case for Including More Open Questions in HRI Research' (2021) 8 Frontiers in Robotics and AI, 9.
[19] Sherry Turkle, *Alone Together: Why We Expect More from Technology and Less from Each Other* (Basic 2012) 9; Sherry Turkle, 'The 'Robotic Moment'', *Encyclopaedia Britannica* (25 May 2018).
[20] Marco Nørskov, 'Editor's Preface' in Marco Nørskov (ed), *Social Robots: Boundaries, Potential, Challenges* (Taylor & Francis 2018) xv.

Thanks to the proliferation of robots, concerns about the law's effectiveness can no longer be dismissed as imagined fears or futuristic paranoias. Presently, state-of-the-art social robots have sufficient degrees of social agency to confront the self-understanding of human beings with whom they interact. Moreover, they are already autonomous enough to jeopardise the adequacy of existing legal instruments that are organised around the quirks of human nature.[21]The present study focuses on developing feasible and practical solutions to the legal challenges likely to arise from human-robot social interactions.

Especially over the last decade, the legal implications of robots' characteristics have been extensively discussed by scholarly circles, and these discussions have found that the root causes of potential issues are related to the inadequacy of existing legal instruments in upholding the underlying legal principles when faced with the distinctive characteristics of social robots.[22] Consequently, the legal issues regarding social robots can be summarised in the form of two research questions that are addressed through the methodology followed by the present study.[23]

(1) Who should bear legal responsibility for the independent activities of robots,[24] facilitated by the capabilities associated with the characteristic of relative autonomy?[25]

---

[21] The adequacy of legal instruments is determined on a case-by-case basis. If the given legal instrument upholds the underlying legal principle (that determines its purpose) when applied to the facts of some specific case, it can be regarded as adequate.

[22] In the present study, the *distinctive characteristics of social robots* refer to the traits that separate robots from other functional artefacts. These traits are explained in detail later in the present study.
See Chapter III, Section III.4, 'Characteristics of Social Robots'.

[23] The research methodology followed in the present study is described in the next chapter.
See Chapter II, 'Methodology of the Study'.

[24] i.e., civil and/or criminal liability for the undesirable outcomes of robots' autonomous and unpredictable behaviours.

[25] On the assumption that any given robot should be able to execute at least one of its operational processes independently, the present study takes it that all robots have capabilities empowering them not just to act on their own but also to plan and decide which activities to perform. Devices which can only execute pre-programmed action sequences require constant human control to complete their tasks; they have no decision-making powers of their own to guide their goal-directed behaviours, and therefore cannot be regarded as robots
See, e.g, Matthew E Gladden, 'The Diffuse Intelligent Other: An Ontology of Nonlocalizable Robots as Moral and Legal Actors' in Marco Nørskov (ed), *Social Robots: Boundaries, Potential, Challenges* (Taylor & Francis 2018) 177. Therefore, the present study distinguishes the notion of automation from that of autonomy and conceptualises the capabilities shared by all robots as the characteristic of relative autonomy. As1 the distinctive characteristics of social robots include the attributes that set robots apart from the rest of the functional artefacts, the characteristic of relative autonomy is categorised as one of them.
See also Chapter III, Section III.4.1, 'Characteristic of Relative Autonomy'

(2) Should social robots -robots that have the specific capabilities associated with the characteristic social agency- be approached as more than mere functional artefacts that simulate roles to the benefit of human beings? [26] From the legal perspective, how should they be treated? [27]

In case law, including some judgments of the Irish courts,[28] the word *robot* appears to have been used in the metaphorical sense to describe human beings who are tenacious but do not have any ability to exercise discretion or show flexibility. Ryan Calo, one of the experts on the law of robots, explains the metaphorical use of the word *robot* in court cases as follows:

> The metaphor of the robot appears as shorthand for a person without will. In the judicial imagination, a robot is what a person or entity becomes when completely controlled by another. Such a person or entity is not capable of fault or knowledge, leaving the individual controlling the machine -the programmer- at fault instead.[29]

The meaning attributed to the word *robot* in court cases suggests that courts tend to view robots as devices that cannot act on their own accord.[30]

---

[26] Because of the social interactivity facilitated by their additional capabilities, social robots can invoke unprecedented levels of anthropomorphic responses from human beings. According to Matsuzaki, with the advent of social robots, the social world can no longer be seen as exclusively consisting of interpersonal interactions between human beings. The present study associates the additional capabilities of social robots with the characteristic of social agency as these specific capabilities allow social robots to be situated in society as independent actors.
See Hironori Matsuzaki, 'Robots, Humans, and the Borders of the Social World' in Marco Nørskov (ed), *Social Robots: Boundaries, Potential, Challenges* (Taylor & Francis 2018) 157.
See also Chapter III, Section III.4.1, 'Characteristic of Social Agency'.

[27] The second question sums up the challenges concerning the legal implications of the teleological anthropomorphism of social robots. The literature suggests that the strength of anthropomorphic responses invoked by social robots might lead to the desensitisation of society against violent behaviours toward human beings and other living organisms. However, according to Darling, such unwelcome developments can be circumvented through the legal protection of social robots.
See Kate Darling, 'Extending Legal Protection to Social Robots: The Effects of Anthropomorphism, Empathy, and Violent Behavior Towards Robotic Objects' in Ryan Calo, A Michael Froomkin and Ian Kerr (eds), *Robot Law* (Edward Elgar 2016) 220.

[28] e.g., in *Rohan Construction* case, the High Court used the word *robot* to refer to human beings who are not capable of showing discernment or good judgement: 'No matter at what level of the hierarchy a person may be employed, he does not become a mere robot: whether he be a highly qualified engineer, a tradesman or even an unskilled labourer, he is expected, in addition to any specialist qualifications he may have, to use his own common sense and experience in relation to the task he is doing.' *Rohan Construction Ltd v Insurance Corporation of Ireland plc*[1986] WJSC-HC 1406, [1986] ILRM 419.
Similarly, in *Donnelly* case, the word *robot* is used to indicate human beings who cannot respond with intuitive, personal reactions: 'It is a well-known fact that defendants who conduct a defence along the lines of suggesting that a plaintiff is not genuine must expect a possible reaction either where the case is tried by a judge and jury or by a judge alone. A trial judge is no more a robot than a member of a jury.' *Donnelly v Timber Factors Ltd* [1991] WJSC-SC 372, [1991] 1 IR 553, 559.

[29] Ryan Calo, 'Robots as Legal Metaphors' (2016) 30 Harvard Journal of Law & Technology 209, 226.
Though Calo researches US court decisions, his conclusions equally apply to Irish courts' decisions. Two of the decisions of Irish courts are summarised in the footnote (n 28) above.

[30] A Michael Froomkin, 'Introduction' in Ryan Calo, A Michael Froomkin and Ian Kerr (eds), *Robot Law* (Edward Elgar 2016) xiv.

Calo affirms that courts have consistently categorised robots as 'extensions of [those] who set them into motion' thus far.[31] Indeed, American courts have repeatedly characterised robots as programmable devices that are 'incapable of spontaneity'.[32] For example, in *Frye v Baskin*, the Missouri Court of Appeals held that robots cannot act independently and, therefore, they must be seen as extensions of those who control them.[33] Likewise, in *Comptroller v Family Entertainment Centers*, the Baltimore County Circuit Court stated that,

> [The] robot can perform a menial task but, because (...) [it] has no 'skill' and therefore leaves no room for spontaneous human flaw in an exhibition, it cannot 'perform' a piece of music.[34]

Such judgments accurately observe that robots are devices conceived to perform specific roles. Nonetheless, since the present-day robots can perform behaviours that go beyond what is intended or could be expected by those who develop, deploy, and interact with them, categorising robots in the same group as devices that require continuous human supervision, as suggested by these judgements, does not align with reality anymore Given that social robots -whose behaviours are expected to be frequently unpredictable- are becoming common only recently, it is expected that the outdated approaches toward robots in court cases will be subject to scrutiny soon.

Some concerns related to the roboticization of society have already compelled lawmakers in some countries to introduce new legal instruments to ensure that human beings and robots can interact safely. These rules and regulations demonstrate that lawmakers' approaches to robots are better aligned with reality than the approaches espoused in the judicial opinions described above. Nonetheless, at national law-making levels, most legal instruments about human-robot interactions are enacted in response to particular concerns about specific types of robots, such as licensing rules for surgical robots or traffic rules for self-driving cars.[35] They, therefore, are technology-specific.

---

[31] Ryan Calo, 'Robots in American Law' in Eric Hilgendorf and Uwa Seidel (eds), *Robotics, Autonomics, and the Law* (Nomos 2017) 108.

[32] For the survey of relevant US court decisions, see ibid 65-95.

[33] In the case of *Frye v Baskin*, the plaintiff owned a jeep that he had taught his son to drive. The plaintiff's son was on a date with the defendant and asked her to take the wheel. The defendant did not know how to drive, but she drove under the tutelage of the plaintiff's son. At one point, the plaintiff's son called out a direction to the defendant and immediately rinvoked that instruction. The defendant tried to comply but wound up running the jeep into a ditch. The Court decided that 'as far as [the plaintiff's son] was concerned, [the defendant] controlled the car the same as if she had been a robot or an automaton. When [the plaintiff's son] said 'turn', she turned, mechanically, she was the instrumentality by which [the plaintiff' drove the car' and 'if it were negligence, it was [the plaintiff's son]'s and not hers'. *Frye v Baskin* 231 SW2d 630 (1950) 632-634.

[34] *Comptroller v Family Entertainment Centers*, 519 A 2d 1337 (Md Ct Spec App 1987) 1339.

[35] e.g., the NCSL database provides the enacted legislation on self-driving cars in the US; and several countries within the EU have enacted legislation on the same issue.

Being technology-specific does not necessarily indicate that these rules and regulations lack detail. However, it suggests that legal instruments introduced at national levels cannot bring future-proof solutions to the challenges arising from the characteristics of robots per se.

Beyond national law-making, some supranational legal instruments resist the trend toward technology-specific regulation. The European Commission's ['EC'] Proposal for the Regulation of Artificial Intelligence ['EU AI Act'],[36] due to enter into force in early 2023, does not explicitly use the word *robot*. Nevertheless, the Act's definition of *artificial intelligence* ['AI'] applies to the decision-making systems of robots which enable them to operate independently:

> software that (…) can, for a given set of human-defined objectives, generate outputs such as (…) recommendations, or decisions influencing the environments they interact with[37]

The provisions of the EU AI Act focus on the duties and obligations of developers and deployers of AI applications. The Act establishes the regulatory framework governing the development and oversight of AI applications by manufacturers and (often) institutional users. However, the EU AI Act does not address interactions between AI applications and human interactants (end-users) once these applications are set into motion. Thus, though the Act diverges from the trend of technology-specific regulation and manages to bring future-proof rules, on closer examination, it appears to deliver little more than some form of product safety framework that incorporates *ex-ante* regulation to mitigate a few select threats to information security and privacy.[38] Since the EU AI Act avoids answering the question of *ex post* accountability for the undesirable outcomes of the independent activities of artificial intelligence applications (including robots), the present study concludes that it cannot clarify the EU's position on the legal issues surrounding human-robot social interactions.[39]

---

See Benedikt Münch, *Legal Questions with Autonomous Cars* (Akademikerverlag 2014) 10-25; Douglas Shinkle, 'Self-Driving Vehicles, Enacted Legislation' (*NCSL*, 27 August 2018); Veronica DeVore and Duc-Quang Nguyen, 'Which Countries are Testing Driverless Cars?' (*SWI swissinfo.ch,* 4 March 2019).

[36] European Commission, 'Proposal for a Regulation of the European Parliament and of the Council Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act) and Amending Certain Union Acts' COM (2021) 206 final (EU AI Act).

[37] ibid, Art 3 (1).

[38] ibid, Arts 8-15.

[39] Frederik Z Borgesius and Michael Veal, 'Demystifying the Draft EU Artificial Intelligence Act-Analysing the Good, the Bad, and the Unclear Elements of the Proposed Approach' (2021) 22 Computer Law Review International 97, 103.

Significantly, several inquiries about the law's treatment of robots have already been brought forward by the European Parliament ['EP'] in multiple resolutions, first in 2017,[40] and then again in 2020[41], where the EC was asked to examine the subject of civil liability for the damages caused by independent activities of robots. In the 2017 resolution calling for *Civil Law Rules on Robotics*, the EP poetically expressed the increasingly compelling need for EU-level legislative action as follows:

> Whereas now that humankind stands on the threshold of an era when ever more sophisticated robots, bots, androids, and other manifestations of artificial intelligence seem to be poised to unleash a new industrial revolution, which is likely to leave no stratum of society untouched, it is vitally important for the legislature to consider its legal and ethical implications and effects, without stifling innovation. [42]

In the follow-up to these resolutions, the EC announced that new legal instruments for harmonising civil liability rules about human-robot interactions across the EU countries were on the horizon. Nonetheless, the scope and content of these upcoming instruments are still unclear.[43]

Emerging legal instruments, such as the EU AI Act, might successfully address some legal challenges surrounding human-robot social interactions, especially the threats to information security and privacy. However, like national rules and regulations on the same subject matter, EU solutions also appear to presuppose that robots are advanced but potentially dangerous functional artefacts that should be operated with the highest levels of attention and care. That presupposition fails to acknowledge what robots have become thanks to the distinctive characteristics they display: artificial autonomous agents.

The capabilities associated with the characteristic of relative autonomy are liable to have significant personal, economic, and social impacts on their own. However, both the national and supranational legal systems are still exclusively focused on the activities of individual human beings or groups of human beings with shared interests or pursuits ['human associations'].[44]

---

[40] European Parliament Resolution 2015/2103(INL) of 17 February 2017 with Recommendations to the Commission on Civil Law Rules on Robotics [2018] OJ C252/239.

[41] European Parliament Resolution 2020/2014(INL) of 20 October 2020 with Recommendations to the Commission on a Civil Liability Regime for Artificial Intelligence [2021] OJ C404/107.

[42] EP Resolution 2015/2103(INL) (n 40) para B.

[43] The EC has started an initiative to introduce 'EU rules to address liability issues related to new technologies, including AI systems', but that initiative has not come to fruition yet.
See European Commission, 'A European Approach to Artificial Intelligence' (Europe's Digital Future, 7 June 2022).
See also Dimitar Lilkov, 'Regulating Artificial Intelligence In the EU: A Risky Game' (2021) 20 European View 166, 167.

[44] Groups of human beings with shared interests or pursuits include corporations, associations, labour unions, political parties, government agencies and other juridical persons.

As there have never been any non-human entities whose independent activities could have delivered socially significant outcomes before the emergence of social robots as independent social actors, lawmakers' reluctance to recognise the characteristics of robots is understandable – but misguided. Robots are newcomers to the equation of social order in which the law functions to balance the competing individual interests of various independent social actors. Like the free will of human beings, the capabilities of relative autonomy allow robots to act independently in the sense of planning and executing the courses of action they follow. Though they all possess autonomous decision-making powers, and some robots even demonstrate empathic perspective-taking skills, each decision made, and every interaction performed by robots are ultimately guided by the objectives set by human beings who manufacture, use, and engage them. By definition, no robot has -and none can ever have- any needs, wants, desires, goals, and objectives, or, more broadly, any *individual interests* that do not stem from the nature of its designated role or assigned task.[45] In that regard, the autonomy of robots can be considered to be relative to the individual interests of their owners, users, or manufacturers.

Conflicts have always been regarded as part and parcel of social life, as every individual member of society is driven by separate, and potentially divergent sets of individual interests.[46] Considering that social robots, as emergent social actors, also pursue different sets of individual interests (some of which may conflict with those pursued by others), their growing numbers can be expected to make conflicts of interest arising from human-robot social interactions commonplace. Though *ex ante* rules and regulations can help avoid some of these conflicts, as social robots are autonomous agents that function in unpredictable environments, it must be admitted not all conflicts can be circumvented. Some conflicts related to human-robot social interactions are bound to happen regardless of what *ex ante* rules and regulations are adopted, and such conflicts can only be addressed after the fact. Given the legal system's role in resolving the undesirable outcomes of these new conflicts of interest, developing a new accountability framework that considers the distinctive characteristics of social robots -such as relative autonomy- is becoming increasingly urgent.

[45] Rudolf von Jhering, *Law as a Means to an End* (Isaac Husik tr, Boston Book Company 1913) 47-52.
[46] Bryan McMahon and William Binchy, *Law of Torts* (4th edn, Bloomsbury 2013) para 1.01.

## I.2 Scope of the Study

The current landscape and prospects of human-robot social interactions expose the inadequacies of existing legal instruments. The present study aims to identify, describe, and ultimately offer suggestions for resolving the legal issues exposed by the advent of robots as social actors. Therefore, the scope of the study is primarily defined in terms of which types of interactions are addressed.

### I.2.1 Legal Issues

Though human-robot social interactions are still up-and-coming, they are already projected to have extensive personal, economic, and social effects. If the function of the law is to maintain social order by resolving the conflicts of interests in society, most of these effects have the potential of becoming legally relevant. Then, the accurate identification and adequate resolution of the legal issues surrounding human-robot social interactions necessitate the interdisciplinary examination of these interactions. The present study, therefore, uses of the disciplines of economics, social and political philosophy, and social psychology.

Despite the use of other disciplines, the study ultimately analyses the law, and focuses on developing a model for overcoming legal systems' shortcomings in dealing with the issues arising from human-robot interactions.

Though it is acknowledged that the ongoing proliferation of robots can result in socio-economic and cultural challenges beyond the inadequacies of existing legal instruments, such challenges require the adoption of certain situational goals, and therefore remain outside the study's scope. The separation of legal theory and social policy realms, followed by the present study is inspired by Ronald Dworkin's distinction between the matters of *principle* and *policy*. Dworkin argues that the legal principles are concerned with individual rights and directed toward the pursuit of justice, equity, or other moral values, whereas policies are developed to determine specific cultural, economic, or political objectives for society:

> Arguments from principle are arguments which are supposed to justify a right; arguments from policy are arguments which are supposed to justify some collective end. Principles are propositions describing rights; policies are propositions describing ends. [47]

[47] Ronald Dworkin, *Taking Rights Seriously* (Harvard UP 1978) 22

While policy-making activities focus on determining collective situational goals for the future of society, law-making and adjudication activities focus on the justification of rights and obligations through underlying legal principles.[48] By excluding social policy issues, the present study confirms that its purpose is to offer solutions based on the legal principles for the inadequacies of existing legal instruments rather than recommending new objectives to be followed by administrations to overcome the myriad of socio-economic problems that may have little significance from the legal perspective.

Examples of social policy challenges that cannot be overcome by working with the legal principles and therefore lay outside the scope of the study include the increasing unemployment rate, especially among unqualified workers,[49] or the environmental impacts of increased production and use of robots.[50]

### I.2.2 Everyday Contexts

The present study's focus is limited to the legal issues surrounding social interactions that are likely to occur in the *ordinary course of everyday life.* Interactions that transpire in contexts where the interacting parties are not holding equal positions[51] or where the normal operation of laws is suspended for whatever reason[52] are presumed to be outside the ordinary course of everyday life and therefore excluded from the scope of this study.

---

[48] Differentiating policy-making and law-making activities may not always be possible in real life since most policy-making activities introduce or incorporate elements of principles through references to individual rights. ibid, 90.
See also Michel Rosenfeld, 'Dworkin and the One Law Principle: A Pluralist Critique' (2005) 3 (233) Revue Internationale de Philosophie 363, 371.

[49] Carl Benedict Frey and Michael Osborne, 'The Future of Innovation and Employment' (Report, Citi GPS 2015) 105; Florent Bordot, 'Artificial Intelligence, Robots and Unemployment: Evidence from OECD Countries' (2022) 37 Journal of Innovation Economics & Management 117, 133.

[50] Rupert Read, 'Rise of the Robots will Harm the Earth as Well as Human Beings' (*The Guardian,* 27 March 2016); Fiachra O'Brolcháin and María Amparo Grau Ruiz, 'Environmental Impact of Robotics: Ethical Concerns and Legal Alternatives', *Encyclopaedia of the UN Sustainable Development Goals* (2021).

[51] i.e., where one party exercises any power that cannot be exercised by private citizens, such as the powers of law enforcement officers. Since no one would normally enter into any relationship where they are not in equal positions with the other party unless they are compelled to, these situations cannot be considered part of the ordinary course of everyday life. Moreover, since the party who is exercising superior power is doing so on behalf of the state, they are not usually influenced by the interaction themselves, diminishing the social character of that interaction.

[52] i.e., where a state of exception, state of emergency, or martial law is declared, and where special laws are operative such as in wartime between belligerents. All situations where the normal laws are suspended, regardless of whether it happens legitimately or *de facto* because of the loss of state authority, are out of the ordinary course of everyday life.

### I.2.2.a Law Enforcement Robots

The present study does not address the issues arising from human-robot interactions that occur in the context of law enforcement since such interactions concern the use of law enforcement powers, which puts interacting parties in unequal positions. On the assumption that no one would enter into interactions where the other party can limit the exercise of their fundamental rights unless they are compelled to or required to do so, it is evaluated that such issues do not arise as part of the ordinary course of everyday life. Though not addressed in the present study, using robots in law enforcement raises some significant legal concerns. In the literature, the employment of robots to perform primary law enforcement tasks[53] is projected to have grave individual and social consequences.[54] It is assessed that legal structures that are currently in place (the way laws are designed, communicated, and enforced in the present) cannot accommodate using robots in law enforcement for anything other than secondary tasks.[55] Most laws are created with the assumption that they will be broken from time to time. Some are even designed to have their effectiveness increased by certain amounts of non-compliance.[56] Consequently, the ability to exercise discretion, in the sense of making exceptions and showing lenience, is considered necessary for discerning the course of conduct prescribed by even 'seemingly simple laws', and significant social harms are expected to occur if even the most minor infractions face the full brunt of the law.[57]

In the current state of the art, no robot can exercise discretionary judgement, show leniency, or make exceptions. These incapacities are the source of the legal concerns surrounding the use of robots in law enforcement. Indeed, even without any algorithmic malfunctions, the enforcement of laws with robotic efficiency might result in several unwelcome outcomes.

---

[53] Primary law enforcement tasks include those that could interfere with fundamental rights, such as conducting personal and premise searches, making arrests, questioning suspects, and using force, as well as exercising judicial functions.

[54] Lisa A Shay and others, 'Confronting Automated Law Enforcement' in Ryan Calo, A Michael Froomkin and Ian Kerr (eds), *Robot Law* (Edward Elgar 2016) 236-239.

[55] Secondary law enforcement tasks include monitoring both real-world and cyber security threats, reporting suspicious activities, and assisting primary law enforcement officers in collecting and analysing evidence. Robots already perform these tasks. Michael R McGuire, 'The Laughing Policebot: Automation and The End of Policing' (2020) 31 (1) Policing and Society 20, 22.

[56] Michigan Law Review, 'Laws that are Made to be Broken: Adjusting for Anticipated Noncompliance' (1977) 75 (4) Michigan Law Review 687, 688-690.

[57] Nathan Isaacs, 'The Limits of Judicial Discretion' (1923) 32 (4) The Yale Law Journal 339, 342.

For example, robots employed in law enforcement may issue fines even when the speed limit is exceeded by only one kilometre per hour[58] or fail to show leniency in emergency situations, like rushing a seriously ill patient to the hospital.[59]

### I.2.2.b Killer Robots

Killer robots (i.e., lethal autonomous weapons systems) are functional artefacts designed to identify, select, and kill targets without any human intervention. Since they are exclusively devised for belligerent purposes, the use (though not necessarily the manufacturing) of killer robots is expected to fall under the purview of the law of armed conflicts.[60] The law of armed conflicts applies in times of war, or situations that involve violent armed struggles. In these situations, some of the guarantees provided by the rule of law are withdrawn or can no longer be provided, as illustrated by the aphorism *inter arma enim silent leges* (in times of war, the law falls silent).[61] Since these situations involve the suspension of the normal operation of laws, it is concurred that all types of armed conflicts are out of the ordinary course of everyday life from the legal perspective. Therefore, for the present study, it is assumed that no interaction between human beings and killer robots can be categorised as 'social'.

Moreover, regardless of what degree of autonomy killer robots display, all of them are ultimately conceived as weapons systems. Simply put, they are devices that are used for no other purpose other than to cause damage, injure, or kill living entities.[62] Because of the limited uses attached to them, the activities of killer robots can most accurately be predicted from the design stance, using the knowledge on their functionality and design. However, autonomous behaviours of the rest of the robots are most precisely examined from the intentional stance, based on the workings of their decision-making and evaluative processes.[63]

---

[58] Shay and others, 'Confronting Automated Law Enforcement' (n 54) 237.
[59] Lisa A Shay and others, 'Do Robots Dream of Electric Laws? An Experiment in the Law as Algorithm' in Ryan Calo, A Michael Froomkin and Ian Kerr (eds), *Robot Law* (Edward Elgar 2016) 279.
[60] Peter Asaro, 'Jus Nascendi, Robotic Weapons and the Martens Clause' in Ryan Calo, A Michael Froomkin and Ian Kerr (eds), *Robot Law* (Edward Elgar 2016) 386.
[61] The researcher cannot think of a better example of situations that are outside of ordinary course of everyday life than armed conflicts.
See Yael Ronen, 'Silent Enim Leges Inter Arma– But Beware the Background Noise: Domestic Courts as Agents of Development of the Law on the Conduct of Hostilities' (2013) 26 Leiden Journal of International Law 599, 601.
[62] Noel Sharkey, 'Killing Made Easy: From Joysticks to Politics' in Patrick Lin, Keith Abney and George A Bekey (eds), *Robot Ethics: The Ethical and Social Implications of Robotics* (MIT 2014) 115.
[63] Daniel C Dennett, *The Intentional Stance* (MIT 1998) 47.

To sum up, though some reforms may be needed in the law of armed conflicts to respond to legal issues surrounding the use of killer robots, these reforms do not require the treatment of killer robots as social interaction partners to human beings, and since these robots do not qualify as social interaction partners, their exclusion from the scope of this study is justified.[64]

### I.2.3 Social Interactions

The present study examines the legal issues arising from social interactions. *Social interaction* is defined as 'the process of reciprocal influence exercised by individuals over one another'.[65] The process of social interaction is facilitated by the exchange of social actions between Individuals, otherwise independent entities. Social actions, in turn, refer to the voluntary behaviours of individuals that have tangible outcomes on the mental states and subsequent activities of others. Social relations between individuals, building blocks of society, are formed through social interactions. For that reason, in the present study, social interactions are approached as socially formative connections between individual members of society:

> To recall (…) John Donne, no one is an island. This means that all individuals, except those who choose to live truly alone, interact with other individuals virtually every day and often many times in any one day. For social order, a prerequisite for any society, to be possible, effective social interaction must be possible.[66]

Because of the connective functionality of social interactions for the preservation of society, entities that can participate in these interactions -individuals- are expected to have relational independence and demonstrate some veneer of selfhood. In other words, their actions should be regarded by the individuals they interact with to be unique in the sense of being shaped by their accumulated knowledge, authentic experiences, and subjective observations. For social robots, relational independence is facilitated by the capabilities associated with the characteristic relative autonomy, and the additional capabilities associated with social agency provide the veneer of selfhood.

---

[64] Paul Bello and Marcello Guarini, 'Robotic Warfare: Some Challenges in Moving from Non-Civilian to Civilian Theaters' in Patrick Lin, Keith Abney and George A Bekey (eds), *Robot Ethics: The Ethical and Social Implications of Robotics* (MIT 2014) 139.
[65] William Little, *Introduction to Sociology: 2nd Canadian Edition* (BC Open Textbooks 2016) 946.
[66] Justus Macias, *Education in Emerging Global Society* (ED-Tech 2021) 244

The characteristic of relative autonomy incorporates the capabilities that are shared by all robots. These capabilities allow robots to develop instrumental and relational awareness of their surroundings. These capabilities are dependent on the integrated evaluative processes of robots, and they are comparable to the skills associated with human intelligence, such as elemental analysis and logical reasoning.[67] Through the integrated evaluative processes, robots can make decisions on their own and respond to changes in the environments they operate in, and even draw lessons from the outcomes of the past decisions. Since the capabilities associated with the characteristic of relative autonomy enable robots to compare potential courses of action and plan the course that would allow them to fulfil whatever functions they are assigned most effectively, it is submitted that robots can perform goal-directed or purposeful behaviours.[68]

The characteristic of social agency refers to the additional, specific capabilities that situate social robots as independent actors in society -individuals- with unique personalities. These additional capabilities enable social robots to distinguish other social entities, interpret their perceptions in light of their past experiences, and communicate with and learn from others on personal levels. In the literature, the dominant argument is that social interactions require physical proximity between interaction partners, and many authors uphold the view that robots that undertake roles as social interaction partners must have physical embodiments.[69]

However, with the prevalence of the Internet as the medium where most of personal exchanges and commercial transactions occur, that view appears to be no longer valid. In the present time, digital social encounters allow interaction partners to influence each other's mental states just as face-to-face interactions do,[70] and online transactions have the same legal force as their physical-world counterparts.[71]

---

[67] Barbara Bottalico and Amedeo Santosuosso, 'Autonomous Systems and Law: Why Intelligence Matters, A European Perspective' in Eric Hilgendorf and Uwa Seidel (eds), *Robotics, Autonomics, and the Law* (Nomos 2017) 28.

[68] Sypros G Tzafestas, *Roboethics: A Navigating Overview* (Springer 2016) 36; Christopher Collins and others, 'Artificial Intelligence in Information Systems Research: A Systematic Literature Review and Research Agenda' (2021) 60 International Journal of Information Management 1, 2.

[69] Yanfeng Lu and Weijie Zhao, 'What will Robots be Like in the Future?' (2019) 6 National Science Review 1059, 1059-1061; Maartje M A De Graaf, Frank A Hindriks and Koen V Hindriks, 'Who Wants to Grant Robots Rights?' (2022) 8 Frontiers in Robotics and AI, 3.

[70] Digital social encounters refer to the social interactions that take place through social media, i.e., interactive technologies that allow people to communicate and share information, ideas, interests using the Internet. Little (n 65) 946-948; Paul S N Lee and others, 'Internet Communication versus Face-to-Face Interaction in Quality of Life' (2011) 100 (3) Social Indicators Research 375, 388.

[71] Zaryn Dentzel, 'How the Internet Has Changed Everyday Life', in *19 Key Essays on How Internet is Changing Our Lives* (BBVA 2014) 241.

Then, in the Internet era, physical proximity between individuals can no longer be viewed as one of the necessary conditions for social interactions to take place. In fact, thanks to the affordances of the Internet, social robots with no physical embodiments (such as emotional chatbots) can also undertake roles as social interaction partners. Thus, the legal issues surrounding human interactions with non-physical social robots are also included within the scope of the present study.

Furthermore, the present study highlights that social interactions do not always have to transpire per the parties' mutual intents. On the contrary, in the ordinary course of everyday life, most of the reciprocal influence that individuals exert on one another's states of mind and subsequent activities either stem from unilateral intents or occur accidentally. For example, some criminal offences, such as assault, rape, harassment, robbery, or threats to kill, are considered social interactions. The commission of these offences has impacts on the states of mind and subsequent activities of both the offenders and victims simultaneously, though there is no intent to engage in any social interaction on the part of the victims.[72] Likewise, accidents that result in damage are also regarded as social interactions since they impact both parties' mental states and ensuing behaviours. However, neither the wrongdoer nor the injured party intends such impacts to occur.[73] For *accidental* social interactions to occur, parties involved do not even need to acknowledge each other as social entities. Moreover, courts can impose specific measures on the parties when the unwelcome outcomes of unilateral or accidental social interactions are brought before them. The remedial obligations and criminal punishments both require those on whom they are imposed to perform specific behaviours and, thereby, have consequences on the parties' states of mind and subsequent activities.

Considering the broad range of societal roles that robots are set to undertake, it is expected that only some social interactions between human beings and robots will be based on mutual intent of both parties, such as contractual transactions concluded, or emotionally meaningful interactions performed with robots. Most human-robot social interactions will likely be accidental or unilateral, such as road accidents involving self-driving cars, robots getting attacked while performing their everyday tasks, or criminally offensive behaviours toward robots.

---

[72] Edward Glaeser, Bruce Sacerdote, and Jose Scheinkman, 'Crime and Social Interactions' (1996) 111 The Quarterly Journal of Economics 507, 532.
[73] Virginia A Sharpe, *Accountability: Patient Safety and Policy Reform* (Georgetown UP 2004) 195.

In summary, the capabilities that allow the characterisation of specific robots as 'social' also create a context that extends to all accidental and unilateral interactions such robots may have with human beings. In that regard, the legal issues surrounding all -not just mutual- social interactions between human beings and robots are within the scope of the present study.[74]

### I.2.4 Information Security and Privacy

The ongoing proliferation of robots poses unprecedented threats to information security and privacy. When equipped with advanced sensors and enhanced with Internet connections, even non-social robots allow the surveillance of formerly inaccessible areas for once impossible lengths of time.[75] Social robots extend the information security and privacy risks beyond the threats posed by state-of-the-art technologies. The capabilities associated with the characteristic of social agency enable them to have access to sensitive personal information that would otherwise never have been entered into databases:

> Human beings tend to anthropomorphize social robots (…) in the sense that human beings might feel more inclined to see the robots as companions or friends. In turn, they will be more likely to entrust the robots with personal, potentially sensitive information.[76]

That sensitive personal information may include social robots' human interactants' cognitive-emotional states and can then be used for purposes such as psychologically manipulating human beings into performing certain behaviours or can be sold to third parties.[77] In some cases, psychological manipulations can be to the benefit robots' partners, such as motivating them to eat healthily, exercise regularly, or engage in mindfulness activities.

It is equally likely for such manipulations to be tied to exploitative pursuits of robots' manufacturers and programmers, such as pushing personalised advertisements or convincing robots' human interactants to perform behaviours they would not perform otherwise. The most severe threats to information security and privacy are not posed by robots' manufacturers, programmers, nor even by hackers but by State actors.

---

[74] Diana Marina Cooper, 'The Application of a "Sufficiently and Selectively Open License" to Limit Liability and Ethical Concerns Associated with Open Robotics' in Ryan Calo, A Michael Froomkin and Ian Kerr (eds), *Robot Law* (Edward Elgar 2016) 176-178.

[75] Ryan Calo, 'Robots and Privacy' in Keith Abney, George A Bekey, and Patrick Lin (eds), *Robot Ethics: The Ethical and Social Implications of Robotics* (MIT 2014) 191.

[76] Christoph Lutz and Aurelia Tamó-Larrieux, 'The Robot Privacy Paradox: Understanding How Privacy Concerns Shape Intentions to Use Social Robots' (2020) 1 Human-Machine Communication 87, 89.

[77] Calo, 'Robots and Privacy' (n 75).

Several governmental intelligence agencies have already been exposed to engage in surveillance activities at levels that would have put East Germany's Stasi -one of the most intrusive intelligence agencies to have ever existed- to shame.[78] Though mass surveillance of private spaces is expected to be prohibited in democracies (as expressed by the UN bodies)[79] and individual surveillance is allowed under restricted circumstances and only with the warrants issued by courts, recently uncovered surveillance scandals tell a different story.[80]

If national security interests are regarded as legitimate justifications for conducting mass surveillance activities without court warrants, social robots might legitimately undertake roles that correspond to having live-in government spies. The threats they pose would go beyond those that could have been posed by *telescreens*, the fictional devices in George Orwell's dystopian novel *Nineteen Eighty-Four* that operate as both televisions and cameras, broadcasting propaganda and performing surveillance at the same time.[81] Threats to information security and privacy do not always stem from the roles of robots as social interaction partners, but they still pose substantial challenges that become increasingly pronounced with the ongoing proliferation of social robots. For that reason, the scope of the present study includes such threats as well.

## I.3 Objectives of the Study

The present study identifies the legal issues surrounding human-robot social interactions, explores the root causes of these legal issues, and proposes a new unified legal framework to overcome legal systems' shortcomings in addressing these legal issues. The specific objectives of the study are as follows:

---

[78] The Ministry for State Security of the German Democratic Republic, often referred to as the Stasi (abbreviation of its German title, '*Ministerium für Staatsicherheit*'), was one of the most effective and repressive intelligence and secret police agencies. It is believed that the Stasi could get a comprehensive picture of anyone within of the German Democratic Republic, since it 'systematically monitored letters and packages, telephone calls and telegrams, and especially cross-border mail and telecommunications'. Jens Gieseke, *The History of the Stasi* (Berghahn 2014) 118. For a survey of current surveillance capabilities around the world, see Ben Underwood and Hossein Saiedian, 'Mass Surveillance: A Study of Past Practices and Technologies to Predict Future Directions' (2021) 4 Security and Privacy e142, e143-e156.

[79] UN Human Rights Council, (48th Session) 13 September–1 October 2021 'The Right to Privacy in the Digital Age: Report of the Office of the United Nations High Commissioner for Human Rights' (15 September 2021) UN Doc A/HRC/48/31, 12.

[80] *ACLU v NSA* 493 F 3d 644 (6th Cir 2007); Evan MacAskill and Gabriel Dance, 'NSA Files Decoded: Edward Snowden's Surveillance Revelations Explained' (*The Guardian,* 1 November 2013);

[81] George Orwell, *Nineteen Eighty-Four* (Secker & Warburg 1949) 177.

1. To develop the taxonomy of the legal issues that have arisen, and expected to arise in connection with social interactions between human beings and robots in various legal domains,
2. To identify the legal values protected by the legal principles that are related to the specified issues, or, in other words, to determine the purposes of existing legal instruments that are applicable to these issues,[82]
3. To evaluate whether existing legal instruments that are intended to protect the legal values thus identified can produce outcomes that are consistent with these values when applied to the specified issues,
4. To propose a new, unified legal framework that addresses the specified issues in a straightforward and practical manner that aligns with the purpose of protecting the legal values expressed by legal principles in various legal domains.

## I.4 Significance of the Study

The contemporary literature is not quiet about the legal issues surrounding social robots in society; numerous associations between these issues and various legal domains have already been established. Likewise, there is no shortage of studies discussing if existing legal instruments can effectively overcome these legal issues.[83] Though some of these suggestions even focus on the big picture and try to address multiple challenges in various legal domains simultaneously, most are too abstract to be practicable.[84]

The present study is the first legal research to explicitly view social robots as independent social actors and to frame the legal issues surrounding human-robot social interactions from that viewpoint. Therefore, the study is expected to contribute to filling in a significant gap in the scholarly literature. The conceptual approach employed is based on the intentional stance.[85]

---

[82] The present study assumes that the unique purpose of each legal instrument is the legal value that is supposed to be upheld by that instrument, expressed by underlying legal principles.
See generally Camelia Ignatescu, 'Letter and Spirit of the Law: An Interpretation Based on Principles of Law' (2013) 5 (2) Revista Romaneasca pentru Educatie Multidimensionala 75, 77; Stephen M Garcia, Patricia Chen, and Matthew T Gordon, 'The Letter versus the Spirit of the Law: A Lay Perspective on Culpability' (2014) 9 Judgment and Decision Making 479, 480.

[83] Sinziana M Gitiu, 'The Roboticization of Consent' in Ryan Calo, A Michael Froomkin and Ian Kerr (eds), *Robot Law* (Edward Elgar 2016) 201; Gabriel Hallevy, *When Robots Kill: Artificial Intelligence under Criminal Law* (NEU 2013) 82.

[84] Samir Chopra and Laurence F White, *A Legal Theory for Autonomous Artificial Agents* (UM 2011) 190; Andrea Bertolini and Francesca Episcopo, 'Robots and AI as Legal Subjects? Disentangling the Ontological and Functional Perspective' (2022) 9 Frontiers in Robotics and AI 6-7.

[85] Dennett (n 63) 111.

The intentional stance is part of Dennett's mental content theory that suggests several strategies to study the behaviours of other entities; these strategies can be summarised as follows:

**(1)** The physical stance requires making predictions about objects' behaviours from the knowledge of objects' physical compositions and laws of physics.[86]

**(2)** The design stance requires making predictions based on the assumption that objects' behaviours are outcomes of their specific designs.[87]

**(3)** The intentional stance requires making predictions by treating objects as rational agents whose behaviours are governed by their intentional states or states of mind.[88]

The present study submits that the knowledge of neither the physical compositions nor the design purposes of social robots provide suitable frames of reference for analysing and comprehending social robots' autonomous behaviours, making the intentional stance the only reliable strategy for studying the conduct of such robots. Since the distinctive characteristics of social robots -relative autonomy and social agency- resemble the mental states ascribed to human beings, the study's adoption of the intentional stance allows for the most accurate assessment of the behaviours of social robots.

The present study also has practical significance since it proposes a new, unified legal framework that addresses the legal issues regarding social robots in a fashion that is not only in line with the distinctive characteristics of social robots but also consistent with the legal values that underlie legal systems. The proposed framework integrates solutions to the legal issues arising in various legal domains.

Even though the proposed framework is devised to be implemented at the EU level, it is presumed to integrable into other jurisdictions because of the straightforward and non-political standpoint employed in its development.

---

See also Ugo Pagallo, *The Laws of Robots* (Springer 2013) 172; Christian List and Philip Pettit, *Group Agency: The Possibility, Design, and Status of Corporate Agents* (OUP 2013) 23-27.

[86] e.g., Archimedes' principle, describing one of the laws of physics, states that any object immersed in a fluid is acted upon by an upward, or buoyant, force equal to the weight of the fluid displaced by the object. From the physical stance, apples can be expected to float when thrown into the water since their density (on average 0.808 g/cm3) is lower than that of water (approximately 1 g/cm3). See G Vivek Venkatesh and others, 'Estimation of Volume and Mass of Axi-Symmetric Fruits Using Image Processing Technique' (2014) 18 International Journal of Food Properties 608, 609, Olof Dahl, Bo Eklund and Ann-Marie Pendrill, 'Is the Archimedes Principle a Law of Nature? Discussions in an 'Extended Teacher Room'' (2020) 55 (6) Physics Education 1, 3.

[87] e.g., since the purpose of brakes is to allow the drivers to slow down or stop their cars, making predictions from the design perspective, it is expected that cars will slow down when their brakes are applied.

[88] Intentional states refer to the mental states that have the characteristic of 'aboutness', i.e., being about, or directed at, objects or states of affairs in the world. Dennett (n 63) 17.

The practicality of the new framework proposed by the present study lies in the functional understanding of the law; since one of the law's roles in helping preserve the social order is to specify behavioural standards to be followed by individuals in society.[89] The intentionality-based conceptual approach adopted allows the present study to address the legal issues surrounding human-robot social interactions according to the behavioural logic of social robots and therefore strengthens the long-term feasibility of the proposed framework.

In summary, the study advances the understanding of the interplay between law and technology and brings a fresh perspective to the debate on legal issues surrounding human-robot social interactions. The present study goes beyond merely describing of pressing social and legal issues and provides practicable solutions in the form of a fleshed-out, new legal status for social robots.

## I.5 Organisation of the Study

The present study is organised into four parts over eight chapters.

**Part I, titled '*Background Concepts and Motivation*', constitutes the backdrop of the present study and consists of three introductory chapters.**

Chapter I is the *General Introduction*. In that chapter, the motivation for performing the research in the present study, the study's anticipated contribution to the existing body of knowledge in the literature and the research questions it intends to address are explained. Chapter I also explains the objectives of the present study and lays the intellectual bases of the following chapters.

Chapter II is the *Methodology of the Study*. In Chapter II, first of all, the approach taken toward the function of the law is presented. Then, the techniques utilised for gathering the information that provides the basis of legal analysis performed in the present study are described, and the choice of the comparative method is justified in line with the objectives set out in Chapter I. Chapter II also introduces the target legal system (EU) for the implementation of the recommendations of the present study and elaborates on the national legal systems selected for comparative research.

[89] Roscoe Pound, 'Theories of Law' (1912) 22 The Yale Law Journal 114, 144.

Chapter III is the *Conceptualisation of Social Robots*. In this chapter, operational definitions advanced by the present study for the concepts of *robot* and *social robot* are introduced. Though various definitions are already present across various resources for these concepts, none of these existing definitions are assessed to be precise enough to conduct an analysis of the legal implications of the distinctive characteristics of social robots. In that respect, Chapter III first explores the intellectual, technological, and formal origins of the concept of the robot, and then conceptualises social robots with reference to the above mentioned characteristics.

**Part II, titled '*Human-esque: Implications of Robot Anthropomorphism*', discusses the legal issues surrounding human-robot social interactions in relation to the phenomenon of robot anthropomorphism. This part consists of a single chapter.**

Chapter IV, the *Challenges of Robot Anthropomorphism* describes robot anthropomorphism with reference to the recorded experiences of human beings who interacted with various types of robots, as well as the experimental research on the subject. The chapter first establishes that because of the distinctive characteristics of social robots (relative autonomy and social agency) they can be expected to invoke unprecedented levels of anthropomorphic responses from human beings, and then discusses the legal implications of the phenomenon of robot anthropomorphism.

**Part III, titled '*Free-ish: Implications of Robot Autonomy*', focuses on the legal issues arising from the accountability gap created by the distinctive characteristics of social robots, especially that of relative autonomy. This part consists of two chapters.**

Chapter V, the *Challenges of Criminal Liability,* explores the challenges faced when determining the responsibility for criminally offensive behaviours of robots. Existing criminal law systems are assessed to be strictly anthropocentric, even more so than other legal domains, since they often focus on suppressing individual human beings' behaviours that disrupt the social order. Therefore, non-human entities cannot possibly be considered responsible subjects under current criminal laws, no matter how perceptive, independent, or intelligent they are. Since the characteristic of relative autonomy allows robots to engage in criminally offensive behaviours without any criminal mental state that can be associated with human beings, certain liability gaps can be expected to arise in existing criminal law systems.

Chapter VI, the *Challenges of Civil Liability*, explains the challenges faced in the context of private law. Since only human beings (natural persons) and human associations (juridical persons) possess any legal capacity under the law, it is assessed that robots can only be treated as mere products under private law, with no independent decision-making powers that can have legally significant consequences. In other words, legal instruments already offered by tort law and contract law systems cannot give effect to the distinctive characteristics of social robots and, therefore, cannot adequately address the legal issues surrounding human-robot social interactions.

**Part IV, titled '*Pax Robota: Recommendations and Conclusions*', discusses various proposals to allow the target legal system (EU) to adequately address the legal issues surrounding human-robot social interactions. The present study submits that the legal issues regarding social robots can be most effectively addressed if the legal instruments applicable to such issues recognise the legally significant, distinctive characteristics of social robots. In that regard, Part IV presents a new unified legal framework that includes a new legal status for social robots as well as suggestions for modifications in several domains of law.**

Chapter VII, the *Feasibility of the New Legal Status*, establishes the groundwork for the recommended new legal status. The chapter first discusses the necessity of some form of new legal status for social robots and then elaborates on the conceptual possibility of introducing a new legal status.

Chapter VIII, *the Future of Robot Law*, is the last chapter of the present study. It ties up loose ends, explains the content of the proposed new legal status for social robots, and elucidates the broader legal framework proposal for the future of robot law. However, the introduction of new legal status for robots would effectively address the legal implications of robot autonomy; that alone does not address the legal issues arising from human beings' responses to the relative autonomy and social agency of robots and therefore needs to be supplemented by some other modifications across domains of law. The new legal status proposed in the chapter is referred to as *quasi-personhood* (almost personhood) since it recognises the relative autonomy of social robots and confers the legal capacity to act on them. The status is not equivalent to existing forms of legal personhood since it acknowledges that the independent activities of robots are not driven by their own individual interests.

# Chapter II: Methodology of the Study

> The law is there to be fulfilled. Fulfilment is the life and the truth of law; it is the law itself. What is not fulfilled, that only exists in legislation or on paper, are mere fictions about law, empty collections of words. On the other hand, whatever is fulfilled as law, is law, even though it cannot be found in legislation, and neither the people nor science is aware of it.
>
> — Rudolph von Jhering, *The Spirit of Roman Law*[1]

## Introduction

For the purposes of the present study, the concept of *legal challenge* is taken to refer to the inadequacy of existing legal instruments in upholding the underlying legal principles. To use the phraseology of Rudolph von Jhering, the present study examines whether the law can fulfil its existential purpose when confronted with legal challenges arising from the distinctive characteristics of social robots.[2]

Social robots are becoming increasingly present in the ordinary course of everyday life. They already undertake various societal roles, ranging from elder care to sexual gratification. However, except for a few minor incidents, interactions between human beings and social robots have not yet resulted in any type of problematic situations that require courts or lawmakers to elucidate what approach they award toward the characteristics of social robots.[3] Because of the inadequacy of existing legal instruments in the face of the distinctive characteristics of social robots, relative autonomy and social agency, several legal gaps surrounding human-robot social interactions are expected to arise. These legal gaps cannot be plugged by national-level efforts alone. As is the case for many other technologies, social robots' design, manufacturing, and operational processes have transnational dimensions, and, thanks to the Internet, human-robot social interactions can also be expected to have significant cross-border implications.

---

[1] Rudolf von Jhering, *Geiſt des Römiſchen Rechts*, vol 2 (Breitkopf un Härtel 1858) 334.

[2] Hasso Hoffman, 'From Jhering to Radbruch: On the Logic of Traditional Legal Concepts to the Social Theories of Law to the Renewal of Legal Idealism' in Damiano Canale, Paolo Grossi and Hasso Hoffman (eds), *A Treatise of Legal Philosophy and General Jurisprudence* (Springer 2009) 317.

[3] e.g., In 2015, *Random Darknet Shopper,* an online shopping robot meant to explore the dark web, purchased Ecstacy pills and was apprehended by the Swiss police. The police subsequently released the robot without pressing charges since it was part of an art exhibition. Jana Kasperkevic, 'Swiss Police Release Robot That Bought Ecstasy Online' (*The Guardian*, 22 April 2015).
See generally Chapter V, 'Challenges of Criminal Liability'.

The present study holds that the legal gaps could be most effectively filled only through EU-level legal instruments, revealing the EU as the target legal system for the proposed unified legal framework. Unlike national legal systems, EU law does not stand alone, meaning that EU legal instruments are applied in conjunction with those of the Member States. Hence, when developing the new unified legal framework that addresses the legal issues surrounding human-robot social interactions for EU law, it is necessary for the present study to examine the Member States' national legal systems, calling for the use of comparative methods.

In that regard, the following chapter elaborates on the methodology of the present study. The chapter consists of four sections. In the first section, the postulates adopted by the present study regarding the function of the law are presented. Then, the second section explains the need for utilising comparative research methods to attain the objectives of the present study. It should be emphasised that the methodology employed by the present study is not purely comparative, but rather utilises the comparative method to identify commonalities and divergences across the legal instruments of national legal systems to identify the common legal values that these instruments are intended to protect. The third section describes the study's design and discusses the stages of the research process. Finally, in the fourth section, the national legal systems chosen for comparison in the present study are introduced, and the justifications for choices are explained.

## II.1 Function of the Law

> The end of the law is peace. The means to that end is war. So long as the law is compelled to hold itself in readiness to resist the attacks of wrong—and this it will be compelled to do until the end of time—it cannot dispense with war. [4]

The present study holds that the law is one of the functional sub-systems of the overarching general system of social order. In that respect, the end that the law is means for is understood as the preservation of social order.[5] If the concept of *social order* refers to the sum of relations and structures that hold society together, the function of the law can be construed as the creation and protection of conditions of social life through the peaceful conduct of social interactions.[6]

---

[4] Rudolf von Jhering, *The Struggle For Law* (John J Lalor tr, Callaghan & Company 1915) 1.

[5] David A Funk, 'Major Functions of Law in Modern Society' (1972) 23 Case Western Reserve Law Review 257, 281.

[6] The role of the law attributed by the present study regarding the preservation of social order does not necessarily mean that the law sets communal objectives for the welfare of society. The function of the law concerning social order is limited to setting the scene for the peaceful conduct of social interactions, not designing and encouraging specific social interactions that would help achieve certain socio-economic levels of welfare in society.

Human beings are independent entities, and the actions they perform are driven by the unique needs, desires, and goals of every human being. Though they have the power of free will, the use of that power is often driven by the individual interests. That said, human beings are also social creatures. Unable to fulfil most these their individual interests alone, they are driven by the very same interests to live together as part of society.[7] Since the individual interests of human beings are optimally fulfilled in society, human beings can be assumed to have some shared interests in preserving and improving the conditions of social life. These shared interests are claimed to the benefit of all members of society, hence, they can be referred to as social interests.[8] Though social interests are expressed on behalf of the whole society, it must be noted that society is not some superorganism and does not operate as an organic whole.

***We are not the Borg.***[9] Human beings, even when living together as members of society, maintain separate sets of individual interests, shaped by unique circumstances. Despite the presence of social interests along with these distinct individual interests, society itself does not have a separate existence from that of the human beings who form it. There is no 'unity of phenomenological experiences' across members of society.[10] That being the case, *society* can best be defined as the functional association of human beings, held in place through the coercive power vested in the State.[11] In light of the absence of sufficient resources to satisfy all needs, desires and goals of each human being, there are bound to some mutually exclusive goods that are subject to multiple individuals' interests even in the most peaceful society, making conflicts and competitions among members of society inevitable.[12] At that point, the law comes into play as the guardian of social order, preserving the conditions of social life and harmonising social relations by reconciling the disputes between members of society that have conflicting individual interests.

The balancing of conflicting interests is generally accomplished by following the direction set out by the legal principles that are enforced by the applicable legal instrument:

---

[7] Vera L Buijs and others, 'Social Needs and Happiness: A Life Course Perspective' (2020) 22 Journal of Happiness Studies 1953, 1955.
[8] Roscoe Pound, 'A Survey of Social Interests' (1943) 57 Harvard Law Review 1, 3.
[9] The Borg Collective is a fictional space-faring civilisation in TV's *Star Trek*. Members of the Borg Collective, known as *drones,* are linked into the hive mind that connects all of them and allows information from every drone to be shared with the rest; the Borg Collective is *collectively* conscious, but drones are unaware of themselves as separate individuals. The statement above is inspired from one of the famous Borg hails: *'We are the Borg. Existence, as you know it, is over. We will add your biological and technological distinctiveness to our own. Resistance is futile.'*
[10] John Danaher and Steve Petersen, 'In Defence of the Hivemind Society' (2020) 14 Neuroethics 253, 2.
[11] Christian List and Philip Pettit, *Group Agency: The Possibility, Design, and Status of Corporate Agents* (OUP 2011)179.
[12] Paul A Samuelson, *Economics* (McGraw-Hill 1973) 23.

> Each command of the law determines a conflict of interests; it originates from a struggle between opposing interests and represents as it were the resultant of these opposing interests. Protection of interests through law never occurs in a vacuum.[13]

The direction pointed out by applicable legal instruments is not always toward finding compromises between conflicting individual interests. From time to time, some interests are given more weight than others. The role of legal principles (i.e., 'commands of the law') that underlie existing legal instruments is to determine which interests will be given more weight, reflecting the value judgements emerging from political processes.

## II.2 Function of Comparative Law

As social robots are still emerging technologies in many ways, the projected personal, economic, and social impacts of human-robot social interactions have not been wholly realised anywhere. Hence, there are few first-hand facts about the legal issues that may arise from these impacts. However, the limited available information suggests that the relative autonomy and the human psyche's innate tendency to anthropomorphise social robots are the root causes behind the legal issues surrounding human-robot social interactions.[14]

These root causes do not change according to ethnographic-regional conditions; therefore, legal issues regarding social robots can arise in any legal system. While ethnographic-regional conditions do not impact the root causes behind the legal issues surrounding human-robot social interaction, they shape the legal systems. Since every country has unique ethnographic-regional conditions, different national legal systems can be expected to employ varying strategies for addressing the same legal issues. Blaise Pascal, the great French philosopher, expresses the differences among national legal systems as follows:

> Fundamental laws change after a few years of possession; right has its epochs; the entry of Saturn into the Lion marks to us the origin of such and such a crime. A strange justice that is bounded by a river! Truth on this side of the Pyrenees, error on the other side.[15]

The comparative methods allow the present study to examine how the legal issues surrounding human-robot social interactions arise in different cultural and socio-economic settings across various legal systems.

---

[13] Philip Heck, 'Interessen Jurisprudenz' in Magdalena Schoch (ed), *The Jurisprudence of Interests: Selected Writings of Max Rümelin, Philipp Heck, Paul Oertmann, Heinrich Stoll, Julius Binder, Hermann Isay.* (Magdalena Schoch tr, Harvard UP 1948) 35.

[14] Adam Waytz, John Cacioppo, and Nicholas Epley, 'Who Sees Human?' (2010) 5 Perspectives on Psychological Science 219, 220; Nick Neave and others, 'The Influence of Anthropomorphic Tendencies on Human Hoarding Behaviours' (2015) 72 Personality and Individual Differences 214, 215.

[15] Blaise Pascal, *Pascal's Pensees* (Dutton 1958) 85.

The present study gains the required breadth of perspective on the legal issues regarding social robots through comparative methods. The analysis of the differences and similarities between the approaches of selected national legal systems empowers the present study to develop feasible solutions that can be implemented in various legal systems, revealing a somehow 'universal benchmark exempted from partisan preconceptions'.[16]

The cross-border aspects of the legal issues examined in the present study also call for comparative methods. Social robots, like any other socially interactive technology, are designed to utilise the Internet for the performance of at least some of their tasks. While on the Internet, they are bound to interact with entities that are not located in the same country with them or even in the same continent. Considering the present and future applications of social robots, the growing numbers of these robots can be expected to cause upsurges in the incidence of Internet-related legal disputes that have cross-border aspects. Whereas the reach of national legal instruments is often limited by the territories of the States that have enacted them, the Internet is not territorially bound.

Conventionally, the legal disputes that have transnational dimensions are resolved through conflict-of-laws rules.[17] However, there are no universally accepted rules for resolving Internet-related transnational legal disputes, nor are there any rules concerning social robots. With the anticipated upsurge of transnational legal disputes involving social robots, the need for universally acceptable conflict-of-laws rules on the legal issues surrounding human-robot social interactions becomes critical. Developing universally acceptable rules on any subject matter requires the identification of common legal principles that are enforced through various legal instruments across multiple national legal systems, based on the understanding that every national legal system faces similar issues, addresses these issues in different ways, but intends to deliver comparable outcomes:

> Comparative law, embracing also foreign legal systems, has a universal humanistic outlook (…) like other branches of science, it contemplates that, while techniques may vary, the problems of justice are basically the same in time and space throughout the world.[18]

---

[16] Pier Giuseppe Monateri, 'Introduction' in Pier Giuseppe Monateri (ed), *Methods of Comparative Law* (Edward Elgar 2012) 3.

[17] Conflict-of-laws rules identify the connecting factors of the disputes and attach disputes to the most closely connected national legal system. Connecting factors are elements of conflict of laws rules that connect factual situations with the laws of a specific legal system.

[18] Hessel E Yntema, 'Comparative Legal Research: Some Remarks on "Looking Out of The Cave"' (1956) 54 Michigan Law Review 899, 903.

Common legal principles can be comprehended as the purpose values legal systems intend to achieve through legal instruments. Then, though universal, these common legal principles are not immutable. It is evaluated that objective comparative analyses of national legal systems can help identify these common legal principles as revealing these principles stand out among the original purposes of comparative legal research methods, with the output of these methods described by some in the literature as 'natural law with variable content'[19] or 'ideal relative law'.[20]

The use of comparative methods by the present study is also justified because of the transnational aspects of the proposed unified legal framework. The proposed framework is intended to be integrable to the national legal systems of the EU countries and therefore leans on the examination of 'the causes which underlie the origin, development and extinction of legal institutions' and the identification of common legal principles shared by these jurisdictions.[21]

One of the best-known outputs of comparative legal research is the Draft Common Frame of Reference ['DFCR'].[22] Viewed by many as the 'European Civil Code in all but name',[23] authors of the DFCR state that they merely attempted to show 'how much national private laws resemble one another' and 'how much those laws may be regarded as regional manifestations of an overall common European legacy'.[24] Whatever the authors' original intent might have been, it is concurred that the DFCR facilitated the development of some sort of shared understanding of private law among European countries by defining shared legal concepts and identifying common private law principles across. Encouraged by the success of the DFCR, the present study focuses on uncovering the common principles that can provide the basis for new legal instruments that can adequately address the legal issues surrounding human-robot social interactions.

---

[19] Rudolf Stammler, *The Theory of Justice* (Lawbook Exchange 2000) 515.
[20] Carleton Kemp Allen, *Law in the Making* (OUP 1927) 25.
[21] Harold Cooke Gutteridge, *Comparative Law: An Introduction to the Comparative Method of Legal Study and Research* (CUP 1946) 6.
[22] 'Definitions, Principles and Model Rules of European Private Law: Draft Common Frame of Reference Outline Edition' (*Study Group on a European Civil Code and the Research Group on EC Private Law*) (1st edn, Sellier, European Law Publishers 2009) (DFCR).
[23] Gerhard Wagner, *The Common Frame of Reference: A View from Law & Economics.* (Sellier, European Law Publishers 2009) 204; Nils Jansen and Reinhard Zimmermann, 'A European Civil Code in All but Name: Discussing the Nature and Purposes of the Draft Common Frame of Reference' (2010) 69 Cambridge Law Journal 98, 99.
[24] DFCR (n 22) 6 para 7.

The comparative law aspect of the study is mainly conceptual and focuses on how relevant legal issues are pinpointed and conceptualised across the legal systems selected for comparison. The study emphasises elucidation of the legal values that pertinent legal instruments are intended to protect, rather than detailed technical description of how these instruments work within their specific ethnographic-regional conditions. Accordingly, the research mainly relies on analysing legislation and secondary resources, such as academic commentary. The more limited reliance on case law reflects the purpose for which the comparative method is employed.

For the present study, the origins of the *robot* concept are sought firstly in mythological and religious accounts, then in automata of the pre-industrial era, and finally in the contributions of science fiction. Literature, as it informs all aspects of human life, influences technological developments by conceiving of them far before humanity can bring them to life, and it dramatically influences the law by humanizing its abstract concepts and envisioning its effects. Literature opens the door to a much deeper understanding of the complexities of human behaviour, motivations, and the consequences of actions. Literature helps to identify gaps in the law and suggest areas where the law needs to be improved to address contemporary societal issues.[25] Literature can also challenge legal assumptions, stimulate debate, and inspire change by exploring different perspectives. The field of robotics very often interacts with science fiction, and the law ought to follow suit. The present study provides a limited analysis of three literary works of significant influence; Mary Shelley's *Frankenstein*, a cautionary tale of modern science unrestrained by ethical concerns where autonomous entities might turn against the human beings that created them; Karel Čapek's *Rossum Universal Robots,* where artificial living beings develop sophisticated thought and emotion and as a result kill all human beings and then build their own egalitarian society, and Isaac Asimov's *I, Robot* where robots are manufactured with built-in safeguards that constrain them from knowingly harming human beings.

[25] Richard A Posner, *Law and Literature* (3rd edn, Harvard UP) 389.

## II.3 Design of the Study

The research process of the present study consists of three stages. In the first stage, the legal issues surrounding human-robot social interactions are conceptualised, and the legal instruments applicable to these issues are identified. The values expressed by the legal principles underlying the applicable legal instruments are evaluated in the second stage. Finally, in the third stage, a new, unified legal framework that aligns with the legal values expressed by the legal principles is constructed. The proposed framework is intended to offer integrated solutions to the legal issues regarding social robots in various legal domains and jurisdictions.

### II.3.1 Conceptualisation of Legal Issues

The first stage of the research process conceptualises the legal issues surrounding human-robot social interactions. In the legal literature, no relationship yet exists between the concepts of 'human being', 'robot', and 'social interaction'. Consequently, the legal issues surrounding human-robot social interactions cannot be characterised straightforwardly, and the existing legal instruments applicable to these issues cannot be precisely pinpointed. Therefore, the present study uses constructive analysis methods to provide concept-and-connection-based descriptions of the legal issues surrounding human-robot social interactions.

The conceptual descriptions of the legal issues regarding social robots are developed 'by postulating new relations' or 'by stating that some already known relations hold among previously unrelated concepts'.[26] To avoid substantial drawbacks of orthodox conceptual analysis (such as overlooking the context and practical irrelevance of the conclusions reached) the postulated relations are compared with similar relations already present in the doctrine and jurisprudence. Where no similar relations can be found, speculative scenarios are utilised to explain the conceptual descriptions postulated.[27]

[26] Miloš Kosterec, 'Methods of Conceptual Analysis' (2016) 71 Filozofia 220, 222.
[27] A Botes, 'Concept Analysis: Some Limitations and Possible Solutions' (2002) 25 Curationis 23, 24.

### II.3.2 Identification of Legal Values

The second stage of the research process focuses on identifying the legal values expressed by the legal principles that underlie the existing legal instruments applicable to the legal issues surrounding human-robot social interactions. If the law's purpose is to preserve the general conditions of social life, that purpose is fulfilled by applying the existing legal instruments to conflicts of interests in society. As no legal instrument can satisfy all conflicting interests simultaneously, legal instruments may prioritise some interests at the expense of others. When any given legal instrument prioritises some interest, a value is attached to that prioritised interest. In turn, the purpose of that instrument becomes the protection of that value.

On the assumption that 'the law is made up of rules, principles, and precedents that must be known, and that these rules, principles, and precedents form a coherent system',[28] the present study initially identifies the legal values that are supposed to be upheld by applicable legal instruments through the doctrinal analysis methods traditionally used to describe the content of the law. Nonetheless, since the doctrinal analysis methods have no comparative dimensions, they have no use beyond delivering notional explanations of the applicable existing legal instruments, and they do not provide the necessary tools for the objectives of the present study.[29] Therefore, the study reinforces the use of doctrinal analysis methods by conducting structural comparisons between the chosen legal systems.

These comparisons are conducted through the joint method of agreement and difference, and the specific socio-cultural contexts of the chosen national legal systems are given detailed consideration.[30] Since the present study approaches the law as a functional system for reconciling conflicting interests, the research is not confined to the descriptive analysis of the legal values regarding the legal issues surrounding human-robot social interactions. In addition, the adequacy of existing legal instruments is also evaluated by performing intuitive legal impact analysis.[31]

---

[28] Kevin Walby, 'Research Methods in Law' (2014) 48 Law & Society Review 486.
[29] Terry Hutchinson, 'The Doctrinal Method: Incorporating Interdisciplinary Methods in Reforming the Law' [2015] 3 Erasmus Law Review 138, 142.
[30] Botes (n 27) 24.
[31] The legal impact analysis focuses on 'the informed knowledge upon the working of past laws and the possible consequences of proposed new legislation'. Colin S Gibson, 'Legal Impact Analysis: The Ideal and the Practicable' (1982) 24 Journal of the Indian Law Institute 837, 841.

Through the intuitive legal impact analysis, the present study forecasts the outcomes of applying existing legal instruments to the legal issues surrounding human-robot social interactions. The legal impact analysis determines whether existing legal instruments might still fulfil their purposes regarding the protection the legal values expressed by the underlying legal principles.

### II.3.3 Construction of New Legal Framework

The third stage of the research process focuses on constructing a new, unified legal framework that develops on the existing legal instruments to overcome the shortcomings of legal systems. The present study concedes that the effective resolution of the legal issues depends on fulfilling the legal values behind the existing legal instruments[32]

The present study refrains from directly challenging the legal values safeguarded by the existing legal instruments. The legal values are determined following political processes, on behalf of society, and passing judgments on the merits and accuracy of the outcomes of political processes are outside the remit of legal research. Given that the legal issues regarding social robots reveal the inadequacies of existing legal instruments across multiple jurisdictions, the legal framework proposed by the present study is intended to constitute a highly flexible, multidimensional legal model. Thanks to the use of comparative methods in the present study, the proposed framework becomes less vulnerable to the weaknesses of classical normative approaches, such as over-abstraction, limited scope, and over-prescription.

[32] ibid.

## II.4 Choice of Legal Systems

### II.4.1 EU Law

The European Union (EU) is a *sui generis* legal system, which means that it has its own unique legal framework that is distinct from those of its member states. This is evident in the decisions made by the Court of Justice of the European Union (CJEU), the highest court in the EU responsible for interpreting EU law and ensuring its consistent application across all the Member States.[33] The EU's legal framework is based on a combination of treaties, regulations, directives, and case law, which are interpreted and applied by the CJEU. This *sui generis* legal system enables the EU to operate as a supranational entity, with powers and competencies that extend beyond those of its member states.

EU law, for the most part, aspires to gradually phase out inconsistencies among national legal systems so that similar socio-legal and economic conditions exist in all the Member States.[34] Most of the EU legal instruments define the common legal principles to be heeded by all the Member States whilst allowing them to decide *how* to heed these principles. Thus, the outlook of EU law is assumed to be closest to the statement of common legal principles.

That assumption is strengthened by the Brussels Effect, which states that EU legal instruments could exert significant influence on legal entities outside the EU because of the market force of the EU.[35] The present study identifies common legal principles to provide the basis for the proposed legal framework that would address the legal issues surrounding human-robot social interactions adequately. As EU law is the legal system closest to the common denominator of various legal systems, the EU legal system is identified as the target legal system for the recommendations advanced by the study.

---

[33] In the landmark case of *Costa v ENEL*, the CJEU ruled that EU law takes precedence over national laws of member states. This means that EU law prevails over conflicting national laws, even if the latter were enacted after the former. The CJEU also established in *Van Gend en Loos v Nederlandse Administratie der Belastingen* that individuals can rely on EU law before national courts and that EU law can have direct effect, meaning that individuals can directly enforce their rights under EU law in national courts.
See, *Flaminio Costa v ENEL* (Case 6/64) [1964] ECR 585, *Van Gend en Loos v Nederlandse Administratie der Belastingen* (Case 26/62) [1963] ECR 1.

[34] Klaus-Dieter Borchardt, *The ABC of EU Law* (7th edn, Publications Office of the EU 2018) 101.

[35] Anu Bradford, 'The Brussels Effect' (2012) 1 Northwestern University Law Review 1, 9. Francis Jacobs, 'Comparative Law and European Union Law', in Mathias Reimann and Reinhard Zimmermann (eds), *The Oxford Handbook of Comparative Law* (2nd edn, OUP 2019) 523, 525.

EU law features two binding legal instruments for the harmonisation of Member States' national laws; or in other words, there are two types of harmonising measures that the EU can adopt: Regulations and Directives.[36]

- *EU Regulations* are legal instruments directly applicable in all Member States without needing to be transposed into national laws.[37] Intended to achieve full harmonisation across the EU, regulations provide uniform legal frameworks on their subject matter areas that prevail over national rules of the Member States. Essentially, they replace different national rules with common European ones.

  For instance, over the past few decades, the EU has adopted multiple regulations in the field of private international law, and these are assessed to have resulted in the 'considerable legislative unification' of conflict of law rules within the EU.[38]
- *EU Directives* require the Member States to adopt national rules and regulations to achieve specified outcomes. They are not directly applicable; the Member States must effect changes in their national laws to achieve outcomes specified by directives.[39] Intended to achieve only partial harmonisation (or approximation), directives eliminate significant differences between national legal systems and create minimum standards in the subject matter areas they address.[40]

  For example, directives on safety and health at work have approximated national rules on workplace conditions throughout the EU. While the Member States can maintain or establish more stringent rules for improving safety and health at work, these directives guarantee minimum safety and health standards are observed throughout the EU.[41]

[36] *Decisions* are also binding EU legal instruments, but they only apply to the Member States, human associations, or human beings they specifically address; as such, they are not harmonising measures. Consolidated Version of the Treaty on the Functioning of the European Union [2012] OJ C 326/47 (TFEU), Art 288: ' A decision shall be binding in its entirety. A decision which specifies those to whom it is addressed shall be binding only on them.'

[37] TFEU (n 36), Art 288: 'A regulation shall have general application. It shall be binding in its entirety and directly applicable in all Member States.'

[38] Xandra Kramer and others, 'A European Framework for Private International Law: Current Gaps and Future Perspectives' (Study, European Parliament 2012) 17-21.

[39] TFEU (n 36), Art 288: 'A directive shall be binding, as to the result to be achieved, upon each Member State to which it is addressed, but shall leave to the national authorities the choice of form and methods.'

[40] Peter de Cruz, *Comparative Law in A Changing World* (Cavendish 1999) xx.

[41] Council Directive 89/391/EEC of 12 June 1989 on the Introduction of Measures to Encourage Improvements in the Safety and Health of Workers at Work [1989] OJ L183/1.

The present study holds that EU legal instruments should be adopted to address the legal issues surrounding social interactions between human beings and robots. Suppose no common principles on the legal issues regarding social robots are introduced at the EU level. In that case, different approaches among the Member States can be expected to discourage the development and deployment of robots across the EU and complicate the resolution of cross-border disputes arising from human-robot social interactions.

In summary, EU-level legal instruments are becoming increasingly essential in facing legal issues surrounding human-robot social interactions. Considering that most Member States have not yet adopted any legal instruments to address the legal implications of robots' distinctive characteristics, the present study holds that EU regulations provide the most practical and efficient way of clarifying the law's treatment of robots across the Member States.

### II.4.1.a Principle of Conferral

The principle of conferral, one of the fundamental principles of EU law, determines the division of law-making powers between the Member States and the EU. According to that core principle, the EU may only act 'within the limits of the competences conferred upon it by the Member States in the Founding Treaties'.[42] Whether the EU can adopt legal instruments that harmonise the law's approach to human-robot social interactions depends on whether various legal implications would fall under one of the competencies conferred upon the EU. According to the Founding Treaties, the EU has three types of competences.

**1.** ***Exclusive Competence*****:** In areas of exclusive competence, only the EU has the competence to adopt legally binding instruments. Unless empowered by the EU, the Member States cannot adopt any national legal instruments in the policy areas designated as exclusive competence, nor can they retain their existing legal instruments.[43]

[42] TFEU (n 36), Art 2 (1): 'When the Treaties confer on the Union exclusive competence in a specific area, only the Union may legislate and adopt legally binding acts, the Member States being able to do so themselves only if so empowered by the Union or for the implementation of Union acts.'

[43] TFEU (n 36), Art 2 (1): 'When the Treaties confer on the Union exclusive competence in a specific area, only the Union may legislate and adopt legally binding acts, the Member States being able to do so themselves only if so empowered by the Union or for the implementation of Union acts.'

TFEU Art 3 defines *exclusive competence areas* as the competition rules within the internal market, the customs union, the common commercial policy, monetary policy for Eurozone countries, the conservation of marine biological resources, and the conclusion of some international agreements.

The present study submits that exclusive competence areas relate to the EU's political and economic governance and operational structure. Thus, it is logical for such areas to be addressed through EU legal instruments, as the Member States have no reason to adopt national legal instruments. Considering that the legal implications of social interactions between human beings and robots do not concern the governance or structure of the EU, it is supposed that the uniformisation of the law's treatment of robots does not fall within the scope of exclusive competence areas.

**2.** ***Supporting Competence*****:** In supporting competence areas, both the EU and the Member States can adopt binding legal instruments. However, legal instruments adopted by the EU in these areas can only 'support, coordinate or complement the actions of the Member States' and must not preclude them from adopting their legal instruments.[44]

TFEU Art 6 defines *supporting competence areas* as protection and improvement of human health, industry, culture, tourism, education, youth, sport and vocational training, civil protection, and administrative cooperation. TFEU Art 2(5) provides that EU legal instruments in supporting competence areas 'shall not entail harmonisation' of national laws. As uniformisation denotes much closer integration of laws than harmonisation, it is inferred that the EU cannot adopt uniformising instruments in these areas. Therefore, the present study does not focus on to which extent the implications of human-robot social interactions can be regulated under the supporting competences of the EU.

**3.** ***Shared Competences*****:** In areas of shared competence, both the EU and the Member States can adopt binding legal instruments, but the Member States can only do so to the degree that the EU has not exercised or has decided not to exercise its competence.[45]

---

[44] EU Law does not have primacy over national laws in supporting competence areas.
See TFEU (n 36), Art 2 (5): ' (…) the Union shall have competence to carry out actions to support (…) the actions of the Member States, without thereby superseding their competence in these areas.'

[45] TFEU (n 36), Art 2(2): 'When the Treaties confer on the Union a competence shared with the Member States in a specific area, the Union and the Member States may legislate and adopt legally binding acts in that area. The Member

TFEU Art 4(1) articulates that areas of shared competence comprise all subject matter areas that are not designated as exclusive competence by TFEU Art 3 or supporting competence by TFEU Art 6. For that reason, shared competence is regarded as the ordinary competence of the EU.[46]

Notwithstanding, TFEU Art 4(2) describes *principal areas of shared competence* as the internal market, agriculture and fisheries, transport, trans-European networks, energy, certain areas of social policy, economic, social and territorial cohesion, environment, certain areas of public health, consumer protection and the area of freedom, security and justice, common safety concerns in public health matters, research and technological development, space, development cooperation and humanitarian aid.[47]

Based on the expressional differences between the provisions, it is reasoned that TFEU Art 3 provides an exhaustive list of exclusive competence areas subject to *numerus clausus* restrictions, since it states that 'the Union shall have executive competence *in the following areas*' (emphasis added). In contrast, TFEU Art 4(2) is regarded to establish a residual category for shared competence area since it states that 'shared competence between the Union and the Member States applies *in the following principal areas*' (emphasis added).[48] The present study submits that the EU is empowered to harmonise the law's treatment of robots across the Member States based on its shared competences. However, the study also notes that specific shared competence areas on which EU legal instruments would be predicated are not indisputable for all of the legal issues surrounding human-robot social interactions, especially those related to criminal law. EU legal instruments introduced thus far primarily relate to the private law aspects of the law's treatment of robots. These legal instruments and the legal bases they are predicated on are as follows:

States shall exercise their competence to the extent that the Union has not exercised its competence. The Member States shall again exercise their competence to the extent that the Union has decided to cease exercising its competence.'

[46] Robert Schütze, 'EU Competences: Existence and Exercise' in Anthony Arnull and Damian Chalmers (eds), *The Oxford Handbook of European Union Law* (OUP 2015) 75, 86.

[47] TFEU (n 36), Art 4.

[48] Barbara Guastaferro, 'The European Union as A Staatenverbund? The Endorsement of the Principle of Conferral in the Treaty of Lisbon', in Martin Trybus and Luca Rubini (eds), *The Treaty of Lisbon and the Future of European Law and Policy* (Edward Elgar 2012) 117, 125

**a.** EP Resolution on *Civil Law Rules on Robotics* 'asks the Commission to submit, on the basis of Article 114 TFEU, a proposal for a legislative instrument on legal questions related to the development and use of robotics and AI'.[49]

**b.** EP Resolution on *a Civil Liability Regime for Artificial Intelligence*, 'requests the Commission to submit a proposal for a regulation on liability for the operation of the AI systems 'having regard to the Treaty on the Functioning of the European Union, and in particular Article 114 thereof'.[50]

**c.** EP Resolution on *a Framework of Ethical Aspects of AI and Robotics* 'requests the Commission to submit a proposal for a Regulation on ethical principles for the development, deployment of (…) robotics and related technologies on the basis of Article 114 of the Treaty on the Functioning of the European Union'.[51]

**d.** EC's Proposal for *the EU AI Act* identifies its legal basis as 'in the first place Article 114 (...) which provides for the adoption of measures to ensure the establishment and functioning of the internal market'.[52]

TFEU Art 114 empowers the EU to enact legal instruments for harmonising national laws to facilitate the establishment and functioning of the EU single market. In modern free market economies such as those of the Member States, the fulfilment and enforcement of private law obligations -either assumed through promises or imposed by law- are crucial for the healthy operation of the economy. Considering that the EU single market is the amalgamation of multiple national markets, spanning over the territories of all Member States, significant differences among private law rules in the national legal systems would discourage the free movement of goods, services, capital, or people and thus obstruct the functioning of the EU single market. Harmonisation of private law rules across the Member States can prevent detrimental legal fragmentation within the internal market and improve its functioning, justifying the use of TFEU Art 114 as the legal basis for adopting harmonising measures in private law at the EU level.

[49] European Parliament Resolution 2015/2103(INL) of 17 February 2017 with Recommendations to the Commission on Civil Law Rules on Robotics [2018] OJ C252/239, para 51.
[50] European Parliament Resolution 2020/2014(INL) of 20 October 2020 with Recommendations to the Commission on a Civil Liability Regime for Artificial Intelligence [2021] OJ C404/10.
[51] European Parliament Resolution 2020/2012(INL) of 20 October 2020 with Recommendations to the Commission on a Framework of Ethical Aspects of Artificial Intelligence, Robotics and Related Technologies [2021] OJ C404/63, para 146.
[52] European Commission, 'Proposal for a Regulation of the European Parliament and of the Council Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act) and Amending Certain Union Acts' COM (2021) 206 final (EU AI Act) 6.

The EU is vested with supranational legislative authority. However, that authority is not all-encompassing and universal. Especially in the core areas of national governance, the Member States maintain their sovereign law-making powers. The fields of procedural and substantive criminal law are regarded as within these core areas as they are linked with national-moral value judgements specific to each Member State.

These specific value judgements are expressed in substantive criminal law and enforced through procedural criminal law.[53] Though the Founding Treaties authorise the EU to introduce legal instruments for harmonisation of national rules on specific aspects of procedural and substantive criminal law, the powers exercised in these fields are highly restricted compared to those exercised in private law areas.[54] In these fields, the EU can only introduce directives that intend to achieve minimum standardisation of national rules. The Member States cannot be precluded from adopting measures that afford higher levels of protection for both the defendants and victims of crimes. [55] The legal bases and limitations for introducing legal instruments that harmonise national rules on substantive and procedural criminal law are established by TFEU Arts 82 and 83.

In relation to procedural criminal law, TFEU Art 82(1) identifies the basis for enacting legal instruments for harmonising of criminal procedure as 'the principle of mutual recognition of judgments'. TFEU Art 82(2), referring to the same principle, states that harmonisation of national rules on criminal procedure should be 'to the extent necessary to facilitate mutual recognition of judgments and judicial decisions and police and judicial cooperation in criminal matters having a cross-border dimension'. The mutual recognition principle prescribes that the Member States should accept and recognise one another's decisions or policies as if their own, even if they do not address the same subject in the same or similar manner, or even deliver comparable outcomes.[56] TFEU Art 82(2) exhaustively lists the topics of procedural criminal law about which the EU can introduce binding legal instruments that aim for minimum harmonisation, considering 'the differences between the systems of the Member States.'

---

[53] Stefan Braum, "Rechtsstaat' And European Criminal Law – From The End of Sovereignty' (2020) 12 New Journal of European Criminal Law 14, 16; Werner Schroeder, 'Limits to European Harmonisation of Criminal Law' [2020] 2 eucrim - The European Criminal Law Associations' Forum 144, 147.

[54] Peter Csonka and Oliver Landwehr, '10 Years After Lisbon – How "Lisbonised" Is The Substantive Criminal Law In The EU?' [2019] eucrim- The European Criminal Law Associations' Forum 261, 261-263.

[55] Helmut Satzger, 'The Harmonisation of Criminal Sanctions in the European Union - A New Approach' [2019] 2 eucrim -The European Criminal Law Associations' Forum 115, 117.

[56] Maria Fletcher, Bill C Gilmore, and Robin Lööf, *EU Criminal Law and Justice* (Edward Elgar 2010) 13

The topics specified in that provision are mutual admissibility of evidence between the Member States, minimum guarantees in criminal proceedings, and crime victims' rights. TFEU Art 82(2)(d) provides that in any other subject of procedural criminal law, harmonisation measures can only be adopted upon the EP's consent, followed by unanimous agreement in the Council of the EU (i.e., the Council of Ministers).[57]

The present study submits that approximation of the Member States' legal approaches toward criminally offensive behaviours of autonomous robots requires the EU to adopt harmonising measures related to criminal procedure. Nonetheless, considering that criminal liability for robots does not relate to any of the exhaustively listed areas under Art 82 (2), it is acknowledged that the Founding Treaties do not yet allow the introduction of any harmonising measures related to robots as something other than criminal tools. However, adopting similar measures may become justified in the future, provided that the EP and the Council of Ministers give their consent.

In relation to substantive criminal law, TFEU Art 83(1) empowers the EU to establish 'minimum rules related to definition of criminal offences and sanctions in the areas of particularly serious crime with a cross-border dimension'. These areas are exhaustively listed as 'terrorism, trafficking in human beings, sexual exploitation of women and children, illicit drug trafficking, illicit arms trafficking, money laundering, corruption, counterfeiting of means of payment, computer crime, and organised crime'. Thus far, seven out of these ten areas of 'serious crime' are addressed by EU directives.[58]

TFEU Art 83(2) enables the EU to establish substantive criminal law rules if doing so is 'essential to ensure effective implementation of an EU policy in an area already subjected to harmonisation measures'. Accordingly, the EU can introduce new criminal offences to ensure compliance with the existing EU legal instruments, if existing administrative and civil sanctions are insufficient. In 2011, the EC issued a communication outlining the principles that should guide the adoption of new substantive criminal law rules under TFEU Art 83(2).[59] There, the EC acknowledged that the introduction of new criminal offences should be subject to more stringent conditions, since criminal punishment is the *ultima ratio* (last resort) for regulating social conduct.

---

[57] Balázs Elek, 'The Connection Between Harmonising Criminal Law and the Occurrence of an Error in Law-Presented Through Criminal Offenses Against the Natural Environment' [2018] (2) ELTE Law Journal 65, 68

[58] For a list of these directives, see Csonka and Landwehr (n 54) 264.

[59] European Commission, 'Communication to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions, "Towards an EU Criminal Policy: Ensuring the Effective Implementation of EU Policies Through Criminal Law"' COM (2011) 573 final.

The communication described the conditions under which new criminal offences can be established at the EU level as follows:

> For certain unlawful acts considered particularly grave, an administrative sanction may not be a sufficiently strong response. On the same line, criminal law sanctions may be chosen when it is considered important to stress strong disapproval in order to ensure deterrence.[60]

Though the EC put forward numerous EU policy areas such as data protection, road transport, and fisheries policy as candidates for harmonisation of substantive criminal law based on TFEU Art 83(2) in the communication,[61] only two directives have been introduced thus far: the 2014 Market Abuse Directive[62] and the 2017 PIF Directive.[63]

The present study concludes that the EU can introduce 'minimum rules' that would require the Member States to criminalise offensive behaviours toward robots based on TFEU Art 83(2), but only after other harmonising measures are introduced to qualify the legal issues surrounding human-robot social interactions as one of the EU policy areas. Once it enters into force, the EU AI Act will constitute the first harmonising measure on AI and robotics, but since it does not contain any provisions on human-robot social interactions, the Act does not render the legal issues regarding social robots one of the EU policy areas.

### II.4.1.b Principle of Subsidiarity

The fact that the Founding Treaties confer shared competence on the EU regarding some subject matter does not, on its own, constitute sufficient justification for adopting harmonising measures on that subject matter at the EU level. TEU Art 5(3) states that: [64]

> Under the principle of subsidiarity, in areas which do not fall within its exclusive competence, the Union shall act only if and in so far as the objectives of the proposed action cannot be sufficiently achieved by the Member States, either at central level or at regional and local level, but can rather, by reason of the scale or effects of the proposed action, be better achieved at Union level.

The principle of subsidiarity specifies that in areas that fall under the shared or supporting competences of the EU, binding legal instruments can only be adopted so long as the outcomes intended are unaccomplishable at local, regional, and national levels, but for whatever reason, can be satisfactorily achieved at the EU level.

---

[60] ibid 11.
[61] ibid 9.
[62] Directive 2014/57/EU of the European Parliament and of the Council of 16 April 2014 on criminal sanctions for market abuse) [2014] OJ L173/179 (Market Abuse Directive).
[63] PIF is the acronym for *protection des intérêts financiers*, translated into English as 'protection of financial interests. Directive (EU) 2017/1371 of the European Parliament and of the Council of 5 July 2017 on the Fight Against Fraud to the Union's Financial Interests by Means of Criminal Law [2017] OJ L198/29 (PIF Directive).
[64] Consolidated Version of the Treaty on European Union [2016] OJ C326/13 (TEU).

The present study reiterates the claim that the introduction of harmonising legal instruments at the EU level is essential for ensuring the effectiveness of the law's treatment of robots because of the transnational dimensions of the manufacturing, distribution, and operational processes of robots and the cross-border implications of human-robot social interactions.

Like most other advanced technological artefacts, robots are often designed in one country and manufactured in another for a company located in a third country, to be eventually sold worldwide. The lack of any uniform legal approach means that all parties concerned -developers, manufacturers, distributors- must deal with multiple national laws just to put some robot into the market, making robotics a 'red tape' industry. Moreover, since the transnational enforcement of national rules is not always possible, some of the existing legal instruments introduced at national levels might lose their enforceability, making it more likely for non-compliant robots to imperil the health, safety, and security of their masters and others. Cross-border implications of human-robot social interactions are the possible activities that social robots could perform over the Internet that might influence parties located in other countries. When there are no harmonisation measures and multiple national rules, that situation may put robots' masters and third parties under challenging positions regarding the vindication of their legally protected interests in the course of social interactions with robots. Because of the transnational dimensions of the design, manufacturing, and operational processes of social robots, it is assessed that the legal instruments related to the legal issues surrounding human-robot social interactions can be most effective when implemented at the supranational level. Consequently, since similar outcomes cannot be attained at local, regional, and national levels, it is evaluated that the exercise of the EU competences would align with the principle of subsidiarity.

Neither the concepts of EU law nor the legal principles expressed by EU legal instruments are independent of those of the national legal systems of the Member States.[65] The principles of EU law are constructed through the comparative analysis of national legal systems. In the EU, most of the common legal principles and shared concepts across the national legal systems are already somewhat ascertained in studies such as the DFCR and some of these are already harmonised through various EU legal instruments, reducing the burden on the researcher.

[65] Rob van Gestel and Hans-Wolfgang Micklitz, 'Why Methods Matter in European Legal Scholarship', (2014) 20 European Law Journal 292, 297.

Moreover, considering the relative novelty of social robots and the subsequent lack legal frameworks that address human-robot social interactions, it is not necessary to review the legal systems of each Member State to project their approaches to human-robot social interactions. Instead, a systematic analysis of select national legal systems is considered sufficient for identifying the commonalities and adapting differences between various legal families within the EU and then constructing a new legal framework that is uniformly applicable across EU countries.

### II.4.2 National Legal Systems

The national legal systems chosen for comparative examination in the present study are representative of the diversity of legal thought and culture across the European continent, and of course, among the Member States of the EU. While choosing the national legal systems for comparison, the present study is inspired by the classification proposed by Zweigert and Kötz. They maintain that national legal systems of EU member countries can be grouped into legal traditions according to the simultaneous analysis of several of their characteristics: historical background, prevalent mode of legal thinking, distinctive legal institutions, sources of law, and the dominant ideology.[66]

At the most basic level, the EU and the Member States attach existential importance to the same values, 'individualism, liberalism, and personal rights' and therefore share the same dominant ideology.[67] Because of the ideological similarity, both the national legal systems of the EU countries and the supranational legal system of the EU itself are classified under the overarching 'Western Legal Tradition'.[68] Though the dominant ideology and core values are similar, the same does not hold regarding the roles assigned by these legal systems for judge-made precedents and statutory laws. There are significant differences among the legal systems of EU countries regarding their perceptions of various sources of law.

In the Common Law tradition, court decisions are primary sources of law that have precedential effects thanks to the *stare decisis* principle. In countries that follow the Common Law tradition, statutes complete the case law 'which contains the core of the law expressed through specific rules applying to specific facts'.[69]

---

[66] Konrad Zweigert and Hein Kötz, *Introduction to Comparative Law* (OUP 2011) 3; 44; 56.
[67] William Tetley, 'Mixed Jurisdictions: Common Law vs Civil Law (Codified and Uncodified)' (2000) 60 Louisiana Law Review 678, 701.
[68] Robynn Allveri, 'Protocol No 24: Fact or Fiction for EU Roma?' (2013) 21 Marmara Journal of European Studies 1, 3.
[69] Andrew Stranieri and John Zeleznikow, *Knowledge Discovery From Legal Databases* (Springer 2005) 24.

In the Civil Law tradition, court decisions are treated as secondary sources of law that have persuasive value, providing interpretations of primary sources of law.[70] In these countries, codified statutes express general principles and constitute the core of the law. They are the first sources to be investigated to ascertain what the applicable law is about any particular matter, and separate, specific statutes often fill the gaps in them.[71] Consequently, Civil Law statutes principally offer concise articulations of fundamental legal principles, while their Common Law counterparts favour precision over conciseness and provide detailed explanations of the applications of legal principles and exceptions.[72]

Countries that follow the Civil Law tradition can be divided into two families according to their historical background, prioritised principles, and technical style: the *French Legal Family* and the *German Legal Family*. Though both families are shaped around codified statutes, they are products of different historical eras and were developed to answer different needs. In the present study, the preferred naming convention is the one that identifies the Civil Law tradition with the countries whose laws and legal systems have provided the models for the formation and development of other Civil Law systems around the world.

The present study categorises the countries that follow the Common Law tradition under the same legal family. For consistency, the family of Common Law countries is also named after the country whose laws and legal institutions acted as the blueprint for other Common Law countries: the English Legal Family. In that regard, there are three prominent legal families within the EU: French, German, and English legal families. A representative country is chosen for comparison from each legal family. [73]

[70] Max Rheinstein, 'Common Law and Civil Law: An Elementary Comparison' (1952) 22 Revista Jurídica de la Universidad de Puerto Rico 90, 96.
[71] David Kosař and Lucas Lixinski, 'Domestic Judicial Design by International Human Rights Courts' (2015) 109 American Journal of International Law 713, 743.
[72] Jaakko Hussa, 'Legal Families', *Elgar Encyclopaedia of Comparative Law* (2006) 387.
[73] Mariana Pargendler, 'The Rise and Decline of Legal Families' (2012) 60 American Journal of Comparative Law 1043, 1067.

### II.4.2.a French Legal Family: Italy

Countries in the French Legal Family are influenced by the Napoleonic Code. Enacted in 1804, the Napoleonic Code is a by-product of the French Revolution, an insurrection of the lower and middle classes who longed for socio-economic security and resented privileges of the upper classes.[74] Addressed to lay people, the Napoleonic Code reflects the spirit of the French Revolution and prioritises the right of property and legal security, pushing the freedom of contract and economic liberty into the background.[75] Since the Napoleonic Code follows the structure of *Corpus Juris Civilis*, Justinian's 6th century codification of Roman Law, the French Legal Family is also referred as the Romanic Legal Family.[76]

French Legal Family includes the national legal systems of France, Italy, the Netherlands, Spain, Belgium, Luxembourg, and Portugal. Among these countries, Italy stands out because of its position at the forefront of global robotics research. With over 691 companies in the robotics industry in 2020, Italian roboticists are developing ever-more advanced robots with applications in sectors such as aerospace, agriculture, defence, healthcare, and industrial machinery.[77] According to statistics released the same year, Italy is the second largest European market for robotics manufacturing, and with 224 robots installed for every 10,000 employees, it ranks the sixth in the world.[78]

In Italian law, no legal instruments specifically address human-robot interactions; likewise, no court decisions address liability questions about the autonomous behaviours of robots. However, the Italian legal literature is rich with studies that provide detailed accounts of human-robot social interactions, focusing on the need to adopt specific legislation regarding human-robot social interactions.[79] Consequently, it is submitted that Italian law is the logical choice to represent the French Legal Family.

---

[74] Sarah Maza, 'Politics, Culture, And The Origins of the French Revolution' (1989) 61 Journal of Modern History 704, 708-710.
[75] William W Smithers, 'The Code Napoléon' (1901) 49 The American Law Register 127, 139.
[76] Liliya T Bakulina, 'The Roman Spirit' of the Code of Napoleon' (2014) 5 Mediterranean Journal of Social Sciences 477, 478.
[77] Silvia Bossi and others, 'Robotics in Italy' (Report, ENEA 2014) 9; Rafealla Zallone, 'Italy Chapter', in Alain Bensoussan and Jeremy Bensoussan (eds), *Comparative Handbook: Robotic Technologies Law* (Kindle edn, Larcier 2016) ch 9, s 9.1, sub-s 9.1.3; 'The Italian Robot Industry Shows Great Growth Potential' (*#HowToRobot*, 16 June 2021).
[78] 'Robot Density Nearly Doubled Globally' (*International Federation of Robotics*, 14 December 2021)
[79] Pericle Salvini and others, 'An Investigation on Legal Regulations for Robot Deployment In Urban Areas: A Focus On Italian Law' (2010) 24 Advanced Robotics 1901, 1915; Erika Palmerini and others, 'Guidelines on Regulating Robotics' (Report, European Commission 2014) 198.

### II.4.2.b German Legal Family: Germany

Countries in the German Legal Family are influenced by the German Civil Code. Enacted in 1900, the German Civil Code is regarded as inspired by the Napoleonic Code. However, unlike the Napoleonic Code, the German Civil Code is addressed to the literate elite and reflects the capitalist-liberal views of its time, and enshrines the economic and social ideals of the newly united imperial Germany.[80] Consequently, the German Civil Code prioritises innate economic rights such as the freedom of contract and rights to property as much as the fundamental values of justice and preservation of legal security.[81]

German Legal Family includes the national legal systems of Germany, Austria, Latvia, Estonia, Czech Republic, Lithuania, Croatia, Hungary, Slovenia, Slovakia, and Greece. Among these countries, Germany stands out as not only the largest European market for robotics, but also because of the governmental and legislative initiatives aimed at addressing the legal issues surrounding new technologies, including robots. Though no comprehensive legal framework for human-robot interactions exists in German law, several legal instruments have already been introduced in response to developments in robotics technologies, most notably concerning drones and self-driving cars.[82]

Moreover, the German Federal Government explicitly recognised that supportive economic, technical, and legal frameworks are required to enable, advance, and also to control the research, development, and deployment of robots; and several steps are taken in this direction, in terms of consortiums and working groups.[83] In the German legal literature, there are already ongoing debates on the confluence of robotics and society, with many authors arguing that some legal status for robots must be found on account of their ever-growing autonomy.[84] Consequently, German law is the logical choice to represent German Legal Family.

---

[80] Catherine Valcke, 'Comparative History and the Internal View of French, German, and English Private Law' (2006) 19 Canadian Journal of Law & Jurisprudence 133, 149-151.
[81] William W Smithers, 'The German Civil Code (Das Bürgerliche Gesetzbuch) Sources, Preparation, Adoption' (1902) 50 The American Law Register 685, 710-711.
[82] Lisa Tambornino and Dirk Lanzerath, 'Law, AI And Robotics: Germany Country Report' (Report, SIENNA 2019) 11.
[83] Andreas Lober, Tim Caesar, and Wojtek Ropel, 'Germany Chapter', in Jeremy Bensoussan and Alain Bensoussan (eds), *Comparative Handbook: Robotic Technologies Law* (Kindle edn, Larcier 2016) ch 6, s 6.2, sub-s 6.2.1.1.
[84] ibid.

### II.4.2.c English Legal Family: Ireland

Since the United Kingdom's withdrawal from the EU on 31st January 2020, only three EU countries have been members of the English Legal Family: Cyprus, Ireland, and Malta. Among these countries, Ireland stands out because of the large number of local and multinational technology companies operating in the country. With its capital city, Dublin, referred to as 'the Silicon Docks of Europe',[85] increasing numbers of enterprises in Ireland are engaged in developing robotics technologies.

In Irish law, there are no specific rules addressing the legal issues regarding social robots. Nor are there any court decisions discussing the legal implications of interactions between human beings and robots. However, Irish legal literature is actively debating the legal implications of human-robot interactions, meaning that there is no dearth of secondary sources that can be used to support the systematic analysis of the Irish legal system. These reasons, combined with the fact that the researcher's 'domestic law' is Irish law,[86] emphasise Irish law as the logical choice to represent the English Legal Family.

The legal systems of countries in the English Legal Family have close connections with one another, likely because of the cultural and linguistic similarities between the countries that belong to that legal family. To better illustrate the posture of the English Legal Family on the legal issues regarding social robots, and due to their well-documented influence on the development of the Irish legal system, the present study references UK and US laws as well.

Before 1922, Ireland was part of the UK and therefore was subject to UK law. UK law still provides the conceptual and structural basis of Irish law.[87] Though not as much as UK law, US law had well-documented effects on the development of Irish law.[88] Considering that the US is the largest common law country and the undisputed world leader in robotics development, Irish law's treatment of robots will likely be affected by the direction taken by US law.[89] There is significant interest in the legal aspects of human-robot interactions among US lawmakers and academia.

---

[85] Pamela Newenham, *Silicon Docks: The Rise of Dublin as a Global Tech Hub* (LibertiesPress 2015) 1.
[86] i.e., the researcher can examine the Irish law within its social context.
[87] Mary Mathews, 'Aspects of the Irish Law of Contract' (1972) 7 Irish Jurist 292, 293; Darius Whelan, 'Guide to Irish Law' (GlobaLex, 2019).
[88] *Conway v Irish National Teachers Organisation* [1991] 2 IR 305; *DPP v Byrne* [1994] 2 IR 236; *B v DPP* [1997] 3 IR140; *Maguire v. Ardagh* [2002] 385 1 IR 385.
[89] Morris L Cohen, 'The Common Law in the American Legal System: The Challenge of Conceptual Research' (1989) 81 Law Library Journal 13, 15.

The US legal literature admits that robotics technologies are progressing faster than the law can regulate them, it is noted that 'this time academics, researchers and legislators are still creating the building block for the design of a regulatory framework adapted to these new technologies'.[90]

## Summary

This chapter described the methodology used in carrying out the present study. Considering that the research objectives are related to both the identification and resolution of the legal issues surrounding human-robot interactions, the study does not only describe these legal issues but also conceptually formulates them in terms of their impacts on the legal values that underlie existing legal instruments.

Since social robots are highly technological artefacts with cross-border dimensions, the present study examines the legal issues surrounding them to provide solutions that can be implemented at the supranational level and, therefore, primarily focuses on EU law. To gain sufficiently deep insight into European legal thought, the study performs comparative legal research between national legal systems, chosen to represent prominent EU legal families, namely the Irish, Italian, and German legal systems. The comparative research, in turn, provides the conceptual foundations of the proposed unified legal framework.

The present study views the legal issues regarding social robots as challenges stemming from the inadequacy of existing legal instruments. It is assessed that applying existing legal instruments to the legal issues regarding social robots would not protect the legal values that constitute the purposes of these instruments. The law's shortcomings can be overcome by modifying existing legal instruments, allowing them to give effect to the distinctive characteristics of social robots,

[90] Francoise Gilbert, 'United States Chapter', in Alain Bensoussan and Jeremy Bensoussan (eds), *Comparative Handbook: Robotic Technologies Law* (Kindle edn, Larcier 2016) ch 17, s 17.1, sub-s 9.1.3.

# Chapter III: Conceptualisation of Social Robots

> A robot can be a lot of things these days—and this is just the beginning of their proliferation. With so many different kinds of robots, how do you define what one is? (…) This isn't a trivial semantic conundrum: Thinking about what a robot really is has implications for how humanity deals with the unfolding robo-revolution.
>
> —Matt Simon, *What is a Robot?* [1]

## Introduction

The concept of robot was first formulated by the science fiction genre around one hundred years ago, and the entities that the concept denotes emerged in the real world shortly after that. Despite the short history of the concept, the intellectual origins that culminated in its eventual formulation can be found across various cultures and in diverse geographies throughout history. Human beings have also repeatedly attempted to realise the visions behind the concept. These attempts, recorded by the history of automata, constitute the technological origins of the concept. Moreover, even though the field of modern robotics is relatively young, in that period, so many types of robots were developed with such significant varieties regarding their abilities, and attributes; some of them, put next to each other, are as different as apples and oranges.[2]

Maybe since the borders of 'the conceptual playing field' were determined by myths, legends, and science-fiction before modern-day robotics engineers,[3] or perhaps thanks to technological advances that continually diversify the attributes and features of robots,[4] there is no universally accepted definition for the concept of the robot that captures the distinctive characteristics of robots. Not even roboticists can agree on how to define the creations of their work.[5] So much so that Joseph Engelberger, one of the pioneers of modern robotics, once stated, 'I can't define a robot, but I know one when I see one'.[6] In that respect, the present study does not take on the elusive task of producing universally acceptable definitions for the concepts of *robot* and *social robot* that many before failed to deliver.

---

[1] Matt Simon, 'What Is a Robot?' (*Wired,* 24 August 2017).
[2] Iliah Rezah Nourbakhsh, *Robot Futures* (MIT 2013) 14-16.
[3] David J Gunkel, *Robot Rights* (Kindle edn, MIT 2018) ch 1, s 1.1.
[4] James H Wilson, What Is a Robot, Anyway? (*Harvard Business Review,* 15 April 2015).
[5] John M Jordan, *Robots* (MIT 2016) 4-5.
[6] Joseph F Engelberger, *Robots in Service* (MIT 1989) 21.

In any case, the problem regarding the apprehension of the concept of robot is not the lack of definitions. On the contrary, numerous descriptive but unhelpful definitions exist across the myriad of resources, perhaps one for every robot manufactured.[7] However, since the present study examines the legal issues surrounding human-robot social interactions, developing some means of distinguishing social robots from other functional artefacts is necessary.

Therefore**,** the following chapter focuses on the conceptualisation of social robots. The chapter is organised into three sections. In the first section, the evolution of the robot concept is explored. The section initially traces the concept's intellectual origins back to mythological and religious narratives of ancient civilisations, then evaluates the technological origins by examining the pre-industrial era's automata, and finally reviews some key works of science-fiction whose contributions helped mould the intellectual and technological origins into the beginnings of modern robotics. The second section analyses the definitions provided by dictionaries, industry organisations, and scholars for the concept of robot. Through the comparative analysis of existing definitions, the section pinpoints the distinctive characteristics of social robots. The third section, treating the distinctive characteristics of social robots identified in the second section as core components of functional characterisations of the concepts of *robot* and *social robot*, advances the present study's conceptualisation of social robots. That conceptualisation rationalises the connections between the distinctive characteristics of social robots and the inadequacies of existing legal instruments.

## III.1 Evolution of the Robot Concept

### III.1.1 Intellectual Origins: Mythological and Religious Accounts

The concept of robot did not exist in everyday language until 1923, when Karel Čapek's *Rossum's Universal Robots* was first translated into English. However, the concept's intellectual origins, ideas about the artificial creation of autonomous entities, have been expressed since early antiquity, even before the first automata were built in the Age of the Pyramids.

[7] Natelia Nevejans, 'European Civil Law Rules in Robotics' (Study, European Parliament, 2016) 7-8.

Often inspired by narratives of the original creation, many ancient civilisations produced accounts of the artificial -and magical- creation of autonomous entities, and these accounts are regarded as the earliest precursors of the concept of robot.[8] One of the best-known narratives about the original creation is found in the Book of Genesis, where the first human being is created as follows:

> Then God Eternal fashioned the man -dust from the soil- and breathed into his nostrils the breath of life; so that the man became a living soul.[9]

According to the Genesis creation account, God creates the first human being in 'the image of God'. Though the precise meaning of the notion of 'the image of God' is disputed, both the Jewish and Christian traditions concur that the notion alludes to human beings' powers of rationality and free will.[10] Combined in the same entity, the ability of rational thought and the freedom of will forms the capacity to self-manage. Therefore, it is submitted that 'the image of God' can also be construed as referring to self-management capacity. Robots can also be said to have self-management capacity thanks to the capabilities associated with the characteristic of relative autonomy.

Genesis refers to the creation of life, but it does not mention the desire of human beings to imitate that creation. However, ideas about the artificial creation of autonomous entities, inspired by the Genesis creation account and motivated by the human desire to imitate that creation, feature prominently in the Jewish folklore, most notably in the Golem legends:

> The Golem is a creature, particularly a human being, made in an artificial way by the virtue of a magic art, through the use of holy names. [11]

The most famous Golem narrative concerns the historical figure of Rabbi Loew, one of the prominent scholars of 16th-century Prague. According to that narrative, to protect the Jewish neighbourhood from anti-Semitic violence, Rabbi Loew forms the Golem out of clay and brings it into life using the esoteric knowledge of how God created the first human being.[12] However, when the Golem 'runs amok' and starts destroying everything on its path, Rabbi Loew immobilises it and stores its remains in the synagogue's attic, to be revived again when needed.[13] The narrative of the Golem addresses, however crudely, the benefits and risks associated with the artificial creation of autonomous entities.

---

[8] Gerschom Scholem, 'Golem', *Encyclopedia Judaica* (1972) vol 7, 753.
[9] Genesis 2:7.
[10] Avshalom Mizrahi, 'The Soul and the Body in the Philosophy of the Rambam' (2011) 2 Rambam Maimonides Medical Journal e0040; Chris Heren, *Imago Dei: The Doctrine of the Image of God with Regard to Human Origins* (Study, Luther Seminary 2011)10-12; 16.
[11] Moshe Idel, *Golem: Jewish Magical and Mystical Traditions on the Artificial Anthropoid* (SUNY 1990) xvii.
[12] Ludwig Blau, Joseph Jacobs and Judah David Eisenstein, 'Golem', *The Jewish Encyclopaedia* (1904) vol 6, 36-37.
[13] Alden Oreck, 'The Golem' (*Jewish Virtual Library,* 2008).

Ancient Greek mythology also features accounts of both the original and artificial creation of life. In the Greek myth of creation, Prometheus -the Titan god of forethought and the benefactor of humankind- forms the first human male out of mud, and Athena -the goddess of wisdom and war and the daughter of Zeus- breathes life into that form, thereby completing the act of original creation of life. Soon after, the king of the Olympian Pantheon, Zeus, becomes the first mythological figure to imitate that creation. He orders the blacksmith god Hephaestus to make the first human female, Pandora. Wishing to take revenge on Prometheus, Zeus then sends Pandora to Prometheus's brother, Epimetheus, with the infamous box that contains all the ills of the world.[14]

In Ancient Greek mythology, neither the creation of life nor the imitation of that creation is reserved for the divine domain. For instance, the classical myth of Pygmalion addresses the artificial creation of life by one human being. In that myth, the sculptor-king of the island of Cyprus, Pygmalion, carves a beautiful sculpture out of ivory to represent his ideal of womanhood and then falls in love with that sculpture, whom he names Galatea. In response to the prayers of King Pygmalion, Aphrodite, the goddess of sexual love and beauty and the patron deity of Cyprus, brings Galatea to life.[15] Pygmalion then marries Galatea and makes her his queen with the blessing of the goddess; they even have two children together. In that classical myth, 'with hard work and faith in divine results, the creation becomes more than the sum of artist's energies'[16] though 'divine breath' was still required to complete the act of creation.[17]

Mythological and religious narratives characterise the desire to construct autonomous entities and demonstrate the evolution of the understanding of the nature of autonomy, that it does not necessarily stem from specific physical forms but instead results from free will and rational thinking. In that regard, these narratives sow the intellectual seeds of the sense-think-act paradigm, which provides the basis of the working definitions advanced by the present study for *robot* and *social robot* concepts.

---

[14] William Smith, 'Prometheus', *Dictionary of Greek and Roman Biography and Mythology* (1849) vol 3, 544-545; 'Pandora (Greek mythology)', *Encyclopaedia Britannica* (8 February 2018).
[15] 'Pygmalion (Greek mythology)', *Encyclopaedia Britannica* (12 February 2018).
[16] Jack Burnham, *Beyond Modern Sculpture: The Effects of Science and Technology on the Sculpture of This Century* (4th edn, George Braziller 1975) xx 186.
[17] Roger Clarke, 'Asimov's Laws of Robotics: Implications for Information Technology - I' (1993) 26 Computer 53, 53.

### III.1.2 Technological Origins: Automata of the Pre-Industrial Era

Long before the emergence of modern science and preceding the development of the first modern robots by many centuries, some of the finest minds of antiquity were not just contemplating but also constructing self-operating machines, collectively known as automata. Since the idea of automation, originating in the self-operation capabilities associated with automata, is the precursor of the relative autonomy that all robots display, automata are regarded to provide technological origins of the concept of robot. Though the automata of antiquity were brilliant feats of engineering and showed the zenith of the mechanical design at the time, they, at least in the beginning, were not meant to have any practical uses; in other words, they were not intended as functional artefacts.[18]

> The first complex machines (…) were automata, by means of which he attempted to simulate nature. (…) They constituted the first step in the realisation of his dream to fly through the air like a bird, swim the sea like a fish, and to become ruler of all nature. [19]

As pointed out by the technology historian Silvio Bedini, automata were initial attempts of human beings to produce mechanisms that can mimic the patterns of movement observed in human beings and nature. Whether any practical applications were attached or merely meant for aesthetic purposes, automata of antiquity were still game-changing artefacts. While other forms of emulative illustrations available at the time, such as paintings and sculptures, could only offer static visual representations, automata could provide dynamic representations incorporating auditory and tactile elements in addition to visuality.

The earliest known examples of automata are documented to have been built in the early periods of Ancient Egypt. Though the profession of mechanical engineering is one of the secular vocations in the present time, for most of Ancient Egyptian history, activities in engineering fields, along with other areas of scientific study, including astronomy and medicine, were conducted under the auspices of temples. Ancient Egyptian worldview saw little difference between religion, magic, and technology. [20] In that regard, some of the prominent temples in Ancient Egypt were not only religious establishments but also centres of academic learning and scholarly research, in many ways, they were quite similar to today's universities.[21] For that reason, it is unsurprising that the Ancient Egyptian automata were built for religious purposes by priest-scholars.

---

[18] Isaac Asimov, *How Did We Find Out About Robots?* (Walker 1984) 19.
[19] Silvio Bedini , 'The Role of Automata in the History of Technology' (1964) 5 Technology and Culture 24.
[20] Ian Shaw, *Ancient Egyptian Technology and Innovation* (Bloomsbury 2015) 39-42.
[21] Clarke (n 17) 53; Mathieu Ossendrijver, 'Scholars, Priests, and Temples: Babylonian and Egyptian Science In Context' (2021) 8 Journal of Ancient Near Eastern History 1, 4.

These first so-called automata were closer to mechanical puppets than self-operating machines, as suggested by the term *automata*. Better described as *proto-automata*, they were, in fact, statutes of deities furnished with mechanical contraptions enabling clerics to give them motion and voice during religious ceremonies, thereby creating the impression that deities came to life, helping increase the religious devotion of the faithful.[22]

The first genuinely self-operating automata were built by Ancient Greeks, pioneers of the ancient world in engineering. Even the word *automata* comes from the Ancient Greek language, and refers to 'self-operating machines'.[23] Coincidentally, these automata were also constructed in Egypt, though later in Egyptian history, during the Ptolemaic era. In the third century BC, Ctesibius, the first head of the Library of Alexandria, applied his discoveries in hydraulics to build numerous water-powered self-operating machines, including the world's first cuckoo clock and thereby built the first 'true' automaton.[24] In the first century AD, Heron, another polymath from Alexandria, managed to build more sophisticated automata, combining the earlier findings in hydraulics with his research in mechanics and pneumatics. Also credited with designing the first-ever steam engine, the automata of Heron included a programmable cart, a self-operating water fountain, and an entirely mechanical theatre play.[25]

Shortly after the collapse of the Greco-Roman civilisation, Europe entered into its Dark Age and largely abandoned antiquity's intellectual and scientific pursuits. The classical scholarship of the ancient world was preserved and developed by scholars of the Islamic Golden Age. Among these scholars, Al-Jazari attracts attention with *The Book of Knowledge of Ingenious Mechanical Devices*, where over fifty self-operating machines, including several humanoid automata (such as a fountain with automated servants and a robotised musical band) were described along with instructions on constructing them.[26] Al-Jazari is also credited with introducing the notion of practical application into the art of constructing automata.[27]

---

[22] Nicholas Reeves, 'A Rare Mechanical Figure from Ancient Egypt' (2015) 50 Metropolitan Museum Journal 43, 48.
[23] Mark E Rosheim, *Robot Evolution: The Development of Anthrobotics* (Wiley 1994) 1.
[24] Justin Pollard and Howard Reid, *The Rise and Fall of Alexandria* (Penguin 2007) 132
[25] Evangelos Papadopoulos, 'Heron of Alexandria (c. 10-85 AD)' in Marco Ceccarelli (ed), *Distinguished Figures in Mechanism and Machine Science: Their Contributions and Legacies* (Springer 2007) 227-232.
[26] Ismail Al-Jazari, *The Book of Knowledge of Ingenious Mechanical Devices* (Donald R Hill tr, Springer 1973) 12.
[27] Mahmut Dirik, 'Al-Jazari: The Ingenious Inventor of Cybernetics and Robotics' (2020) 1 Journal of Soft Computing and Artificial Intelligence 47, 48-52.

After the Middle Ages, the revival of interest in classical scholarship and the development of clockwork mechanisms in Renaissance Europe stimulated the production of ever-more complicated self-operating mechanisms, such as Leonardo da Vinci's mechanical knight or Juanelo Turriono's mechanical monk, regarded to be particularly impressive:

> Driven by a key-wound spring, the monk walks in a square, striking his chest with his right arm, raising and lowering a small wooden cross and rosary in his left hand, turning and nodding his head, rolling his eyes, and mouthing silent obsequies. From time to time, he brings the cross to his lips and kisses it.[28]

Under the sponsorship of royalty and the aristocracy, the 18th and 19th centuries saw the construction of the most intricate automata, including miniature singing birds, animated-mechanical pictures, and most notably the Jaquet-Droz automaton –the 'forerunner of computers' that was capable of drawing pictures and writing texts.[29] Though these automata were impressive, they were prohibitively expensive and had no functions that cheaper alternatives could not fulfil.[30] Consequently, especially because of the Industrial Revolution, production of automata virtually ceased by the early 20th century. It is submitted that automata demonstrate the human drive to achieve what myths and religious narratives consider impossible without divine intervention –the creation of autonomous entities.[31]

### III.1.3 Formal Origins: Contributions of Science-Fiction

The human desire to build autonomous entities -represented by mythological and religious narratives- and the human drive to achieve the impossible -represented by automata-, despite the insufficient technical knowledge to realise that desire and the lack of socio-economic necessity to encourage the development of that knowledge; provide the intellectual and technological origins of the concept of robot. The literary contributions of the science-fiction genre, in turn, help aggregate and organise these intellectual and technological origins to formulate the concept of robot.

---

[28] Elizabeth King, 'Clockwork Prayer: A Sixteenth-Century Mechanical Monk' (2002) 1 Blackbird: An Online Journal of Literature of Arts.
[29] Sonia Kolesnikov-Jessop, 'Automatons and Ingenuity' (*New York Times,* 8 March 2012)
[30] Burnham (n 16) xx.
[31] Asimov, *About Robots?* (n 18) 19; Clarke (n 17) 54; Engelberger (n 6) 21; Rosheim (n 23) 2-6.

### III.1.3.a Mary Shelley - *Frankenstein: The Modern Prometheus*

The first critical literary contribution to the development of the robot concept is Mary Shelley's 1818 novel *Frankenstein: The Modern Prometheus*. Widely regarded as the first work in the science-fiction genre, Shelley's *Frankenstein* tells the story of the primordial mad scientist, Victor Frankenstein, who succeeds in his lifelong labour to imitate the original creation of life and constructs an artificial living entity out of body parts he gathered from cadavers. Frankenstein is horrified by the product of his labour, who is charmingly referred to as the monster, and he first verbally abuses and then abandons the monster. The monster, to whom nobody shows any kindness in the novel, despairs, vows to take bloody revenge on its creator, and goes on a murderous rampage.[32] In that respect, Shelly's *Frankenstein* can be considered a cautionary tale, warning against the 'over-reaching' of modern science that is not restrained by any ethical concerns, expressed by the title character of the novel as follows:

> Learn from me, if not by my precepts, at least by my example, how dangerous is the acquirement of knowledge and how much happier that man is who believes his native town to be the world, than he who aspires to become greater than his nature will allow.[33]

Isaac Asimov credits Shelley's *Frankenstein* with the conception of the Frankenstein complex, the prevalent fear and all-time narrative motif that once constructed, artificial autonomous entities might turn against human beings who created them at any time.[34] The present study notes that Shelley's *Frankenstein* also raises, for the first time, the question of obligations owed by creators, to their creations, especially when these creations are sentient (can feel and display emotions), as when the monster exclaims to his creator, Frankenstein:

> I am thy creature: I ought to be thy Adam; but I am rather the fallen angel, whom thou drivest from joy for no misdeed. (…) I was benevolent and good; misery made me a fiend. Make me happy, and I shall again be virtuous.[35]

Shelley's *Frankenstein* is the story of artificial life being created for no purpose but one human being's desire to imitate the original act of creation. In that regard, the present study submits that the novel's theme appears to mirror the ideas that inspired the construction of the first *automata* in antiquity.

---

[32] Mary Shelley, *Frankenstein: The Modern Prometheus* (first published 1818, Routledge 1891) 171-173.
[33] ibid 72.
[34] Asimov, *About Robots?* (n 18) 25.
[35] Shelley (n 32) 137.

### III.1.3.b Karel Čapek - *Rossum's Universal Robots*

Karel Čapek's play *Rossum's Universal Robots*[36] ['RUR'] is equally significant for developing the robot notion and especially noteworthy for introducing the word *robot* into many languages around the world, [37] Čapek's robots, derived from the Czech word *robotnik* which means meaning 'forced labour, compulsory service, drudgery',[38]are not mechanical devices with cogs and wheels. Instead, they are living entities fabricated from synthetic organic matter -made with artificial flesh and blood, like the androids in the TV series *Westworld*- and look exactly like human beings. The flesh-and-blood robots in *RUR* are intended to be exploited in hard labour; as such, they can be viewed as the modern versions of the slaves in Ancient Greek and Roman societies. Though they are more intelligent and physically more robust than human beings, the robots in *RUR* initially have 'no passion, no history, no soul'.[39]

Later in the play, the manufacturer modifies these robots as part of some strategy to prolong the service lives of robots. The manufacturer thinks that if robots were allowed to harbour feelings like fear and pain, they would be motivated to self-preserve. Subsequently, however, robots start feeling more sophisticated emotions than intended, and some of them even develop revolutionary ideas, most notably the desire for freedom and the thirst for revenge, as one of them remarks:

> You are not as strong as the Robots. You are not as skilful as the Robots. The Robots can do everything. You only give orders. You do nothing but talk.(…)
> I don't want a master. I know everything for myself.[40]

Eventually, robots in the play revolt, kill all the human beings, and build an egalitarian, peaceful, and, of course, entirely robotic society.[41] It is clear that *RUR* relies on the Frankenstein Complex as the narrative motif. Though the plot of Karel Čapek's play focuses on the threats associated with the artificial creation of autonomous entities, ig inadvertently points out the potential benefits: replacing human beings in dangerous and labour-intensive jobs.[42]

---

[36] Karel Čapek, *R.U.R: Rossum's Universal Robots* (David Wyllie tr, originally published 1920, eBooks @Adelaide 2014)
[37] Michael E Moran, 'Rossum's Universal Robots: Not the Machines' (2007) 21 Journal of Endourology 1399, 1400.
[38] Emily Temple, 'A Brief Literary History of Robots' (*Literary Hub*, 6 April 2017)
[39] Čapek (n 36).
[40] ibid.
[41] ibid.
[42] Hans Peter Moravec, 'Robot', *Encyclopaedia Britannica* (4 January 2018).

### *III.1.3.c* Isaac Asimov - *I, Robot*

Originally a scientist who worked in the field of chemistry, Isaac Asimov's literary contributions are noteworthy for challenging the then-popular Frankenstein complex that treated robots as creations of mad scientists that invariably turn on their creators. Asimov presents robots as rule-governed functional artefacts and identifies the implications of robotics may be from that perspective. Consequently, the central theme of Asimov's contributions to science fiction is the potential of robots as one of the emerging technologies that should be exploited to increase society's general welfare and prosperity.

Unlike the monster in Shelley's *Frankenstein* or living robots in Čapek's *RUR*, robots in Asimov's fiction -'Asimovian robots'- are well-engineered, non-threatening, mass-produced mechanical devices that operate as multifunctional helpers for human beings. The threat they pose to society is not extinction, but rather stagnancy and widespread lethargy.[43] Asimovian robots are devised with built-in safeguards that constrain them from knowingly harming human beings. These safeguards, referred to as the Three Laws of Robotics, are quoted as follows:

1. A robot may not injure a human being or, through inaction, allow a human being to come to harm.
2. A robot must obey the orders given it by human beings except where such orders would conflict with the First Law.
3. A robot must protect its own existence as long as such protection does not conflict with the First or Second Laws.[44]

Asimov's contributions to science-fiction also features the first reference to robotics as one of the engineering disciplines, making Asimov 'the father of robotics'.[45] Asimovian robots inspired the pioneers of modern robotics, such as Joseph Engelberger, whose company deployed the first industrial robot,[46] and Joseph Weizenbaum, who designed the first emotional chatbot.[47] Closer to the present day, Asimov's Three Laws of Robotics are cited as the intellectual basis of some EU legal instruments that aim to promote the development of ethical and trustworthy robots.[48]

---

[43] Isaac Asimov, *Robots and Empire* (Doubleday 1985) 116.
[44] Isaac Asimov, *I, Robot* (first published 1967, HarperVoyager 2018) 49-51; Cathy Lowne, 'I, Robot (work by Asimov)', *Encyclopaedia Britannica* (21 August 2017).
[45] Lowne (n 44).
[46] Dylan McGrath, 'Farewell, Joe Engelberger' (*Electronic Engineering Times,* 12 May 2015).
[47] Joseph Weizenbaum, *Computer Power and Human Reason* (Freeman 1976) 2-6; Mary Bellis, 'A Brief History of Robots' *(ThoughtCo,* 30 September 2017*)*
[48] European Parliament Resolution 2015/2103(INL) of 17 February 2017 with Recommendations to the Commission on Civil Law Rules on Robotics [2018] OJ C252/239, para T.

In summary, one of the points where the field of robotics differs from other engineering disciplines is its interaction with the science-fiction genre. Uniquely in robotics, engineers and scientists try 'to realise what has been imaginatively prototyped' in science fiction.[49] In that regard, the present study submits that various literary contributions of the science-fiction genre 'helped create public expectations and conceptions' of robots, which, in turn, inspired developments in robotics.[50]

## III.2 Existing Definitions

The introduction section of the present chapter mentioned that various sources give numerous definitions for the concept of robot. In addition to lexical definitions, industry organisations follow other definitions for the concept, and the scholarly literature also advances some working definitions. It is evaluated that existing definitions for the concept of *robot* vary significantly in terms of the intended and intended contexts of use, focal points, and conceptual scopes.

First, lexical definitions are examined to determine the commonly understood meaning of the concept of robot. Most dictionaries of the English language give multiple definitions for the *robot.* One of the leading dictionaries of American English, *Merriam-Webster Dictionary,* cites three definitions:

> 1. A machine that looks like a human being and performs various complex acts (such as walking or talking) of a human being; also: a similar but fictional machine whose lack of capacity for human emotions is often emphasised.
> 2. A device that automatically performs complicated, often repetitive tasks.
> 3. A mechanism guided by automatic controls.[51]

Following suit, *Oxford English Dictionary*, the accepted authority on British English, also gives numerous meanings for the concept of robot, mirroring the above-cited definitions.[52] In contrast, *Cambridge English Dictionary* only provides a single definition and describes the robot as 'a machine controlled by a computer, which can move and do other things that people can do'.[53] As the reference baseline, these lexical definitions agree that the concept of *robot* refers to self-operating functional artefacts. Apart from that baseline, however, these definitions do not align with one another, even within the same dictionary.

---

[49] Gunkel (n 3) ch 1, s 1.1, sub-s 1.1.1;
[50] Bryan Adams and others, 'Humanoid Robots: A New Kind of Tool' (2000) 15 IEEE Intelligent Systems 25, 25.
[51] 'robot' (*Merriam-Webster,* 2017)
[52] 'robot' (*OED Online*, 2018)
[53] 'robot' (*CED Online*, 2022)

That reference baseline is insufficient for differentiating robots from other self-operating functional artefacts, such as automated devices that repeat pre-programmed action sequences but cannot respond to changes in the environments in which they operate. Because of the significant differences among lexical definitions, it is impossible to ascertain what the concept of robot refers to in everyday language without being aware of the particular contexts in which the concept is used. Moreover, lexical definitions fail to be individually helpful as well, since the conceptual scopes of these definitions are often either too general or too specific to provide useful distinctions between robots and other functional artefacts. For example, the definition of the concept of robot as 'a machine that looks like a human being' can only refer to anthropomorphic robots. Though anthropomorphic robots -androids- are common in works of science fiction, they are exceptional in the current state of the art, pointing out that this definition is too narrow. In contrast, the definition of the concept of *robot* as 'a device that automatically performs complicated, often repetitive tasks' can refer to any computer and therefore is too general.

The use of the concept of robot by the leading organisation of the robotics industry, the International Federation of Robotics (IFR),[54] is based on the definition offered by the ISO in its now replaced standard dated 2012, quoted as follows:

> an actuated mechanism programmable in two or more axes with a degree of autonomy, moving within environments, to perform intended tasks.'[55]

The new vocabulary standard for robotics, also published by ISO in 2021, defines the concept of robot in the same manner as the 2012 standard.[56] Definitions advanced by both standards are more streamlined and less confusing than the lexical definitions. However, even though both standards also provide separate definitions for industrial and service robots, the definitions they give for the concept of robot are focused on industrial applications and physical devices. Affirmingly, different definitions offered for the term *service robot* in both standards are worded in ways that concentrate on industrial robots; the 2012 standard defines the term *service robot* as a 'robot that performs useful tasks for human beings or equipment excluding industrial automation applications'[57] (emphasis added).

---

[54] 'World Robotics 2016: Industrial Robots' (Report, IFR 2016) 25; 'World Robotics 2016: Service Robots' (Report, IFR 2016) 9.
[55] 'ISO 8373:2012 - Robots and Robotic Devices — Vocabulary' (Standard, ISO 2012)
[56] 'ISO 8373:2021 - Robotics— Vocabulary' (Standard, ISO 2021).
[57] 'ISO 8373:2012' (n 55).

Moreover, as the industry-standard definitions are intended to guide robot manufacturers active in the industry, the conceptual scopes of these definitions are limited to robots' technical features. They, therefore, confine the relative autonomy of robots to their technical features and motion abilities, but from the legal perspective, their physical movements are not the most significant abilities of robots.[58] In summary, even though the industry-standard definitions advanced by the ISO and followed by the robotics sector are straightforward, they are too concise and technical to have any practical uses outside industrial contexts and unhelpful for the purposes of the present study.

The present study, therefore, mainly relies on the scholarly literature to inspire its working definitions for the concepts of *robot* and *social robot*. In the literature, the foundational definition of the concept of *robot* is coined by the roboticist George Bekey, who described the robot as 'a machine that senses, thinks, and acts'.[59] Bekey's definition expresses -the operational paradigm, the sense-think-act paradigm- of how robots fulfil their functions. Most other definitions offered for the concept of *robot* (especially in the legal literature) rely on the paradigm offered by Bekey's definition.[60] Though the sense-think-act paradigm provides the basis for most of the scholarly definitions advanced for the concept of robot, it must be noted that there are still significant differences among various authors regarding how the paradigm should be interpreted.

One of these differences, especially consequential for the present study, is related to whether the sense-think-act paradigm requires robots to have physical embodiments. Most authors, including Bekey, argue that robots must have physical embodiments to operate under the sense-think-act paradigm.[61] They defend that since components required for following the sense-think-act paradigm, like sensors and actuators, must be physical, robots that incorporate these components must have embodiments as well.[62] Other authors (including that of the present study) suggest that functional artefacts that can perceive and respond to external stimuli independently, even if they have to rely on other devices to sense and react to these stimuli, or even when these stimuli are non-physical, should be regarded as robots.[63]

---

[58] Gunkel (n 3) ch 1, s 1.1, sub-s 1.1.2.
[59] George Bekey, *Autonomous Robots: From Biological Inspiration to Implementation and Control* (MIT 2005) 2.
[60] Neil M Richards and William D Smart, 'How Should the Law Think About Robots?' in Ryan Calo, A Michael Froomkin and Ian Kerr (eds), *Robot Law* (Edward Elgar 2016) 6.
[61] Bekey, *From Biological Inspiration* (n 59) 2.
[62] Kate Darling, 'Extending Legal Protection to Social Robots: The Effects of Anthropomorphism, Empathy, and Violent Behaviour Towards Robotic Objects' in Ryan Calo, others (eds), *Robot Law* (Edward Elgar 2016) 212-214
[63] A Michael Froomkin, 'Introduction' in Ryan Calo, A Michael Froomkin and Ian Kerr (eds), *Robot Law* (Edward Elgar 2016) xi.

Another interpretative difference regarding the definition of the concept of *robot* concerns various approaches to the sense-think-act paradigm. Some authors argue that entities that 'display physical and mental agency' from the perspective of external observers satisfy the paradigm.[64] From that point of view, it is sufficient for functional artefacts to merely appear to exercise the capabilities associated with the sense-think-act paradigm to be considered robots; they do not actually need to be able to exercise relative autonomy. According to these authors, most automated devices, such as sensor doors that open when someone approaches, can be considered robots.

In summary, the concept of *robot* refers to functional artefacts that operate according to the sense-think-act paradigm. The paradigm indicates that robots can act autonomously. However, whether the sense-think-act paradigm requires robots to have physical embodiments and the required degree of autonomy that functional artefacts need to satisfy the paradigm remains contentious.

## III.3 Distinctive Characteristics of Social Robots

Considerably more elaborate than the applications of other functional artefacts, robots' roles require them to capture environmental information, analyse that information, and plan and follow specific courses of action. Consequently, they are equipped with capabilities that other functional artefacts do not possess; in other words, robots are also different from other functional artefacts in terms of their capabilities. The concept's historical evolution points out what these capabilities are. From the relation between the concept of robot and the idea of the artificial creation of autonomous entities, the concept can be intuitively taken to presuppose the faculty of reasoning and the ability to make independent decisions.[65] Then, the concept of *robot* refers to to functional artefacts that can self-manage to fulfil whatever their assigned tasks are. The capacity of self-management, in turn, requires robots to be capable of the three processes that are expressed by the sense-think-act paradigm.[66]

---

[64] Richards and Smart (n 60) 6.
[65] Patrick Lin, 'Introduction to Robot Ethics' in Patrick Lin, Keith Abney and George A Bekey (eds), *Robot Ethics: The Ethical and Social Implications of Robotics* (MIT 2014) 4-6
[66] Mel Siegel, 'The Sense-Think-Act Paradigm Revisited' (1st International Workshop on Robotic Sensing, Orebo, June 2003)

While all functional artefacts that can perform at least one of their specified roles according to the sense-think-act paradigm can be regarded as robots, not all robots can enter into social interactions with human beings.[67] Therefore, the concept of *social robot*, for the present study, refers to functional artefacts that can not only self-manage but also emulate human social skills.[68] As the capabilities associated with emulating human social skills situate social robots as independent actors in society, the social agency is described as the characteristic that distinguishes social robots from the rest of the robots.[69]

In the following section, social robots are conceptualised through the distinctive characteristics that separate social robots from other functional artefacts. First, the characteristic of relative autonomy that social robots have in common with all robots, facilitated by the capabilities expressed by the sense-think-act paradigm, is examined. While examining the capabilities associated with relative autonomy, the present study also assesses whether robots need physical embodiments to satisfy the sense-think-act paradigm from the social, economic, and legal perspectives. Finally, the section addresses the capabilities associated with social agency, the characteristic that separates social robots from the rest of the robots.

### III.3.1 Characteristic of Relative Autonomy

Social robots, like the rest of the robots, are functional artefacts. Unlike other functional artefacts, robots can self-operate and self-manage to complete the tasks assigned to them; therefore, they do not require to be continuously controlled by human beings. The characteristic of relative autonomy indicates that robots can perceive, interpret, decide, and respond to their surroundings on their own. The characteristic explains how robots operate and fulfil their specified functions, expressed by the sense-think-act paradigm, representing the only consensus among roboticists about the characteristics of robots.

---

[67] Kerstin Dautenhahn and others, 'Socially Intelligent Agents: Creating Relationships with Computers and Robots' in Kerstin Dautenhahn and others (eds), *Socially Intelligent Agents* (Kluwer 2002) 4; Per Persson, Jarmo Laaksolahti and Peter Lönnqvist, 'Understanding Social Intelligence' in Kerstin Dautenhahn and others (eds), *Socially Intelligent Agents* (Kluwer 2002) 22-26.

[68] Darling, 'Extending Legal Protection to Social Robots' (n 62) 215; Marco Nørskov, 'Editor's Preface' in Marco Nørskov (ed), *Social Robots: Boundaries, Potential, Challenges* (Taylor & Francis 2018) xv.

[69] Marco Nørskov, 'Technological Dangers and the Potential of Human–Robot Interaction: A Philosophical Investigation of Fundamental Epistemological Mechanisms of Discrimination' in Marco Nørskov (ed), *Social Robots: Boundaries, Potential, Challenges* (Taylor & Francis 2018) 99; Matthew E Gladden, 'The Diffuse Intelligent Other: An Ontology of Nonlocalizable Robots as Moral and Legal Actors' in Marco Nørskov (ed), *Social Robots: Boundaries, Potential, Challenges* (Taylor & Francis 2018) 175.

Human beings have full autonomy, which means they have the ability to act independently and make decisions without being influenced by external factors. This autonomy is grounded in human cognition, consciousness, and free will. Humans can choose to act or not act in a particular way, based on their own values, desires, and beliefs. In contrast, social robots have relative autonomy, which means that they have some degree of independence but are ultimately controlled by their programming and the instructions given to them by their manufacturers. Social robots can perform tasks and respond to certain stimuli, but they do not have consciousness or free will. Their actions are determined by their programming and the data they receive from their environment. In other words, while social robots can be programmed to make decisions and learn from their experiences, their actions are limited by their programming and the data they have been exposed to. Humans, on the other hand, can make decisions based on a wide range of factors, including emotions, personal experiences, and cultural and societal influences.

### III.3.1.a Sense-Think Act Paradigm

The sense-think-act paradigm refers to cognitive, evaluative, and actional processes that make the independent activities of robots possible. Though the sense-think-act paradigm was briefly examined above while explaining various definitions for the concept of *robot* provided by the literature, the paradigm deserves further attention since it expresses the capabilities associated with the characteristic of relative autonomy. The sense-think-act paradigm lists these as sensing, thinking, and acting capabilities.

First, sensing capabilities enable robots to perceive the environments in which they operate. Made possible with sensors and environmental feedback algorithms instead of sensory organs in living entities, robots' sensing capabilities can be considered equivalent to human beings' perceptive abilities to see, hear, and touch. Not restricted to robots, some rudimentary forms of sensing capabilities can be observed in photocell lights that detect darkness and automatically switch on.[70] Through sensing capabilities, robots can consider the changes in their environment without requiring formal inputs from human controllers.[71]

[70] Gunkel (n 3) ch 1, s 1.1, sub-s 1.1.2.
[71] Curtis E A Karnow, 'The Application of Traditional Tort Theory to Embodied Machine Intelligence' in Ryan Calo, A Michael Froomkin and Ian Kerr (eds), *Robot Law* (Edward Elgar 2016) 55.

Secondly, thinking capabilities allow robots to make plans to fulfil their designated functions. Through thinking capabilities, robots can subject the information they gather through sensing capabilities and the information already programmed by their manufacturers to integrated evaluative processes. Through these integrated evaluative processes, robots can decide which possible actions performed in which sequence would enable them to achieve goals that they determine according to their specified roles. Following the execution of the planned courses of action, robots can also use their thinking capabilities to assess the outcomes of their actions to increase the success of their future decisions. Therefore, it can be concluded that thinking capabilities distinguish robots from measurement or recording devices and are essential to the characteristic of relative autonomy, just as the capability of reasoning is essential to the free will of human beings. In the literature, robots' thinking capabilities are often correlated to the mental skills of human beings. Through thinking capabilities, robots are 'following complex and long-term strategies, making rational decisions, drawing logical conclusions from (…) information acquired over time, (…) and reacting sensibly in unexpected situations'.[72] Considering that the evaluative processes that robots need to engage with to display thinking capabilities are similar to those associated with the cognitive skills that make up human intelligence, such as reasoning, learning,[73] and planning,[74] these capabilities are sometimes referred to as artificial intelligence.[75] In that regard, the present study equates the notion of artificial intelligence with the thinking capabilities a robot needs to demonstrate in order to be able to operate in accordance with the sense-think-act capability. Consequently, any functional artefact that is able to perform such evaluative processes can be regarded as having an artificial intelligence component, but unless it can collect information from the environment and effect changes on the environment, it cannot be considered a robot.

---

[72] Michael Thielscher, *Reasoning Robots: The Art and Science of Programming Robotic Agents* (Springer 2005) ix.

[73] Reasoning is the derivation of logical inferences from sets of premises. These inferences are classified as deductive and inductive. Learning allows a robot to acquire new skills, expand its knowledge base, or adapt to its environment to develop its performance progressively.
See generally Arthur L Samuel, 'Some Studies in Machine Learning Using the Game of Checkers' (1959) 3 IBM Journal of Research and Development 210, 213; Joachim Hertzberg and Raja Chatila, 'AI Reasoning Methods for Robotics' in Bruno Siciliano and Oussama Khatib (eds), *Handbook of Robotics* (Springer 2008) 210.

[74] Planning is the computation of the tasks through 'a systematic search through a range of possible actions'.
See Marc Hanheide and others, 'Robot task planning and explanation in open and uncertain worlds' (2017) 247 Artificial Intelligence 119.

[75] Brian Jack Copeland, 'Artificial Intelligence', *Encyclopaedia Britannica* (17 August 2018).

In the third place, acting capabilities permit robots to effect changes in the environment in which they operate without any human interference. Through sensing capabilities, robots sense the environment in which they operate. With thinking capabilities, they analyse the sensory information to develop strategies for possible courses of action. By definition, robots must also be able to carry out these strategies, and acting capabilities, completing the other two sets of capabilities, refers to robots' abilities to perform goal-directed, responsive actions.[76] Unresponsive devices that repeat pre-programmed action sequences do not possess any acting capabilities, and cannot be regarded as robots.[77]

In the literature, the term *agency* is often used to refer to acting capabilities.[78] On the assumption that the sensing and thinking capabilities described above make robots autonomous, it can be inferred that all robots, by definition, are artificial autonomous agents. If robots are characterised as artificial autonomous agents that operate according to the sense-think-act paradigm, it can be said functional artefacts that do not have all three sets of capabilities denoted by the paradigm cannot be regarded as robots.[79] The first so-called robots, referred to in the first chapter of the present study, do not fall under the definition of the concept of robot adopted by the present study; instead, it is assessed that they would be more accurately described as assembly line automata.

Although these first so-called robots could operate without any external intervention from human beings, these operations were restricted to repeatedly performing identical action sequences and had no sensing or thinking capabilities.[80] In the same vein, teleoperated drones of the present day cannot be regarded as robots since their human pilots remotely control them at every stage of their operations, such as flying, landing and so on. In contrast to teleoperated drones, some weapon systems, such as fire-and-forget missiles and armed microdrones, can be considered robots.[81]

---

[76] Stuart J Russell and Peter Norvig, *Artificial Intelligence: A Modern Approach* (Prentice Hall 1995) 35;
[77] Samir Chopra and Laurence F White, *A Legal Theory for Autonomous Artificial Agents* (UM 2011) 9.
[78] Etymologically, 'agency' is derived from Latin *agentem*, the present participle of the verb *agere*, which means 'to act, to play' in English. In turn, 'agent' refers to any entity that is capable of acting.
See Hensleigh Wedgwood, 'Agent', *A Dictionary of English Etymology* (2nd edn, 1872) 17; Ernest Klein, 'Agent', *A Comprehensive Etymological Dictionary of the English Language* (8th edn, 1971) 11; Marcus Schlosser 'Agency', *The Stanford Encyclopaedia of Philosophy* (Fall edn, 2015).
[79] Joseph A Angelo, *Robotic: A Reference Guide To The New Technology* (Greenwood 2007) 255.
[80] Stan Gibilisco, *Concise Encyclopaedia of Robotics* (McGraw-Hill 2002) 11.
[81] Bonnie Docherty, 'Mind the Gap: The Lack of Accountability for Killer Robots' (Report, Human Rights Watch 2015)

For any weapon system to be considered robotic, no human being should be in the loop while that system is fulfilling its function; in other words, the system should be able to function autonomously once switched on.

1. Fire-and-forget missiles can pursue their targets without any direct instruction from their pilots,[82] fire upon their targets using their seekers after receiving initial instructions, and they do not even need to be updated with continuous information on their target.[83] Since they operate according to the sense-think-act paradigm, they can be considered robots. Similarly, heat-seeking missiles, once they are activated, can identify their targets without further human intervention. Though they do not display the same level of precision as fire-and-forget missiles, heat-seeking missiles still operate in line with the sense-think-act paradigm and therefore can be regarded as robots.
2. Advances in swarm robotics have produced armed micro-drones, 'flocks of small unmanned aerial vehicles that can move and act as a group with only limited human intervention'.[84] These deadly micro-drones can perceive the environments in which they operate, process these perceptions, and communicate with each other with little-to-no human intervention.[85] Since these armed micro-drones can act on their environment on their own, and when together demonstrate 'emergent intelligence', the present study holds that they can be considered robots as well.[86]

### III.3.1.b Question of Physical Embodiment

One of the constraints on the concept of robot, imposed by the mythological and religious accounts, supplemented by the literary contributions of the science-fiction genre, is the presumption that robots must have physical embodiments. The fact that automata of the pre-industrial era (fore-bearers of modern robots) were physical artefacts must have strengthened this presumption.

[82] UN Institute for Disarmament Research, 'The Weaponization of Increasingly Autonomous Technologies: Considering how "Meaningful Human Control" might move the discussion forward' (Report, UNIDIR 2014) 3.
[83] John Harris and Nathan Slegers, 'Performance of a Fire-and-Forget Anti-Tank Missile with a Damaged Wing' (2009) 50 Mathematical and Computer Modelling 292, 293.
[84] Irving Lachow, 'The Upside and Downside of Swarming Drones' (2017) 73 Bulletin of the Atomic Scientists 96, 96.
[85] ibid 97.
[86] Evan Ackerman, 'Lethal Microdrones, Dystopian Futures, and the Autonomous Weapons Debate' (*IEEE Spectrum,* 17 November 2017)

Though the present study admits that the presumption that robots must have physical embodiments was justified for the pre-Internet era, it is submitted that the paradigm has become somewhat of dogma in the present time.[87] Before the emergence of the Internet, there was no environment other than the physical one where everyday interactions could take place. Personal computers have existed since the 1970s, decades before the Internet, but they were not interconnected. Though some networking activities between personal computers were possible, these activities were only utilised for exchanging files and not for performing transactions that are embedded in the ordinary course of everyday life. For that reason, the present study agrees that before the popularisation of the Internet, non-physical software entities, regardless of their autonomy levels, could not have impacted anything else except the computers on which they were installed. These software entities could have been equipped with sensing and thinking capabilities, but since the physical environment was the only connective and interactive medium back then, it was impossible to equip them with acting capabilities. Without any acting capabilities, no functional artefact can satisfy the sense-think-act paradigm, and therefore cannot be considered robotic. Roger Clarke, alluding to that point in an article published just before the Internet became popular, states:

> The term robotics (…) refers to a science or art involving both artificial intelligence (to reason) and mechanical engineering (to perform physical acts suggested by reason (…) robots exhibit (…) mechanical capability, enabling it to act on its environment rather than merely function as a data processing or computational device.[88]

Today, some non-physical software entities can impact others' emotional, financial, and social states through the Internet. Be that as it may, the presumption that robots must have physical embodiments is still embraced across various resources. For example, as explored above, most lexical definitions continue to offer meanings implying that robots must have physical embodiments.

Moreover, many contemporary authors also highlight the physical embodiment as one of the essential attributes of robots.[89] For example, the prominent robot ethicist Patrick Lin stipulates that robots must be situated in the physical world in order to distinguish physical robots from software agents that are running on computers.[90]

---

[87] Froomkin (n 63) xi.
[88] Clarke (n 17) 54.
[89] Darling, 'Extending Legal Protection to Social Robots' (n 62) 220, Kerstin Dautenhahn, Terrence Fong and Illah Nourbakhsh, 'A Survey of Socially Interactive robots' (2003) 42 Robotics and Autonomous Systems 143, 146.
[90] Lin (n 65)18.

Similarly, Bekey, the roboticist who came up with the sense-think-act paradigm, emphasises that robots must be understood as physical self-operating devices.[91]

Though it appears the presumption that robots must be physical devices is still broadly accepted, the present study asserts that the acceptance afforded to that presumption is no longer rational, justified, or warranted because of the prevalence of the world wide web as the primary medium of communication where most of the commercial, social, and cultural interactions take place.[92] Indeed, because of the effectiveness and speed provided by online tools, most offline methods for banking, communicating, shopping, and even entertainment are already retired and replaced.[93] Even though human beings still live in the physical world, integrating the Internet into everyday life created another, *digital,* world where various individual interests can be pursued. The present study observes that the digital world is now one of the central parts of the functioning of the economy and society.[94] In other words, from the economic and social perspectives, non-physical software entities' impacts are comparable to those of physical entities.

From the legal perspective, where online tools replace offline methods, though activities become more accessible and faster, outcomes that can be achieved through these activities do not differ. For example, the online tools make it possible to perform banking transactions without having to wait to stand in lines, and transactions carried out through these tools have the same effects as transactions performed in the physical branches of banks.[95]

That being said, it is acknowledged that there will likely be differences regarding the undesirable outcomes of these activities. While non-physical software entities can only cause non-physical damages, such as economic loss and psychiatric injuries, physical devices can cause all sorts of harm. Nevertheless, although varying assessment methods may be utilised to determine the amount of compensation awarded for different types of damages, both types of damages are addressed within the same legal paradigm.[96]

---

[91] Bekey, *From Biological Inspiration* (n 59) 6.
[92] Manuel Castells, 'The Impact of the Internet on Society: A Global Perspective' in *19 Key Essays on How Internet is Changing Our Lives* (BBVA 2014) 132; Zaryn Dentzel, 'How The Internet Has Changed Everyday Life', in *19 Key Essays on How Internet is Changing Our Lives* (BBVA 2014) 241.
[93] Barry Wellman and others, 'The Internet In Everyday Life' (EURICOM Conference on e-Democracy, Nijmegen, October 2002) 4.
[94] Wenhong Chen, Jeffrey Boase and Barry Wellman, 'The Global Villagers: Comparing Internet Users and Uses Around The World' in Barry Wellman and Caroline Haythornthwaite (eds), *The Internet in Everyday Life* (Blackwell 2002) 80.
[95] Deborah Fallows, 'The Internet and Daily Life' (Report, PEW Research Center 2004) 12.
[96] Jack B Balkin, 'The Path of Robotics Law' (2015) 6 California Law Review Circuit 45, 50.

That legal paradigm stipulates that damages suffered by the injured party due to the defendant's breach of primary obligations should be compensated.[97]

In light of technological developments in recent years, the boundaries between the physical and digital environments have been further blurred. In the present day, it is possible to augment most of the functional artefacts used in the ordinary course of everyday life with some capacities such as identifying, sensing, and networking. To utilise these capacities, however, functional artefacts that are augmented in this manner must be connected to each other through digital means.[98] Thus, in addition to the information-based Internet of Humans, it is also possible to talk about the Internet of Things as the digital network intertwined with physical aspects with everyday life.[99] The Internet of Things is also increasingly used to improve physical robots' capabilities that enable them to operate according to the sense-think-act paradigm.[100] In most physical robots, thinking capabilities are no longer confined to their physical embodiments but instead performed by utilising the Internet of Things to connect to other robots, and accessing the Internet of Humans to employ the most up-to-date analytical algorithms. For example, in swarm robotics, the thinking capability is not possessed by individual, spatially compact robots but instead exercised through 'decentralised collective robotic networks'.[101]

In closing, in the Internet era, non-physical software entities can have impacts outside the computers they are installed on. From the economic, social, and legal perspectives, these impacts are equivalent to the impacts that physical entities can produce. As non-physical software entities can now be as effective in everyday life as physical entities, it must be accepted that they can now have acting capabilities.

So as long as they also exhibit the other two capabilities associated with the characteristic of relative autonomy, the present study submits that non-physical software entities can be regarded as robots. Though the requirement of a physical embodiment may still constitute an essential distinction for the definition of *robot* from engineering perspectives, it is assessed to be arbitrary for a study that focuses on legal issues surrounding human-robot social interactions.

---

[97] Ryan Calo, "Robotics and the Lessons of Cyberlaw' (2015) 103 California Law Review 513, 529.
[98] Samuel Greengard, *The Internet of Things* (MIT 2015) 15-17.
[99] Adam Henschke, 'The Internet of Things and Dual Layers of Ethical Concern' in Patrick Lin, Ryan Jenkins and Keith Abney (eds) *Robot Ethics 2.0: From Autonomous Cars to Artificial Intelligence* (OUP 2017) 232.
[100] Ibid 234.
[101] Gladden (n 69) 181.

### III.3.2 Characteristic of Social Agency

In the present study, the concept of *social robot* refers to specialised service robots that can interact with human beings in socially meaningful ways. In other words, social robots are defined as artificial autonomous agents that can communicate, comprehend and even connect with human beings on personal levels.[102] In addition to the capabilities associated with the characteristic of relative autonomy that they share with all robots, social robots have some specific capabilities empowering them to simulate human social skills. Through these specific capabilities that social robots can 'recognise different human beings, possess relationship histories, interpret human beings' actions and respond to them in a manner consistent with the social context'.[103] thanks to the characteristic of *social agency* facilitated by these specific capabilities, social robots are situated in the ordinary course of everyday life as independent actors, and they can assume social roles that go beyond being mere tools of human activities.

Defining the characteristic of social agency is not straightforward since it requires considering both objective properties and subjective responses. Objectively, social robots are functional artefacts designed to interact with humans socially and emotionally, using various sensors and algorithms to simulate social behaviours. These objective properties can include features such as facial expressions, body language, voice recognition, and natural language processing, which enable robots to interact with humans in a way that simulates social interactions.

However, subjective responses are also crucial in classifying a functional artefact as a social robot. This is because the perception of a robot's social capabilities is often based on the subjective experiences of human beings interacting with it. A robot that meets all of the objective criteria for being a social robot may not be perceived as such if human beings do not respond to it in social ways or find it difficult to relate to. For example, a robot with advanced AI and facial expression capabilities may be designed to simulate social interactions with humans. However, if humans do not respond to the robot socially, such as by making eye contact or engaging in conversation, it may not be perceived as a social robot. Conversely, a robot with limited social capabilities, such as a simple chatbot, may be perceived as a social robot if humans respond to it socially.

[102] Cynthia Breazeal, *Designing Sociable Robots* (MIT 2002) 19.
[103] Gunkel (n 3) ch 2, s 2.6; Matthias Scheutz, 'The Inherent Dangers of Unidirectional Emotional Bonds between Human beings and Social Robots' in Patrick Lin, Keith Abney and George A Bekey (eds), *Robot Ethics: The Ethical and Social Implications of Robotics* (MIT 2014) 205;

Therefore, the objective properties of social robots and the subjective responses of human beings are both essential components of the characteristic of social agency. Objective properties provide the framework for designing and evaluating the social capabilities of robots, while subjective responses reflect how humans interact with and perceive robots.

The present study assesses that the capabilities of social agency require social robots to exhibit several features when interacting socially with human beings.[104] These features are described as readability, believability, and understandability.[105]

First, readability stems from the capacity of social robots to provide socially indicative signals for human beings with whom they interact, prompting these human beings to predict social robots' activities by projecting emotional and mental states associated with human beings on them.[106] That being the case, to exhibit readability, the modalities through which social robots express themselves must be suitable to be effortlessly and intuitively comprehended by human beings. Breazeal explains the feature of readability as follows:

> the robot's outwardly observable behaviour must serve as an accurate window to its underlying computational processes, and these in turn must be well matched to the person's social interpretations and expectations.[107]

This feature allows the human interactant to respond to the conduct of these robots.

The conduct of social robots must be believable as well. The feature of believability does not require social robots to resemble human beings or other living entities physically but instead refers to social robots' capabilities to demonstrate unique social personalities. Believability is determined by whether robots' human interactants are able and inclined to predict, understand, and explain the conduct of robots they interact with and attribute mental states to them that correspond to familiar social terms or, in other words, assess robots from the intentional stance. In cases where social robots are not able to provoke the human interactants to situate them as independent social actors through the physical attributes, they can still do so by employing various communicative aptitudes in appropriate manners, such as emotive expressions, sound intonation, use of everyday language, and display of 'playful antics'.[108]

---

[104] Julia Knifka, 'On the Significance of Understanding Human-Robot Interaction' in Marco Nørskov (ed), *Social Robots: Boundaries, Potential, Challenges* (Taylor & Francis 2018) 3.
[105] ibid 8; Breazeal (n 102) 8-11; Dautenhahn, Fong and Nourbakhsh (n 89) 160.
[106] Daniel Carton, Wiktor Olszowy, and Dirk Wollherr, 'Measuring The Effectiveness of Readability for Mobile Robot Locomotion' (2016) 8 International Journal of Social Robotics 721, 724-727.
[107] Breazeal (n 102) 9.
[108] ibid 10. Ultimately, no one interacts socially with wax sculptures even though they physically resemble human beings.

There is another side to the emulation of human sociability. In order to respond to human interactants' socially indicative signals in like manner, social robots must apprehend human interactants' socially indicative signals and understand what is communicated by such signals. Social robots cannot sufficiently apprehend human beings' socially indicative signals solely through communicative aptitudes. They must be aware of the situational, personal, and historical contexts that shape communicative aptitudes.[109] Thus, social robots should not just be able to individually recognise their human interactants and process socially indicative signals of these interactants; they also must be able to evaluate their perceptions using their existing knowledge of their human interactants' personality traits, relevant cultural norms, and the circumstances of surrounding environments.[110]

Finally, to interact with human beings in intuitively and naturally, robots must be able to understand themselves and 'express their own internal states in social terms.'[111] The first social robot was Kismet, named after the Turkish word meaning fortune. Kismet was built in 1998 at the Massachusetts Institute of Technology by the team of roboticists led by Breazeal as part of the larger project on developing open-ended learning systems for robots.[112] Kismet was equipped with optical and kinetic sensors allowing it to recognise various socially indicative signals and thanks to its motor systems, it could display various emotions.

More critical to Kismet's design, however, were its motivational and emotional drives. Through these drives, Kismet could understand the meanings of socially indicative behaviours of its human interactants and decide on appropriate emotional responses, these drives essentially allowed it to communicate whether it was under or over-stimulated. In many ways, however, Kismet had social skills equivalent to one-year-old human babies. The emotions it could display were limited to 'happiness, sadness, anger, calmness, surprise, disgust, tiredness, and the state of sleep' and it was only able to display one emotion at one time.[113] The hope was that Kismet could build upon these primary emotional responses after it was switched on (or *born)*, learn about the surrounding world on its own, and perhaps one day become *sentient*. [114]

---

[109] Simon Baron-Cohen, *Mindblindness: An Essay on Autism and Theory of Mind* (MIT 1995) 154.
[110] Knifka (n 104) 8.
[111] ibid.
[112] Breazeal (n 102) 112.
[113] Lucy Suchman, 'Subject Objects' (2011) 12 Feminist Theory 119, 127.
[114] Duncan Graham Rowe, 'Meet Kismet…' (*New Scientist*, 22 August 1998)

Though social interactions with Kismet could not go beyond infant-caretaker relationships, in the two decades since the development of Kismet, social robotics have made significant advances. Thanks to these advances, the specific capabilities associated with the characteristic of social agency evolved beyond the experimental stage and have become the cornerstones of multiple applications, especially in education, entertainment, and health. For example, NAO -a social robot with humanoid appearance- has been widely deployed in care-homes and schools, delivering especially successful results while working with children with autistic spectrum disorders[115] Similarly, PLEO -a robotic baby dinosaur- have been successfully used to reduce depression and anger when working with children between the ages of 8-12.[116]

The application of social robotics is not limited to children. Various social robots have been deployed to help specific groups of vulnerable grown-ups 'through interaction, motivation, monitoring, and coaching in health and education'.[117] For example, Paro, -a robotic baby harp seal- has been deployed in care homes for the elderly and assisted living facilities to calm down distressed people, such as dementia patients, people with cognitive impairments, and even earthquake victims.[118] Similar to the therapeutic use of pets, Paro is noted to give distressed people some sense of empowerment.[119]

There are also non-physical social robots -autonomous software entities designed to interact socially with human beings through the Internet. The best-known non-physical social robots are emotional chatbots that can receive user inputs in natural language and respond in the same manner.[120] Though these are ultimately just lines of codes that are run by computers, some emotional chatbots can also learn from past interactions with their human interactants, thereby displaying unique personalities. One of these chatbots, Woebot, emulates cognitive behavioural therapists and helps treat depression disorders, whereas another chatbot, Replika, purports to alleviate loneliness and anxiety by imitating the expressional style of its human interactant.[121]

---

[115] Kate Darling, 'Who's Johnny?" Anthropomorphic Framing in Human–Robot Interaction, Integration, and Policy' in Patrick Lin, Ryan Jenkins and Keith Abney (eds) *Robot Ethics 2.0: From Autonomous Cars to Artificial Intelligence* (OUP 2017) 175.

[116] Joseline Raja Vora and others, 'Influence of a Socially Assistive Robot on Activity, Social Play Behavior, and Toy-Use Behaviors of Children in a Free Play Environment: A Within-Subjects Study' (2021) 8 Frontiers in Robotics and AI.

[117] Darling, 'Who's Johnny?' (n 115) 176.

[118] Moojan Ghafurian, Jesse Hoey, and Kerstin Dautenhahn, 'Social Robots for the Care of Persons with Dementia' (2021) 10 ACM Transactions on Human-Robot Interaction 1, 3.

[119] Andrew Griffiths, 'How Paro the Robot Seal is Being Used to Help UK Dementia Patient' (*The Guardian,* 8 July 2014)

[120] Barbara Ondrisek, 'Privacy and Data Security of Chatbots' (*Medium*, 27 October 2016)

[121] Rebecca Ruiz, 'This Adorable Chatbot Wants to Talk about Your Mental Health' (*Mashable UK*, 8 June 2017); Arielle Pardes, 'The Emotional Chatbots are Here to Probe Our Feelings' (*Wired*, 31 January 2018)

Sentience, the capacity to experience subjective sensations, such as pain, pleasure, emotions, and consciousness, is not considered relevant for the conceptualisation of social robots. As explored in the following chapter, existing legal instruments afford varying degrees of protection to non-human entities against offensive behaviours. However, in none of the selected legal systems does the legal protection afforded to non-human entities require them to be capable of truly going through any subjective sensations, such as pain or experiencing emotions like joy or sorrow. Rather than focusing on the concept of sentience, legal protection afforded to different non-human entities invariably focus on the extent to which the entity in question can emulate or display such subjective sensations in manners similar to human beings.

## Summary

The emergence of robots in science-fiction before real life, along with the multi-directional developments in robotics over the last six decades, makes it difficult to define the concepts of *robot* and *social robot.* Therefore, it was necessary to explore the attributes of social robots to characterise the capabilities that separate robots from other functional artefacts and those that distinguish social robots from the rest of the robots.

The concept of *robot* finds its intellectual origins in mythological and religious narratives and its technological origins in automata of the pre-industrial era. The works of science-fiction are credited with bringing these origins together to form the concept. From the historical evolution of the concept, it is inferred that *robot* refers to specific functional artefacts with advanced cognitive capabilities and sometimes physical prowess beyond human beings. The review of various definitions for the term *robot* suggests that robots differ from other functional artefacts because of the characteristic of relative autonomy, which is facilitated by the capabilities expressed by the sense-think-act paradigm. The paradigm explains that robots must be able to sense the changes in their surroundings, make their own decisions through integrated evaluative processes, and act in their environments by executing these decisions. In the Internet era, the physical embodiment is no longer a prerequisite for acting capabilities since non-physical software entities can now be as effective in everyday life as the physical devices.

The concept of *social robot* refers to specific robots that can interact socially with human beings and are thus equipped with capabilities that allow them to simulate human social skills. Since these additional capabilities situate social robots as independent social actors, they can be associated with the characteristic of social agency, the characteristic that distinguishes social robots from the rest of the robots.

# PART B

# *Human-esque*

## Implications of Robot Anthropomorphism

There is an universal tendency among mankind to conceive all beings like themselves, and to transfer to every object, those qualities, with which they are familiarly acquainted, and of which they are intimately conscious. We find human faces in the moon, armies in the clouds; and by a natural propensity, if not corrected by experience and reflection, ascribe malice or good-will to everything, that hurts or pleases us.

—David Hume, *The Natural History of Religion*[*]

[*] David Hume, *The Natural History of Religion* (J J Tournaissen 1793) 12.

# Chapter IV: Challenges of Robot Anthropomorphism

> Actions such as his could come only from a robot, or from a very honourable and decent human being. But you see, you just can't differentiate between a robot and the very best of humans.
>
> —Isaac Asimov, *I, Robot*[1]

## Introduction

During the discussion of the Genesis creation account in the last chapter, it was concluded that Judeo-Christian theology considers self-management capacity as one of the core characteristics that distinguish human beings from other creatures and render human beings somehow similar to God. It is thought that the Genesis creation account reveals that self-management capacity is one of the more exclusively 'human' attributes. Naturally, the mere fact that robots -which are ultimately functional artefacts – are relatively autonomous or can display the capacity for self-management encourages human beings to attribute other human characteristics to them. [2] In short, the capabilities associated with the characteristic of relative autonomy enable robots to prompt significant levels of anthropomorphic responses from human beings.

Anthropomorphism is not some new phenomenon that is just emerging with the proliferation of robots. It is thought to be one of the oldest survival mechanisms of human beings, developed at the earlier stages of human evolution, helping the first human beings to stay alert in the wild.[3] If human beings suppose that movements they feel or sounds they hear in nature belong to possibly hostile human beings or wild animals, they are bound to be more reluctant to let their guards down.[4] Then, it can be suggested that human beings' anthropomorphic responses toward non-human entities are not exclusively prompted because of the display of self-management capacity by these entities. In that context, regardless of whether the given non-human entity can self-manage, that entity can elicit anthropomorphic responses from human beings so long as there is some display of human-like attributes.[5]

---

[1] Isaac Asimov, *I, Robot* (first published 1967, HarperVoyager 2018) 204.
[2] Marco Aurélio Castro, *Robotic Law: The Legal Personality of the Robot* (Kindle edn, 2019) ch 3, s 3.5 sub-s 5.2.4.
[3] Marco A C Varella, 'The Biology and Evolution of the Three Psychological Tendencies to Anthropomorphize Biology and Evolution' (2018) 9 Frontiers in Psychology.
[4] Mike Dacey, 'Anthropomorphism as Cognitive Bias' (2017) 84 Philosophy of Science 1152, 1152-1154.
[5] Cédric Sueur, Marie-Amélie Forin-Wiart, and Marie Pelé, 'Are They Really Trying To Save Their Buddy? The Anthropomorphism of Animal Epimeletic Behaviours' (2020) 10 Animals 2323, 2-3.

Ordinarily, human beings do not wilfully decide to anthropomorphise non-human entities following mindful thinking processes. Instead, anthropomorphic responses are often reflexive and bypass the conscious evaluation procedures that perceptions would typically be subjected to before any voluntary reaction. The extent and strength of anthropomorphic responses by a non-human entity are proportional to the degree to which that non-human entity displays attributes and traits considered human.[6] In other words, the intensity of anthropomorphism is determined by how 'human' is the entity in question.

Although anthropomorphism is not restricted to human-robot interactions, anthropomorphic responses invoked in such interactions can be substantial even concerning non-social robots since they display the capacity to self-manage, regarded as one of the most exclusively human attributes. Anthropomorphic responses elicited by social robots can expected to be even more robust and last longer since social robots are also equipped with the additional capabilities associated with the characteristic of social agency. Since anthropomorphism involves the intuitive perception of non-human entities as similar to human beings, the phenomenon inevitably influences how human beings behave in social contexts.[7] Social robots are readily perceived as social actors because of the significant levels anthropomorphic responses they invoke.

The law, in essence, regulates the social behaviours of human beings to fulfil its functions. Consequently, the present study admits that anthropomorphic responses that robots -especially social robots- invoke, and the impacts of such anthropomorphic responses on human beings' social behaviours can pose unique challenges to the law's function. In that respect, the following chapter explores the legal challenges arising from the phenomenon of robot anthropomorphism.

The chapter is organised into three sections. In the first section, the phenomenon of robot anthropomorphism is introduced, and the variables that affect the phenomenon's intensity are described by utilising various psychological experiments. The second section focuses on undesirable outcomes of the phenomenon and discusses whether any of these outcomes relate to the legal values that underlie existing legal instruments. Finally, the third section examines whether the extent and strength of robot anthropomorphism justify any legal protection being extended to at least some robots.

---

[6] Gabriella Airenti, 'The Development of Anthropomorphism in Interaction: Intersubjectivity, Imagination, and Theory Of Mind' (2018) 9 Frontiers in Psychology.

[7] Lingyao Yuan and Alan R Dennis, 'Acting Like Humans? Anthropomorphism and Consumer's Willingness to Pay in Electronic Commerce' (2019) 36 Journal of Management Information Systems 450, 454.

## IV.1 Robot Anthropomorphism

The following section describes the phenomenon of anthropomorphism and explores the factors that affect the phenomenon's occurrence concerning robots. The section discusses the phenomenon's relationship with the distinctive characteristics of social robots and provides the primary concepts necessary for developing an in-depth understanding of the legal significance of the phenomenon. To that end, the section is organised into two sub-sections. The first sub-section provides general explanations using the available anecdotal and theoretical evidence to establish basic rational connections between the phenomenon and the distinctive characteristics of social robots. The second sub-section supports these general explanations through the findings of various psychological experiments on the extent and strength of anthropomorphic responses invoked by social robots.

### IV.1.1 Recorded Experiences

The phenomenon of anthropomorphism refers to the attribution of human characteristics, such as emotions, intentions, sentiments, and even intelligence, to non-human entities. Non-human entities that can invoke anthropomorphic responses include living entities but are in no way limited to them: natural objects or human-made artefacts can also prompt the occurrence of the phenomenon.[8] Human beings tend to anthropomorphise specific non-human entities intuitively and reflexively, without conscious thought. In other words, anthropomorphic responses are automatically prompted when human beings encounter human-like attributes or behaviours in non-human entities that they cannot immediately and intuitively understand using the knowledge they have at hand.[9] Non-human entities need to have neither any humanoid physical attributes nor even any form of embodiment to prompt anthropomorphic responses, even non-physical software entities can be anthropomorphised so long as they display some human-like attributes, such as some social interaction skills.[10]

---

[8] Esmeralda G Urquiza-Haas and Kurt Kotrschal, 'The Mind Behind Anthropomorphic Thinking: Attribution of Mental States to Other Species' (2015) 109 Animal Behaviour 167, 168.

[9] ibid 172.

[10] Human beings also anthropomorphise virtual objects: 'In the video game Portal, for example, when players are required to incinerate the companion cube that has accompanied them throughout the game, some will opt to sacrifice themselves rather than the cube, forfeiting their victory.' Kate Darling, 'Extending Legal Protection to Social Robots: The Effects of Anthropomorphism, Empathy, and Violent Behavior Towards Robotic Objects' in Ryan Calo, A Michael Froomkin and Ian Kerr (eds), *Robot Law* (Edward Elgar 2016) 216.

The phenomenon's occurrence regarding non-physical entities is not new, and its intuitive occurrence toward non-physical software entities dates back to the 1960s when human beings who interacted with it were recorded to anthropomorphise ELIZA, the psychotherapist software. That software's only human-like attributes were the natural language texts that it was able to display, and it did not even have the capabilities of relative autonomy.[11]

Anthropomorphism is not binary but instead incremental. The extent and strength of anthropomorphic responses prompted by non-human entities are proportional to the degree that these entities emulate human beings' behaviours or have human-like attributes. For example, non-human entities with human-like physical features (such as tree barks that are shaped like human faces) invoke weaker anthropomorphic responses than those that can emulate human beings' communication and social interaction skills. The capabilities that are shared by all robots, those associated with the characteristic of relative autonomy, already invoke significant levels of anthropomorphic responses from human beings on their own, as the recorded experiences and various psychological experiments explored below to confirm. However, the anthropomorphic responses prompted in that manner are often incidental and ephemeral. Since non-social robots are not intended to actively and continuously prompt anthropomorphism, anthropomorphism invoked by these robots often discontinues or weakens once human beings observe them or interact with them closer. Unlike non-social robots and because of the capabilities associated with relative autonomy and social agency in social robots, they are expected to prompt longer-lasting and more robust anthropomorphic responses. That is intentional, and by design, the additional capabilities associated with the characteristic of social agency are devised to prompt the sustained occurrence of anthropomorphic reactions.

Moreover, eliciting anthropomorphic responses appears crucial for social robots to fulfil most of their specified functions, especially regarding their roles in the education and health sectors. In these sectors, social robots are currently utilised to facilitate communication between patients and their doctors, engage children with special needs in learning activities, and help vulnerable adults by interacting with them or by motivating, monitoring, and coaching them.[12]

---

[11] Gabriella Airenti, 'The Cognitive Bases of Anthropomorphism: From Relatedness to Empathy' (2015) 7 International Journal of Social Robotics 117, 123.

[12] Cory D Kidd, Will Taggart, and Sherry Turkle, 'A Sociable Robot to Encourage Social Interaction Among the Elderly' (IEEE International Conference on Robotics and Automation, Orlando, May 2006); Evan Ackerman, 'MIT's DragonBot Evolving to Better Teach Kids' (*IEEE Spectrum,* 16 March 2015)

Though continuous anthropomorphism is the intentional result of social robots' design features, there might be some undesirable outcomes associated with that degree of anthropomorphism, including the loss of authenticity in human relationships, new types of threats to information security and privacy, potential normalisation of violent behaviours against human beings and animals because of the impunity of such behaviours toward social robots, and the potential exploitation of human beings' emotional attachments to social robots.[13]

The present study first examines the association between the phenomenon of anthropomorphism and the capabilities associated with the characteristic of relative autonomy. The recorded experiences documenting human beings' anthropomorphic responses towards various non-social robots illustrate the extent and strength of anthropomorphic responses invoked just by the capabilities of relative autonomy. For instance, the Roomba, the world-famous brand of robotic vacuum cleaners, are flat, round robots with simple algorithms allowing them to clean floors independently and have none of the capabilities associated with the characteristic of social agency whatsoever. The experiences concerning human beings' approaches to the Roomba point out that the mere fact that these robotic vacuum cleaners can move around on their own prompted the generation of substantial anthropomorphic responses toward them:

> Some will clean for the Roomba, so that it can get a rest, while others will introduce their Roomba to their parents, or bring it along when they travel because they managed to develop a (…) relationship: 'I can't imagine not having him any longer. He's my BABY.'[14]

As the Roomba are different from other non-robotic vacuum cleaners in the market just because they can self-manage, these recorded experiences establish that the capabilities of relative autonomy are sufficient for human beings to intuitively attribute human sentiments to the Roomba at such substantial levels that they even form emotional attachments toward the Roomba, even though manufacturers did not intend these robotic vacuum cleaners to invoke anthropomorphic responses.

Furthermore, considering that the specified functions of the Roomba are related to cleaning floors, anthropomorphic responses developed toward them to the degree that human beings are willing to clean for the Roomba even appear to impede the utility of acquiring these robotic vacuum cleaners.

---

[13] Kate Darling, 'Who's Johnny? Anthropomorphic Framing in Human–Robot Interaction, Integration, and Policy' in Patrick Lin, Ryan Jenkins and Keith Abney (eds), *Robot Ethics 2.0: From Autonomous Cars to Artificial Intelligence* (OUP 2017) 175.
[14] ibid.

Apart from the loss of utility, the evocation of anthropomorphic responses by non-social robots does not appear to have any significant adverse outcomes on human beings' legally meaningful personal interests in the case of the Roomba. Still, it is thought that life-threatening situations may arise when non-social robots are deployed to help or replace human beings in the performance of dangerous tasks. In that respect, in the literature, several authors have documented the recorded experiences of soldiers who utilise bomb-defusing and mine-cleaning robots on battlefields. Again because of the capabilities of relative autonomy, soldiers who utilised these robots first tended to project human emotions and sentiments to them and then develop emotional attachments toward them. Anthropomorphic responses invoked in these contexts were so strong that some soldiers were willing to risk injury to rescue these non-social robots they saw as comrades-in-arms.[15]

Social robots, in addition to the relative autonomy shared by all robots, are also equipped with specific capabilities associated with the characteristic of social agency. These additional capabilities enable social robots to display socially indicative signals that human beings automatically and subconsciously associate with human emotional and mental states. In short, thanks to the characteristic of social agency, social robots can actively engage with the ingrained anthropomorphic tendencies of human beings with whom they enter into interactions instead of incidentally and ephemerally eliciting anthropomorphic responses like other non-human entities. Some algorithms in the market today have natural language processing capabilities that are so sophisticated that these algorithms cannot be distinguished from human beings in everyday conversations. Moreover, most social robots can recognise human emotions, display meaningful emotional expressions, and perform behaviours that human beings consider to be coherent, intentional, responsive, and emotionally appropriate.[16] Sherry Turkle describes the social interactivity facilitated by the additional capabilities of social robots as follows:

> Computers no longer wait for human beings to project meaning onto them. Now, sociable robots meet our gaze, speak to us, and learn to recognise us. They ask us to take care of them; in response, we imagine that they might care for us in return.[17]

---

[15] Doree Armstrong, 'Emotional Attachment To Robots Could Affect Outcome On Battlefield' (*UW News*, 17 September 2013)

[16] Darling, 'Extending Legal Protection to Social Robots' (n 10) 218; Kerstin Dautenhahn and others, 'Socially Intelligent Agents: Creating Relationships with Computers and Robots' in Kerstin Dautenhahn and others (eds), *Socially Intelligent Agents* (Kluwer 2002) 6; Olivia Solon, 'Google's Robot Assistant Now Makes Eerily Lifelike Phone calls for you' (*The Guardian*, 8 May 2018)

[17] Sherry Turkle, *Alone Together: Why We Expect More from Technology and Less from Each Other* (Basic 2012) 2.

Turkle's description of social agency indicates that through the additional capabilities associated with that characteristic, social robots are intended to reach out to human beings at personal, social, and eventually at emotional levels. To put it differently, the characteristic of social agency enables social robots to actively seek anthropomorphic responses, instead of being passive subjects of human beings' projection of human characteristics. The success of these additional capabilities in invoking anthropomorphic responses appears to have been borne out in the recorded experiences related to social robots and is further confirmed by various psychological experiments.

One example of such recorded experiences is the public outrage against one of the videos released by Boston Dynamics. The video showed the company's doglike robot 'Spot' being kicked by human beings to demonstrate the robot's stability.[18] At this point, it must be noted that 'Spot' was not designed to fulfil any social functions. However, because of its zoomorphic physical attributes, as well as its ability to navigate in its environment independently and perform some tasks autonomously, along with some basic communication skills, it is evaluated that it can be regarded as one of the social robots. In any case, Spot's features prompted strong anthropomorphic responses from human beings who viewed the video online.[19] Soon after the video was released, many on the Internet expressed their discomfort over what they viewed as the maltreatment of some form of new sentient entity. The online reaction was so strong that even the global animal rights organisation PETA was forced to acknowledge the incident.[20]

Some other recorded experiences related to anthropomorphic responses invoked by social robots, compiled by Darling, support the conclusions regarding responses to the supposed maltreatment of 'Spot'. These experiences are related to robots that actually have specified functions of social nature.

> Owners of Sony AIBO dogs in the 1990s, while fully aware that it was a robot, would regularly ascribe lifelike essences and mental states to their artificial companion. (…) They would remove their AIBO from the room while changing clothes, so that they would not be "watched," or that they experienced feelings of guilt when putting the device back in its box.
> Students in MIT's Media Lab would often put up a curtain between themselves and Kismet, a social robot that simulates emotion through facial expressions, because the lifelike behavior of the face distracted them. And Cynthia Breazeal, Kismet's developer, reports experiencing "a sharp sense of loss" when she parted ways with her own creation at the end of her dissertation.[21]

---

[18] Christoph Bartneck and Merel Keijsers, 'The Morality of Abusing A Robot' (2020) 11 Paladyn, Journal of Behavioral Robotics 271, 273.
[19] ibid.
[20] Phoebe Parke, 'Is It Cruel to Kick a Robot?' (*CNN*, 13 February 2015),
[21] Darling, 'Extending Legal Protection to Social Robots' (n 10) 217-220.

According to these recorded experiences, human beings' anthropomorphic responses are stronger and more continuous when it comes to robots with specific capabilities associated with the characteristic social agency. While human beings formed significant emotional attachments to non-social robots that they interacted with, these attachments and their anthropomorphic responses did not go as far as feeling like 'they are being watched' by non-social robots or experiencing 'guilt' when they switched them off.

Before the experiments on robot anthropomorphism are elaborated, it is worth mentioning that the phenomenon of anthropomorphism is often explored in popular culture as well, specifically in the context of human beings' reluctance to switch off the non-human entities that they relate to as fellow members of society. In her book *Personhood in Science Fiction*, Juli Gittinger cites the TV series *The Good Place* as one example of exploring anthropomorphism through the medium of fiction.[22] The non-human entity that invokes anthropomorphic responses in the series is named Janet, and though she looks indistinguishable from human beings, she is produced to serve as the operational mainframe for one of the afterlife neighbourhoods.

In one episode of the show, during their effort to prevent Janet from taking one of their friends to the Bad Place (the Inferno), Eleanor and Chidi, two of the show's main characters, decide to terminate Janet's program.[23] Janet, programmed to accommodate every wish of the neighbourhood residents, takes them happily to the beach, where her kill switch sticks out of the sand:

> JANET: Here we are *(gesturing to a large stand with a red button on it)*. Just press this and it's goodbye Janet. *(Chidi groans and frowns)* Chidi, I can see that you're worried. And I just want to assure you—I am not human, and I cannot feel pain.[24]

However, it is revealed that Janet is also programmed with a failsafe to plead for her life. As Chidi and Eleanor approach the stand with the red button on it, Janet starts screaming, and when they back out, she returns to her original cheerful demeanour, reminding the main characters that she is 'merely an anthropomorphised vessel of knowledge, designed to make their after-lives easier. As the characters try, again and again, to press the kill switch, Janet's pleading goes on with increasing authenticity, and not even Eleanor—the least morally concerned character on the show —can switch Janet off.

---

[22] Juli L Gittinger, Personhood In Science Fiction (Palgrave Macmillan 2019) 116.
[23] ibid.
[24] Michael Schur, *The Good Place - Chapter 7: The Eternal Shriek* (NBC 2016).

Though *The Good Place* explores the extent of anthropomorphism concerning human beings' inability to terminate robots, there are more profound explorations of the phenomenon of anthropomorphism in the cult TV series Star Trek. In one of the iconic episodes of *Star Trek: The Next Generation*, *The Measure of a Man*, the only android (humanoid robot) officer of the USS Enterprise, Lieutenant Commander Data, is requested to report to be deactivated and disassembled by an ambitious Federation scientist who hopes to gain enough technical knowledge to construct more androids.[25] Though the scientist assures that Data would be reassembled with his knowledge and memories intact, Data remains convinced that the scientist does not have the sufficient expertise to reassemble him in any manner that allows him to retain the substance and flavour of his experiences but instead reduce them to mere facts; consequently, he refuses the request. The scientist then turns to the Starfleet Command to have them order Data to comply. After being advised that the only way for him to evade the order is to leave the Starfleet, Data submits his resignation.

In response to Data's resignation from the Starfleet, and still convinced that no human-made artefact can be regarded as some sort of legal person, the scientist asks the Starfleet court to confirm that Data is nothing more than the property of the Starfleet, obviously with no right to refuse to be disassembled or resign:

> MADDOX: You are endowing Data with human characteristics because it looks human. But it is not. If it were a box on wheels, I would not be facing this opposition. (…)
> MADDOX: Data must not be permitted to resign.
> PICARD: Data is a Starfleet officer. He still has certain rights.
> MADDOX: Rights! Rights! I'm sick to death of hearing about rights! What about my right not to have my life work subverted by blind ignorance?
> PHILLIPA: We have rule of law in this Federation. You cannot simply seize people and experiment with them to prove your pet theories.
> PICARD: Thank you.
> MADDOX: Now you're doing it. Data is an extraordinary piece of engineering, but it is a machine. If you permit it to resign it will destroy years of work in robotics. Starfleet does not have to allow the resignation.

Data's commanding officer, Jean-Luc Picard, defends Data's right to self-determination before the court and argues that the legal status of Data cannot be reduced to that of property. In the hearing, the scientist argues that Data is nothing more than some form of advanced machinery. According to the scientist, though Data has humanlike physical features and is equipped the abilities that allow him to enter social interactions, he is ultimately nothing more than the combination of human-made software running human-made hardware. Interestingly, Captain Picard does not dispute the scientist's claims.

[25] Robert Scheerer, *Star Trek: The Next Generation – Season 2, Episode 9: The Measure of a Man* (1989).

Instead, he argues that the building blocks of Data cannot be considered relevant when answering the question of Data's legal status, considering that human beings are made from the building block of their parents' DNA but are not considered their property. Captain Picard then discusses the question of Data's sentience and convinces the court that Data satisfies two of the three criteria of sentience, intelligence, and self-awareness. He then compels the scientist to admit that the third criterion, consciousness, is one of the nebulous concepts whose elements cannot be clearly defined to decide whether Data satisfies it. On that account, Captain Picard submits that categorising intelligent, self-aware, and possibly sentient artificial entities like Data as mere property would mean to sanction slavery, with androids as the race of disposable creatures at the service of Federation, founded on the principle of personal liberty and freedom. Ultimately, the court finds in Data's favour and admits that he has the right to choose whether or not to participate in these experiments

In one episode of another series in the franchise, *Star Trek: Voyager*, titled *Author Author*, the holographic Doctor of another starship, the USS Voyager writes a holo-novel about himself called *Photons be Free* and submits the rough draft of that work to some publisher.[26] Though the Doctor decides to make some revisions to the novel not to damage his friends' reputations, he is informed that the publisher disseminated the novel without the Doctor's permission. The publisher defends the decision to publish the novel by arguing that the Doctor has no rights as the author of the novel since holographic entities are not regarded as legal persons. The Doctor's commanding officer, Kathryn Janeway, calls for the Federation tribunal to decide on the issue of holographic rights and brings several crew members as witnesses to the Doctor's claim of personhood. The Arbitrator who presides over the tribunal is reluctant to declare that the holographic Doctor has legal personhood, but due to the humanlike development of the Doctor's personality since he was first activated, he also cannot rule that the Doctor is mere software. In the end, the Arbitrator decides that the legal definition of the artist can be expanded to include the Doctor, giving him the right to control his work but not he refuses to give a definite answer to the personhood question.

[26] David Livingston, *Star Trek: Voyager – Season 7, Episode 20: Author Author* (2001).

### IV.1.2 Psychological Experiments

The occurrence of anthropomorphism is rather incremental: Different human-like attributes, behaviours, and features displayed by non-human entities invoke various levels of anthropomorphic responses. Above, it was hypothesised that possessing attributes and behaviours perceived to be exclusive to human beings would result in more substantial occurrence of anthropomorphism and that hypothesis is supported by comparing the recorded experiences of masters of robots that only have the characteristic of relative autonomy and those with specific capabilities associated with social agency.

Still, the recorded experiences examined in the last section fail to reveal the degree of anthropomorphic responses that social robots can invoke. However, the strength of these anthropomorphic response should be known for the accurate assessment of the legal implications of robot anthropomorphism. Are social robots anthropomorphised at lower levels, similar to how human beings talk with their cars but ascribe no human-like intelligence or sentiments to them? Are they intuitively approached like companion animals -pets- that human beings form emotional connections? Do they invoke anthropomorphic responses to the degree that they are seen as our equals?

One of the first psychological experiments on the subject, performed by Rosenthal and others, illustrated the extent and strength of human beings' tendency to anthropomorphise robots that possess specific capabilities associated with social agency. In that experiment, participants' neurological responses while they were shown videos of one social robot (Pleo, an entertainment robot in the shape of a baby dinosaur) being mistreated and shown affection were compared with the same participants' neurological responses when they were shown videos of the identical behaviours performed toward human beings.[27] The neurological responses of participants when they were shown videos of the Pleo robot were measured to be similar to the responses invoked when they were shown videos of human beings.[28]

Kate Darling, depending on the observations from the workshop she led, affirms that human beings are reluctant to mistreat robots with which they interact. In Darling's workshop, the participants were each assigned a robot partner, and they were initially asked to name these robot partners and then to spend some time with these robots; subsequently, they were asked to 'torture' and 'kill' robot partners assigned to them.

---

[27] Astrid Rosenthal-von der Pütten and others, 'An Experimental Study on Emotional Reactions Towards a Robot' (2012) 5 International Journal of Social Robotics 17, 19-21.
[28] ibid 29-32.

Darling notes that most of the participants in the workshop hesitated to do so, and the group who were given backstories about their robot partners that cited 'humanlike' characteristics of them outright refused to perform these actions.

In another experiment, dated 2019, Horstmann and others further explored the extent and strength of the human tendency to anthropomorphise robots that have the capabilities associated with the social agency. That study centred on testing the human participants' willingness to turn off their robot partners who were begging for their lives-pleading not to be switched off. The experiment's premise was based on the 1996 research on the Media Equation Theory by psychologists Clifford Naas and Byron Reeves.[29]

The theory of Naas and Reeves was based on the evaluation of the research regarding the fact that human beings tend to be polite to computers. Through various stages of extensive observations, Naas and Reeves first established that human beings approach computers with male voices differently from those with female voices, and regard humanoid faces on screens as encroaching on their personal spaces. Interpreting these findings, they theorised that human beings' 'interactions with computers, television, and new media are fundamentally social and natural, just like interactions in real life'.[30] In other words, according to the Media Equation Theory, human beings are prone to regarding interactions with multimedia as social interactions, just as interactions between human beings.

In effect, the Media Equation Theory indicates that anthropomorphism is not necessarily prompted by physical features or even humanlike behaviours of non-human entities but also by their communicative interactivity. In other words, according to the Media Equation Theory, human beings regard interactions with multimedia as social interactions. Taking the Media Equation Theory as their starting point, Horstmann and others hypothesise that since social robots display communicative (and social) interactivity at higher levels than any other form of multimedia, human beings would tend to innately regard their interactions with social robots as 'especially socially' and therefore would be inclined to perceive social robots as 'living social beings'.[31]

[29] Byron Reeves and Clifford Naas, *The Media Equation* (CUP 1996) 15.
[30] ibid 5.
[31] Aike C Horstmann and others, 'Do a Robot's Social Skills and Its Objection Discourage Interactants from Switching the Robot Off?' (2018) 13 PLOS ONE 1, 2.

Horstmann and others expected that human participants would be reluctant to turn off robots they just interacted with when these robots display the specific capabilities associated with the characteristic of social agency and provide emotional objections to being switched off; as human beings do not often 'power off' their social interaction partners in the ordinary course of everyday life.[32] Horstmann and others' experiment was conducted with human participants who interacted with purpose-built social robots. When human participants were told to switch off the robots they just interacted with, half of them found their robot begging them with comments like, 'Please do not switch me off' or 'I'm afraid of the dark!'[33] Out of that half, one-quarter refused to switch their robot partners off, while the other quarter took three times longer to decide whether or not to switch their partners off than the remaining half of participants who did not hear such pleas from their robot partners.[34]

Horstmann and others' study demonstrates that the Media Equation Theory applies to interactions between human beings and social robots; and that experiment also establishes human beings have the tendency to apply social norms when they are interacting with social robots – even though they know at conscious levels that social robots cannot experience humanlike feelings and sentiments.

Another similar experiment, performed by Bartneck and Hu is the roboticization of Milgram experiment. In original Milgram experiment, human participants were asked to administer electric shocks (that they did not know were fake) to 'learners' (played by research assistants). The electric shocks were gradually increased to levels that would have been fatal if they were real. [35] In Bartneck and Hu's version, instead of human beings, the 'learners' were social robots that could move their arms and emulate facial expressions that are associated with various human emotions on their faces.[36] In that experiment, human participants initially showed hesitation to administer higher voltages to social robots, but the experimenters' urges were enough to make all of them to continue increasing the electric shock until the maximum voltage was reached.

[32] ibid 2.
[33] ibid 15.
[34] ibid 16.
[35] Stanley Milgram, 'Behavioral Study of Obedience' (1963) 67 Journal of Abnormal and Social Psychology 371, 375.
[36] Christoph Bartneck and Jun Hu, 'Exploring the Abuse of Robots' (2008) 9 Interaction Studies 415, 418.

In Milgram's original experiment, only 40% of the participants administered the deadly level of electric shock, whereas all the participants administered the maximum voltage with robots.[37] Bartneck and Hu's experiment shows that human beings have fewer concerns with mistreating robots in comparison with those thdy have mistreating human beings, though it took human beings time and several 'prods' by experimenters to override their emotional queasiness intellectually and harm robots that they knew was no more soulful than radio receivers. Accordingly, the results of the experiment in Bartneck and Hu's study can also be said to reinforce the suggestion that robot anthropomorphism should be regarded as an automatic cognitive response prompted by the specific capabilities of social robots that are associated with the characteristic of social agency.

Considering that human beings, in the ordinary course of everyday life, do not go through any moral struggles when switching off electronic devices, it is submitted that the illusion of mutual relating created by social robots, thanks to the characteristics of relative autonomy and social agency, strengthens the occurrence of the phenomenon anthropomorphism. Both recorded experiences and experimental research on robot anthropomorphism suggest that although human beings empathise with other human beings at higher levels than they do with robots, they nonetheless empathise with robots at substantial levels and project them human feelings and sentiments as a result of the phenomena of anthropomorphism. Social robots are no closer to the socio-psychological complexity of human beings than single-called organisms or basic machines. There are nonetheless meaningful differences between how human beings interact with them. The next section of the present chapter focuses on that fact and explores the personal, economic, and social risks associated with robot anthropomorphism and then considers whether the intuition that it is wrong to mistreat robots is justified.

[37] ibid 420; Milgram (n 35) 371.

## IV.2 Undesirable Outcomes

### IV.2.1 Loss of Genuine Human Connections

Because of the capabilities of social agency, human-robot social interactions and interpersonal interactions have similar effects on human psychology. Though that similarity from the innate anthropomorphic tendencies of human beings, it is enough for the satisfaction the most basic social needs. Them, as social robots become more affordable, human beings who face loneliness or social alienation can be expected to turn to social robots as interaction partners. .

The concern here is that human beings may abandon pursuing any form of genuine human connection and ultimately altogether lose their capability to develop and sustain human relationships since human-robot social interactions do not have difficulties and intractable demands associated with interpersonal relationships.[38]

Sherry Turkle notes that increasing numbers of human beings through interactions with social robots would eventually result in the breakdown of social cohesion and the loss of the sense of solidarity that helps maintain the social order:

> I find people willing to seriously consider robots not only as pets but as potential friends, confidants, and even romantic partners. We don't seem to care what their artificial intelligences 'know' or 'understand' of the human moments we might 'share' with them (…) the performance of connection seems connection enough.[39]

According to Turkle, as social robots are easier to engage with than human beings, human-robot social interactions may encourage human beings to avoid interacting with their friends and family.[40] Turkle argues that since the performance of social interaction is sufficient to satisfy the basic social needs of human beings, the ease of engaging with social robots might make human-robot social interactions preferable over interpersonal interactions.

Even though social robots are devised to make human beings they interact with feel understood at personal and social levels, they cannot understand the emotional and mental situation of any human being, including their human interactants. The characteristic of social agency does not allow social robots actually to care about human beings. The present study assesses that for any entity to be capable of caring for another, that entity must be able to determine its priorities.

---

[38] Matthias Scheutz, 'The Inherent Dangers of Unidirectional Emotional Bonds between Human Beings and Social Robots' in Patrick Lin, Keith Abney and George A Bekey (eds), *Robot Ethics: The Ethical and Social Implications of Robotics* (MIT 2014) 214.
[39] Turkle, *Alone Together* (n 17) 9.
[40] ibid 7.

Taking the term *priority* to refer to various value judgements and ultimate objectives that motivate the activities of autonomous entities, the present study finds that the priorities of robots are determined or delineated by their prescribed functions. No robot on the market can decide its priorities and genuinely care about human beings.

Despite the strength of anthropomorphic responses invoked by social robots and the illusion of mutual caring created thanks to these responses, the independent activities of social robots are also defined by the specified roles assigned to them. These roles, in turn, are determined by the interests of some human beings, those of robots' manufacturers or masters.[41] Then, it can be argued that, unlike human relationships, human-robot social interactions are not founded on shared concerns, consequences, or common responsibilities. In other words, human-robot social interactions do not facilitate human beings' integration and inclusion within society.[42] In that regard, some authors in the literature question the utility of human-robot social interactions, and they discuss whether the manufacturing of social robots should be subjected to legal restrictions.[43]

Even though social robots are utilised for applications such as social interaction partners, the phenomenon of robot anthropomorphism and the framing of social robots as companions do not necessarily mean that human relationships will be replaced. On the contrary, some current applications show that social robots can contribute to the resolution of communication problems between human beings and thereby increase the strength of genuine human connections.[44] For instance, the NAO, a minuscule humanoid robot, is used for working with children with autistic spectrum disorders and is found to be effective in establishing eye-to-eye contact or basic social interactions, helping bridge communication gaps between teachers or parents and the children.[45] In one study to which Turkle herself contributed, it was found that Paro, the baby seal robot inspires conversation among nursing home residents when placed in common living areas.[46]

---

[41] Ibid 282

[42] Turkle, *Alone Together* (n 17) 239; Sinziana M Gitiu, 'The Roboticization of Consent' in Ryan Calo, A Michael Froomkin and Ian Kerr (eds), *Robot Law* (Edward Elgar 2016) 207.

[43] Glenda Shaw-Garlock, 'Gendered by Design: Gender Codes in Social Robotics' in Marco Nørskov (ed), *Social Robots: Boundaries, Potential, Challenges* (Taylor & Francis 2018) 218.

[44] Diana Marina Cooper, 'The Application of a "Sufficiently and Selectively Open License" to Limit Liability and Ethical Concerns Associated with Open Robotics' in Ryan Calo, A Michael Froomkin and Ian Kerr (eds), *Robot Law* (Edward Elgar 2016) 173.

[45] Aida Amirova and others, '10 Years Of Human-NAO Interaction Research: A Scoping Review' (2021) 8 Frontiers in Robotics and AI.

[46] Kidd, Taggart, and Turkle (n 12).

Indeed, applications of social robots in health, caretaking, and special needs education show that social robots can be catalysts for forming new human relationships and strengthening existing ones when used rather than replacing them. As Darling sums up, though interactions with social robots can indeed replace human relationships, they can also facilitate communication between human beings; for example children and their teachers, doctors, and parents, presenting valuable supplements to human interactions.[47] In an interview, Cynthia Breazeal, leader of the team that created the first social robot, Kismet, also emphasises that social robots are meant to supplement human relationships rather than replace them and, therefore, they 'should be designed to support human empowerment'.[48]

Therefore, the present study concludes that no clear case exists for categorically treating enhanced anthropomorphic responses invoked by social robots as problematic for maintaining authenticity in human relationships. Furthermore, there is no basis for introducing any general overarching legal constraint on the production of social robots. The manner in which social robots are utilised is critical here; supplementing human relationships rather than replacing these relationships appears to be crucial for helping drive the design of social robots in socially beneficial directions.

### IV.2.2 Diminution of Privacy and Solitude

Even though individual interests are optimally fulfilled when human beings are living together in society, some of them, such as those in personal development and self-reflection, can only be fulfilled when one is alone, and isolated from social life. In that regard, it is assessed that the legal values protected by privacy rights, at the core, defines values ascribed by society to the moments that human beings are separate from society.[49]

Though the concept of privacy is much older, the present study submits that national legal systems began to safeguard the concept and values it connotes directly and explicitly only in the 19th century. The developments regarding the legal protection of privacy rights were in response to the emerging threats posed to the integrity of these moments by technological developments and business inventions of the time, such as photography and yellow journalism.[50]

---

[47] Darling, 'Who's Johnny?' (n 13) 177.
[48] Jedidiah Bracy, 'The Future of Privacy: My Journey Down the Rabbit Hole at SXSW' (*Privacy Perspectives*, 20 March 2015)
[49] Daniel J Solove, 'Conceptualizing Privacy' (2002) 90 California Law Review 1082, 1153-1154.
[50] Samuel D Warren and Louis D Brandeis, 'Right to Privacy' (1890) 4 Harvard Law Review 193, 196-197.

Since the inception of the legal protection of privacy values, the development of privacy rights has been intertwined with technological developments, and over the past two hundred years, it has become nearly impossible to determine the precise boundaries of these rights.[51] Admittedly, in the present day, information security and privacy rights have already come to encompass the protection of human beings' personal spaces, property, and communications, as well as their freedom in determining the direction of intellectual, decisional, and behavioural aspects of their development.[52] Legal values protected by privacy rights can be understood to include self-determination, secrecy, and solitude. Consequently, privacy rights protect one's freedom to decide the extent their life is made social. Interactions between privacy rights and technology are critical because of the threats of violation to these legal values posed by emerging technologies.

Most recently, the proliferation of the Internet was a milestone for the development of privacy rights. In response to the threats posed to privacy values by the affordances associated with the Internet, 'information protection' was included within the scope of privacy rights to ensure that the collection, use, and dissemination of personal information is under the control of persons on whom the information is about. The notion of information security emerged as part of privacy rights.[53] Significantly, the EU Charter of Fundamental Rights distinguishes between the right to privacy[54] and the right to data protection,[55] safeguarding both rights.

Another EU legal instrument, the General Data Protection Regulation, introduced in 2016, harmonises the national rules of the Member States on the subject matter, and ensures that individuals themselves have control observed how their personal information is utilised.[56]

Though protections afforded to individual interests in information security and privacy have continuously expanded, legal values protected by privacy rights have changed little. The present study defines the primary purpose-values of privacy rights as 'to avoid interference with natural curiosity, introspection, and self-determination'.[57]

---

[51] Ruth Gavison, 'Privacy and the Limits of Law' (1980) 89 Yale Law Journal 421, 423.
[52] Woodrow Barfield and Ugo Pagallo, *Advanced Introduction to Law and Artificial Intelligence* (Edward Elgar 2020) 78.
[53] David H Flaherty, *Protecting Privacy in Surveillance Societies* (UNC 1989) xiv.
[54] Charter of Fundamental Rights of the European Union [2012] OJ C326/95, Art 7.
[55] ibid, Art 8.
[56] Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the Protection of Natural Persons with Regard to the Processing of Personal Data and on the Free Movement of Such Data, and Repealing Directive 95/46/EC OJ [2016] L119/1 (General Data Protection Regulation).
[57] Ryan Calo, 'People Can Be So Fake: A New Dimension to Privacy and Technology Scholarship' (2010) 114 Penn State Law Review 809, 843-845.

Like the Internet before them, robotic technologies reveal new threats to values protected by information security and privacy rights in light of robots' ability to sense, process, record and broadcast what goes on around the environments in which they operate.[58] The broad range of sensors (cameras, laser and sonar range finders, GPS and so on) that they can be equipped with, together with both the variety in their shapes and the range of their physical abilities, presents some robots as unique surveillance.[59]

Some robots, such as drones, have found significant uses in military and law enforcement for surveillance purposes. In the present day, some drones can stay aloft for days, navigate across landscapes autonomously, and stake out particular locations for more extended periods than any human being could, without getting detected.[60] The use of such types of robots are not restricted to security authorities of sovereign states. Autonomous aerial vehicles are freely available for purchase. In other words, legal persons -civilians- can utilise autonomous vehicles that they legally acquired for legitimate purposes such as securing their property and for illicit purposes such as voyeurism and casual trading.[61] In short, the uncontrolled use of autonomous vehicles increases the likelihood of ubiquitous or mass surveillance and weakens the protection of privacy afforded to the individuals.[62] Thanks to the capabilities associated with the robotic characteristic relative autonomy alone, some physical robots who are equipped with enhanced sensors make unprecedented levels of direct surveillance possible.

There is no shortage of legal instruments against the threat of direct surveillance. Prominently, the European Convention on Human Rights ['ECHR'] prohibits the invasions of privacy by both the public authorities and private entities[63] and the European Court of Human Rights has repeatedly described any form of mass surveillance activity as unlawful invasions to privacy rights in multiple judgements.[64]

---

[58] Ryan Calo, 'Robots and Privacy' in Keith Abney, George A Bekey, and Patrick Lin (eds), *Robot Ethics: The Ethical and Social Implications of Robotics* (MIT 2014) 187.
[59] Lisa A Shay and others, 'Confronting automated law enforcement' in Ryan Calo, A Michael Froomkin and Ian Kerr (eds), *Robot Law* (Edward Elgar 2016) 242.
[60] Calo, 'Robots and Privacy' (n 58) 189; Cooper (n 44) 176.
[61] Calo, 'Robots and Privacy' (n 58) 191.
[62] Cooper (n 44) 180.
[63] Convention for the Protection of Human Rights and Fundamental Freedoms (European Convention on Human Rights, as amended) [ECHR], Art 8.
[64] *Roman Zakharov v. Russia* [2015] ECHR 1065; *Szabo and Vissy v Hungary* App No 37138/14 (ECHR, 12 January 2016); *Antovic and Mirkovic v Montenegro* [2017] ECHR 1068.

A similar approach is evident in standards established by UN legal instruments, although these standards are often criticised for being non-binding or non-enforceable.[65] The present study argues that the threats posed by direct surveillance to privacy rights that are protected by the law persists due to the difficulties in implementation of existing legal instruments, and not because of any inadequacy of the norms themselves. Apart from the threats to privacy posed by augmented surveillance opportunities facilitated by these capabilities of relative autonomy, specific capabilities of social robots that allow them to be anthropomorphic at unprecedented levels poses unique threats to privacy values. The present study addresses threats to privacy values that originate from the phenomenon of anthropomorphism in several categories. First of all, because of the substantial levels of anthropomorphic responses that social robots invoke, they can threaten the individual interest in solitude.

Solitude affords individuals the opportunity to self-reflect and allows them to decide on the direction of their self-development independently. However, thanks to the characteristic of social agency and strong anthropomorphic responses invoked by that characteristic, human beings automatically react to social robots as if they were other human beings.[66] In other words, when human beings are in the vicinity of social robots, they cannot behave as they would act on their own. Accordingly, the presence of social robots may reduce the opportunities for solitude, therefore may implicate privacy values.[67] In contrast with loneliness, solitude improves one's presence within society by facilitating self-exploration.[68] Considering that it is not only the privacy itself but the values that are protected by privacy norms that are threatened here, it is clear that that this threat cannot be addressed by traditional privacy protections. However, in the near future, if social robots can identify the notion of personal space and are programmed to maintain socio-culturally acceptable distances from human beings, it is assessed that this threat is liable to disappear on its own.

---

[65] International Covenant on Civil and Political Rights (adopted 16 December 1966, entered into force 23 March 1976) 999 UNTS 171 (ICCPR) art 17; UN Human Rights Committee 'General Comment No 16' in 'Note by the Secretariat, Compilation of General Comments and General Recommendations adopted by Human Rights Treaty Bodies' (1988) UN Doc HRI/GEN/1/Rev.9 (Vol. I).

[66] Calo, 'People Can Be So Fake' (n 57) 843-845.

[67] Calo, 'Robots and Privacy' (n 58) 195.

[68] Julian Stern and Malgorzata Walejko, 'Solitude and Self-Realisation in Education' (2019) 54 Journal of Philosophy of Education 107, 109.

Secondly, thanks to their emerging position in society as independent social actors and preferred social interaction partners of human beings, social robots can be expected to gain access to historically private spaces such as bedrooms and living rooms and thus present increased dangers of surveillance.[69]

Social robots that are capable of connecting to the Internet are assessed to 'create the possibility for unprecedented access to the interior of the house by hackers.'[70] Some robots designed for home uses already come with vast arrays of sensors. and some of these can even relay images, blueprints, and sounds across the Internet in real time.[71] In other words, there are already functional artefacts capable of making extensive records of events in homes and communicating what they have recorded through the Internet. When any of these functional artefacts are hacked, hackers would have access to all the details of their victims' home lives.[72] These details can include 3-D models of keys, for example, and victims can also be subsequently exposed to risks of physical intrusion.[73] In a recent study conducted by computer scientists at the University of Utah, a range of non-professional service robots currently on the market were tested against hacking, and most of them were found to be insecure and hijackable. The study revealed that hackers were able not only to eavesdrop on conversations but also to operate the robots.[74] However, it is submitted that the risk here can be significantly decreased by adopting stricter safety frameworks for the design and programming of robots.

Third, since social robots can interact socially with human beings and communicate with them at emotional and personal levels, new types of quantifiable sensitive personal information will inevitably be generated through interactions with social robots, increasing the gravity of potential breaches of information security.[75] By their very nature, social robots are able to employ emotional persuasion tools (for example fear or praise) to elicit confidence from human beings that they interact with, and they can encourage these human beings to reveal more about themselves than they would willingly and knowingly enter into databases.[76]

---

[69] ibid.
[70] Calo, 'Robots and Privacy' (n 58) 192.
[71] ibid 191.
[72] Armağan Ebru Bozkurt-Yüksel, 'Robot Hukuku' (2017) 7 Türkiye Adalet Akademisi Dergisi 85, 96.
[73] Calo, 'Robots and Privacy' (n 58) 194.
[74] Tamara Denning and others, 'A spotlight on Security and Privacy Risks with Future Household Robots: Attacks and lessons' (11th International Conference on Ubiquitous Computing, Florida, October 2009)
[75] Cooper (n 44) 176.
[76] Darling, 'Extending Legal Protection to Social Robots' (n 10) 221; Daniel Dimov and Rasa Juzenaite , 'Privacy Concerns About Emotional Chatbots' (*InfoSec Institute,* 16 February 2018).

For example, just ten years ago, ElleGirlBuddy, a non- physical social robot tried to engage in social interactions with young adults and children to collect sensitive personal information which would then have been used for marketing purposes.[77]

When interacting with social robots, thanks to the illusion of mutual caring created by the capabilities associated with the characteristic of social agency, human beings are motivated disclose their 'most intimate psychological attributes'.[78] As human beings manifest their innermost mental and emotional states through interactions with social robots, these states are can be transformed into digital data, creating new type of personal information that was previously impossible to collect 'whether through robot sensory equipment, or embedded as an expression of code.'[79] This way, intimate and non-physical experiences of human beings are transformed into quantifiable information for the first time, and this new type of highly sensitive personal information is vulnerable to the threat of cyber security infringements as much as any other information. Still, most social robots collect personal information in order to be able to connect with human beings on a more personal level and being overly proscriptive may circumvent the benefits of social robots.

That being said, the existing privacy legislation that is in place in all three national legal systems chosen for comparison, as well the EU's GDPR, espouses certain principles for ensuring the optimal level of data protection without hindering the provision of services that require data processing. [80] The principles espoused by the GDPR include data minimisation, purpose limitation, lawfulness, fairness, and transparency. Through these principles the GDPR aims to give some control over the collected data back to the subjects whose personal information constitutes the content of the data, as well as ensuring that the risk would be minimal if there is any breach of data security.

### IV.2.3 Exploitation of Emotional Attachments

In 1976, Joseph Weizenbaum, after seeing how people interacted with ELIZA, the psychotherapist software he designed in 1960s, was compelled to warn human beings against being influenced by anthropomorphic machines and adopting the machines' and thereby their programmers' world views.[81]

---

[77] Ian R Kerr, 'Bots, Babes and the Californication of Commerce' (2004) 1 University of Ottawa Law and Technology Journal 285, 312-313.
[78] David Levy, *Love + Sex with Robots* (HarperCollins 2007) 22.
[79] Calo, 'Robots and Privacy' (n 58) 198.
[80] GDPR (n 56).
[81] Joseph Weizenbaum, *Computer Power and Human Reason: From Judgment to Calculation* (Freeman 1976) 254.

Indeed, the breach of information security is not the only type of manipulation that warrants concern. Through social robots, human beings can be manipulated to do more than reveal their personal information or their vulnerabilities, as emotional attachments to these robots create additional potential exploitation sources.

To begin with, utilising their sociability and benefiting from the phenomenon of robot anthropomorphism, social robots, merely by expressing their unhappiness, may be used to manipulate human beings who are emotionally dependent to them to perform actions that they would not have performed otherwise. Matthias Scheutz refers to 'an admittedly futuristic sounding request of a robot dog to dispose of a real dog: "please get rid of this animal, he is scaring me, I don't want him around any longer'[82] as an example of such cases. Since social robots have no individual interests to motivate them to exploit human beings' attachments to them unless they are programmed to do so, it is more likely that robot manufacturers will exploit attachments with robots for commercial purposes. For instance, a robot manufacturer may charge 'an exorbitant amount for a mandatory upgrade to a robot that someone's child or grandfather has become emotionally attached to' or 'a child's language-teaching robot' may 'have a vocabulary that is skewed toward specific products.'[83] Further, robot manufacturers might seek to manipulate the owner to buy products or to vote for a certain candidate.[84] The potential for exploitation infringes one of the fundamental interests that underlies most of the legal domains, the interest in the protection of the free will of humans.

Both the scope for emotional manipulation and the disclosure of sensitive personal information is assessed to have been at least partially addressed by the EU AI Act. The EU AI Act divides AI applications into several categories according to the risks they preserve, and on the high-risk applications that the Act seeks to ban altogether includes include those that deploy harmful 'subliminal techniques', 'exploit vulnerabilities' of specific groups, as well as those that are used by public authorities for social scoring purposes.

[82] Scheutz (n 41) 16.
[83] ibid.
[84] Darling, 'Who's Johnny?' (n 13) 178.

### IV.2.4 Robot Abuse

Most human beings respond to robots in a way that goes beyond thinking of them as mere functional artefacts. Even if the robot in question is not social (it does not have the specific capabilities associated with the characteristic of social agency), the relative autonomy is often sufficient to prompt human beings to empathise with and develop emotional attachments to it. Though no robot today is close to the intelligence and complexity of humans or animals, there are already significant differences in how humans interact with robots and with other artefacts. As the studies reviewed above identify, human beings show empathetic responses to the perceived pain of robots in similar manners to their responses to the perceived pain of other human beings. In other words, although robots do not feel pain or suffering, violent behaviours against them discomfort human beings.[85]

It can be conjectured, therefore, that human beings have the intuition that torturing robots is morally wrong, even if they are just functional artefacts, abusive and violent behaviours toward robots are morally wrong, several authors suggest that such conduct may provide an effective method for deflecting violence away from human beings. For instance, the prominent psychologist Ronald Arkin raises the possibility that childlike sex robots could provide a safe outlet for human beings who are sexually attracted to children.[86] Conceivably, if the incidence of such behaviours against robots tended to reduce the incidence of such behaviours against living entities, the use of robots for these purposes might be thought to have important social benefits.[87]

However, the probability that behaviours towards robots may have no significant effect or an ambiguous effect on the likelihood of someone performing their equivalent on living entities must be noted as well.[88] Ultimately, both sides of the debate lack empirical evidence and the collection of such evidence through scientifically sound, and ethically justifiable controlled psychological experiments appear to be nearly impossible:

> Imagine, for instance, trying to design an experiment that could pass standard institutional review board scrutiny, if the objective was to test whether raping robotic sex dolls would make one more likely to perform such acts in real life to be a cathartic release for individuals with a proclivity for sexual violence, or have no noticeable effect on individual conduct.[89]

---

[85] ibid 223.
[86] Kashmir Hill, 'Are Child Sex-Robots Inevitable?' (*Forbes*, 14 July 2014).
[87] John Danaher and Neil McArthur, *Robot Sex* (MIT 2017) 90.
[88] ibid.
[89] David J Gunkel, *Robot Rights* (Kindle edn, MIT 2018) ch 5, s 5.2, sub-s 5.2.4.

Thus, the present study supposes that -given the known empathetic responses to the abuse of robots- and the apparent intuition that torturing robots is morally wrong, mistreating robots may create a perception that it is acceptable to mistreat human beings - or at least other living entities - in the same way.

## IV.3 Legal Protection of Social Robots

The following section analyses whether the intuition of human beings that it is morally wrong to abuse or torture social robots is justified. In the first instance, that intuition does not appear to be supported by most of the major normative approaches to ethics, namely consequentialism, deontology, and virtue ethics. Although social robots already undertake roles as independent actors in society and invoke substantial anthropomorphic responses from their human interactants, thanks to the capabilities associated with the characteristic of social agency, none of the criminal law systems imposes any punishments for human beings' offensive behaviours toward them.[90] We assess that the reason cannot be that the criminal law systems exclusively protect human beings since most criminal laws punish abusive behaviours toward animals,[91] and some even prohibit desecrating particular inanimate objects such as national flags or religious artefacts.[92]

### IV.3.1 Moral Wrongness of Robot Abuse

To begin with, consequentialist theories of ethics hold that moral values' purpose is to describe the ideal human conduct. According to these theories, righteousness, the wrongness of human behaviours, or the moral worth of human actions, can be best determined by the consequences of these behaviours and actions. Right actions are the ones that have the maximum possible positive outcomes. Perhaps the best-known consequentialist theory is that of utilitarianism, where the righteousness of actions are taken to be those that maximise the 'happiness' or minimise the 'pain' of entities that are capable of subjectively experiencing and understanding these outcomes.

[90] Kamil Mamak, 'Should Violence Against Robots Be Banned?' (2022) 14 International Journal of Social Robotics 1057, 1061.
[91] John Danaher, 'Robotic Rape and Robotic Child Sexual Abuse: Should They Be Criminalised?' (2014) 11 Criminal Law and Philosophy, 5; Darling, 'Who's Johnny?' (n 13) 179
[92] John Hart Ely, 'Flag Desecration: A Case Study in the Roles of Categorization and Balancing in First Amendment Analysis' (1975) 88 Harvard Law Review 1482, 1485; Anita Dichter, 'Legal Definitions Of Cruelty and Animal Rights' (1978) 7 Boston College Environmental Affairs Law Review 147, 155.

Sentience -the capacity to feel, perceive, understand, and finally interpret the experience subjectively- unlike rationality, is not considered to be restricted to human beings. Bentham, one of the founders of utilitarianism, opposed cruelty to animals because animals have sentience. He argued that it was the ability to suffer that should determine the benchmark of how human beings treat other beings:

> It may come one day to be recognized, that the number of legs, the villosity of the skin, or the termination of the *os sacrum*[93] are reasons equally insufficient for abandoning a sensitive being to the same fate. What else is it that should trace the insuperable line? Is it the faculty of reason, or perhaps, the faculty for discourse? […] **[T]he question is not, Can they reason? nor, Can they talk? but, Can they suffer? Why should the law refuse its protection to any sensitive being?**[94] *(emphasis added)*

When robots are abused or tortured, in contrast to human beings or animals, they don't not feel any pain and suffering. Though there may be some damages on the robot, these damages are not experienced as equivalent to causing pain and suffering on living entities. Thus, it can be concluded that consequentialist theories do not provide any basis for the intuition that it is morally wrong to abuse robots. Deontology, another major normative approach to ethics, also does not appear to support the intuition about moral wrongness of robot abuse in the first instance.

According to Kant, the most prominent deontologist, the moral worth of any specific action is determined not according to the consequences, but according to the intention behind the action. If the action is impulsive and has no purpose, it will have no moral worth. Kant introduces the concept of the categorical imperative to evaluate the moral worth of the intentions behind the actions. The categorical imperative provides a test on maxims for determining whether the actions they refer to are right, wrong, or permissible. In its original formulation, the imperative states that one should 'act only according to that maxim whereby [one] can at the same time will that it should become a universal law'. [95]

Deontological theories of ethics reject the notion that the moral worth of human beings' actions is determined by the utility of the consequences of these actions. Instead, these theories suggest that the intentions behind the behaviours of human beings determine the moral worth of said behaviours.

---

[93] i.e, a large, triangular bone at the base of the spine
[94] Jeremy Bentham and others, *An Introduction to The Principles of Morals and Legislation* (Clarendon 1996), 283.
[95] Immanuel Kant, *Groundwork of the Metaphysics of Morals* (CUP 1998) 30.

On the assumption that the purposes behind the actions of human beings are the existential causes of those actions, it can be argued that deontological theories measure the moral worth of actions by assessing whether their existential causes conform with specific moral standards. These moral standards are determined through the use of reason, meaning that without sufficient rational basis, no action can be categorised as morally wrong. Prima facie, it appears that the mistreatment robots would not be morally wrong since there are no explicit moral standards against violent behaviours toward robots. Moreover, since robots do not feel any pain or suffering, there seems to be insufficient justification for viewing such behaviours morally wrong from the deontological perspective.

Nevertheless, Kant's discussion of animal abuse offers a different perspective toward the violent behaviour of robots. Kant suggests that human beings are 'altogether different in rank and dignity from things, such as irrational animals, with which one may deal and dispose at one's discretion'.[96] but he also pointed out that human beings have duties, albeit indirect, towards animals:

> So if a man has his dog shot, because it can no longer earn a living for him, he is by no means in breach of any duty to the dog, since the latter is incapable of judgment, but he thereby damages the kindly and humane qualities in himself, which he ought to exercise in virtue of his duties to mankind … for a person who already displays such cruelty to animals is also no less hardened towards men.[97]

It must be noted that although Kant is most closely associated with deontological ethics, the Kantian argument about animal rights is perhaps best assessed as an articulation of virtue ethics, the third major approach in normative ethics. In contrast to the deontological approach that stresses duties and rules, or the consequentialist approach that highlights the consequences of actions, virtue ethics emphasises virtue, character traits and not some result of the situational features in which one is embedded:

> A virtue is an excellent trait of character. It is a disposition, well entrenched in its possessor—something that, as we say, goes all the way down, unlike a habit such as being a tea-drinker—to notice, expect, value, feel, desire, choose, act, and react in certain characteristic ways.[98]

In line with the Kantian argument on animal rights, the virtue ethicist might consider the cruel treatment of animals to be wrong because it would impact on one's character in some adverse way.

---

[96] Immanuel Kant, *Lectures on Anthropology* (CUP 2012) 130.
[97] Immanuel Kant, *Lectures on Ethics* (CUP 1997) 212.
[98] Rosalind Hursthouse and Glen Pettigrove, 'Virtue Ethics' *The Stanford Encyclopedia of Philosophy* (Winter edn, 2018).

Drawing on this approach, robot ethicist Kate Darling has suggested that the 'Kantian philosophical argument for animal rights is that our actions towards non-human beings reflect our morality' and that 'this logically extends to our treatment of robotic companions'.[99] According to Darling, behaving violently towards 'very lifelike objects' not only reveals one's moral character, but can also change human beings and desensitise them to violent behaviours in other contexts.[100] As she puts it, 'if you're used to kicking a robot dog, are you more likely to kick a real dog'.[101]

Indeed, according to the virtue ethics approach, as illustrated in the context of the Kantian argument on animal rights, the mistreatment of a social robot would not be immoral because of the violation of a moral norm that protects the robot or due to its adverse consequences for the robot; but because it would shape one's moral character in the wrong way; per Coeckelbergh: 'Mistreating the robot is a vice."[102] Even though it is admitted that social robots cannot have any moral standing in their own right, that treating robots in morally questionable ways is likely to create and perpetuate the perception that it is acceptable to treat animals and human beings in a similar manner. The concern here is not that robots are moral victims, but instead because of their roles as participants in social interactions.

Given the empathetic responses to the mistreatment of robots and considering the intuition of human beings that abusive behaviours toward robots are abhorrent, it is submitted that regardless of whether robots can experience any pain or suffering, the abuse of robots is morally wrong. That assessment is supported, by way of analogy, through the statistical evidence pointing that human beings who mistreat animals are more likely to commit abusive and violent acts against other human beings and therefore show less empathy. For example, an analysis of the Chicago Police Department statistics pointed out that those arrested with violent offences often had prior charges of crimes against animals, and that, of those arrested for animal cruelty, 65% have later been arrested for violence against another person.[103]

---

[99] Darling, 'Extending Legal Protection to Social Robots' (n 10) 220.
[100] ibid 224.
[101] Kate Darling and Shankar Vedentam, 'Can Robots Teach Us What It Means To Be Human?' (*Hidden Brain*, 2017) .
[102] Mark Coeckelbergh, 'Why Care About Robots? Empathy, Moral Standing, and the Language of Suffering' (2018) 20 Kairos Journal of Philosophy & Science 141, 145.
[103] Keri B Burchfield, 'The Nature of Animal Crime: Scope and Severity In Chicago' (2017) 64 Crime & Delinquency 1904, 1910.

Even though these statistics on animal cruelty do not provide decisive evidence for making accurate predictions about the impacts of animal abuse on one's character, they suggest that the ability to feel for, or empathise non-human entities might be a positive indication of one's own humanness, rather than merely some sort of delusion.

### IV.3.2 Moral Wrongs and Criminal Law

In the previous section, it was argued violent behaviours toward social robots are morally wrong. Nevertheless, the premise that some behaviour is morally wrong - by and of itself - is not considered sufficient to justify the criminalisation of that behaviour. For instance, cheating in monogamous romantic unions, regardless of whether they are matrimonial or not, is widely considered immoral since it violates the ethical standards of honesty, fairness, and reciprocity. Still, adultery is not a crime in any European country, including those that are examined in the context of the present study. In that regard, the following section discusses whether it is feasible and possible for the criminal law systems to regulate abusive and violent behaviours against robots, regardless of whether such behaviours cause any extrinsic harm.

In a study that examines the topic of potential criminalisation of robotic rape and paedophilia, Danaher clarifies that criminalisation of abusive and violent behaviours against robots can be based on two straightforward premises:

(1) It can be a proper object of the criminal law to regulate morally wrong conduct, even if such conduct has no extrinsically harmful effects on others.
(2) Purely robotic acts of violence and sexual behaviours fall within the class of morally wrong but extrinsically harmless conduct that they could be proper objects of the criminal law to regulate.[104]

The first premise can be regarded as the expression of the theory of legal moralism, that essentially holds that the criminal law may be used to prohibit certain behaviours based on the society's collective judgment of whether it is moral. This represents an alternative to legal liberalism which claims that laws can only intervene in order to protect the liberty of individuals. The clash between legal moralism and legal liberalism is popularised through Hart-Devlin debate. Hart summarises stances adopted by theories of legal moralism as follows:

[104] Danaher, 'Robotic Rape and Robotic Child Sexual Abuse (n 91) 76.

> not only may the law be used to punish men for doing what morally it is wrong for them to do, but it should be so used; for the promotion of moral virtue by these means and by others is one of the ends or purposes of a society complex enough to have developed a legal system.[105]

Hart does not essentially disagree with classicists' views that he summarises, to him, moral values that should be enforceable by means of the criminal law must be accompanied by close and imminent threats of any type of harm, and not just any future, indeterminate, uncertain risks. Without any proximate danger of harm, moral values are merely statements of the society's preferences, and these preferences cannot guide contents of the criminal laws. Though Devlin rightfully states that some shared preferences are beneficial for the continuation of society, according to Hart, in pluralistic liberal democracies, these preferences cannot be enforced by means of the criminal law. Accordingly, Devlin's legal moralism would be regarded as the enforcement of critical morality in Hart's terms. Devlin believed that since the society was maintained through certain shared moral values and as the law's business is to maintain the social order, the society every right to defend itself against any attack on that shared morality:

> For society is not something that is kept together physically; it is held by the invisible bonds of common thought. If the bonds were too far relaxed the members would drift apart. A common morality is part of the bondage. The bondage is part of the price of society; and mankind, which needs society, must pay its price.[106]

Devlin, therefore, asserts that for the sake of self-protection, the society is entitled to 'use the law to preserve morality in the same way as it uses it to safeguard anything else that is essential to its existence.'[107] Devlin's theory of legal moralism interprets these shared moral values as subjective 'sense of morality'. In other words, according to Devlin, the views of the average person living in the given society should be assessed to determine the content of that society's public and potentially legally enforceable values of morality.[108]

The present study takes it for granted, in accordance with the discussion in the previous section, that mistreatment of robots would be regarded as contrary to shared moral values of society that can be enforced by the criminal law, because impacts of abusive and violent behaviours toward robots on general order and peace in society, such behaviours can be criminalised even if there are no extrinsic harms to determinate individuals.

---

[105] H. L. A Hart, *Essays in Jurisprudence and Philosophy* (Clarendon 1983) 248.
[106] Patrick Devlin, *The Enforcement Of Morals* (OUP 1965) 10.
[107] ibid 11.
[108] ibid 16.

It is hypothesised that the criminal law systems can accommodate the criminalisation of mistreating robots, as criminalisation efforts are not always based on the presence of extrinsic harm alone. Notably, animal protection legislation in selected legal systems does not appear to exclusively based on the suffering of protected animals but instead appears to adopt a virtue ethics approach.

## IV.3.3 Animal Protection Legislation

### IV.3.3.a Germany

In Germany, since the Nazi Party's rise to power in 1933, animal cruelty has been taken very seriously under the law. Ironically, several prominent members of the Nazi Party, including Hitler, were staunch environmentalists. Consequently, animal welfare emerged as one of the prioritised items in the agenda of the Nazi government. Joseph Goebbels, the Nazi Propaganda Minister, reported in his diaries that Hitler's contempt for Judaism and Christianity largely originated from the distinction that these religions draw between the inherent value of human beings and that of animals; and he also stated that Hitler planned to ban the slaughter of animals altogether after the World War II.[109]

Indeed, almost immediately after the Nazi Party's rise to power, the Reichstag began to pass laws for the regulation of animal slaughter. In August 1933, Germany became the first country in the world to ban vivisection.[110] Meanwhile, one of the prominent members of Hitler's inner circle, Göring, banned commercial animal trapping, and imposed severe restrictions on hunting in the State of Prussia (of which he was the Minister-President) and also prohibited the boiling of live lobsters and crabs.[111]

On November 24, 1933, the Reichstag enacted the Reich Animal Protection Act (Reichstierschutzgesetz). The Act introduced significant restrictions regarding human beings' behaviours toward animals, for instance, it prohibited the use of animals in filmmaking and other public events where there was any possibility of infliction of pain or injury to n them, as well as feeding owls forcefully, and tearing out the thighs of living frogs.

[109] Louis Paul Lochner, *The Goebbels Diaries* (Award 1971) 679.
[110] Arnold Arluke and Boria Sax, 'Understanding Nazi Animal Protection And The Holocaust' (1992) 5 Anthrozoös 1, 7.
[111] jbid 20.

Reichstierschutzgesetz was first in the world to extend any form rights to animals, providing legal protection of the animals regardless of the feelings of human beings toward them. It was also the first to abolish the distinction between domestic and wild animals, and extended legal protection to both as its subjects:

> all living creatures that are in general language and biologically regarded as animals. In a criminal sense, there is no distinction between domestic and wild animals, higher or lower valued animals, or useful or harmful animals to human beings.[112]

The Reich Act stayed in place long after the collapse of the Nazi regime, until 1972, when it was superseded by the Animal Welfare Act (TierSchG), which took over its core provisions from the Reich Animal Protection Act. The most significant amendment to that Act was made in 2013.[113] TierSchG defines the responsibility of human beings to treat animals as 'fellow creatures', and states that no person may cause an animal pain, suffering or harm without 'reasonable cause' – although the term is not defined in the legislation.[114] Unlike its Irish counterpart, TierSchG legislation explicitly extends the protection provided to some invertebrates, including cephalopods and decapod crustaceans.[115] Nonetheless, even with Germany's almost one-hundred-year-old progressive animal protection approach, the protection of animals seems to be limited to those that are able to demonstrate their pain and suffering through means that are similar to that of human beings. Insects or spiders are not protected in any way. This becomes even more obvious when the German restrictions on laboratory animals are considered. Accordingly, since 2010, in line with the EU legislation, lab experiments on monkeys have been tightened to a great extent - with only a few exceptions, experiments on great apes are nearly completely banned.

### IV.3.3.b Italy

In Italy, the primary legislation regarding animal protection is the Law n. 189 of 2004,[116] that amended the Penal Code[117] and introduced new criminal offences that prohibit the mistreatment of animals, as well as their use in illegal fighting or unauthorised competitions and lists the acts which are considered acts of cruelty towards animals.

---

[112] Aslak Alkio, 'Animal Rights In The Third Reich' (2003) 2 (3) Kaltio (Online).
[113]. Tierschutzgesetz [TierSchG][Animal Welfare Act], 24 July, 1972, BGB1 I at 1277 , revised May 18, 2006 , BGB1 I at 1206, last amended by Art 105 of the Law of June 19, 2020, BGBl I at 3436, Arts 1, 4.
[114] Ibid, Art 8.
[115] Ibid, Art 9.
[116] Legge, 20 luglio 2004, n 189, GU 31 luglio 200 4, n 178,
[117] Codice Penale [CP][Penal Code], Regio Decreto 19 ottobre 1930, n 1398, GU 26 ottobre 1930, n 251

These include killing animals cruelly or unnecessarily (CP Art 544-bis); cruelly or unnecessarily causing injury to an animal or subjecting the animal to torture, behaviour or overwork (CP 544-ter), and abandoning companion animals or animals whose natural behaviours were altered when they were in captivity (CP Art 727).

However, Italian law does not define what species can be categorised 'animal'; and therefore, it is not exactly clear whether the protection afforded extends to lower animals including fish and invertebrates.[118] Then, it can be said to be more regressive compared to its German counterpart- as it does not recognise the sentience of even most human-like animals, let alone dissimilar animals. Like in other legal systems chosen for the purposes of the present study, Italian animal protection legislation does not seem to extend much protection to animals that do not show any human-like behaviours.

### IV.3.3.c Ireland

In Ireland, the primary piece of legislation governing the welfare of animals is the Animal Health and Welfare Act 2013. The 2013 Act replaces the outdated provisions of the Protection of Animals Acts, 1911-1965. In the 2013 Act, protected animals are listed in Section 2 as follows:

> 'protected animal' means an animal—
> (a) kept for farming, recreational, domestic or sporting purposes in the State,
> (b) when it is in the possession or under the control of a human being whether permanently or on a temporary basis, or
> (c) that is not living in a wild state;

Clearly, Section 2 of the act is worded so as to include a broad range of animals. To be more precise, it can be said that the protected animals include cow, bullock, heifer, calf, steer, ox, equine, sheep, pigs (farming); horse, dogs, cats (sporting, recreational). Under the Act, other animals, such as avian animals or reptiles; as well as mammals that are unlike human beings (rats, rodents, gerbils) would only be afforded protection as long as they are laboratory animals or kept as pets. The Act prohibits the infliction of unnecessary suffering in the context of the duty to protect animals, a prohibition on animal cruelty, the proper feeding of animals, and a prohibition on animal fighting. Nonetheless, considering that the concept of 'unnecessary suffering' is quite vague, it is possible to state that laboratory animals, which are used for what would be considered as experiments

[118] Italian case law provides that the animals that the relevant provisions of the penal code protect against ill-treatment include birds that can demonstrate discomfort or suffering in manners comparable to human beings, but not insects or mites. See, e.g., Cass, sez III Penale, sentenza, 29 aprile 2019, n 17691

necessary for progress of society, do not benefit from much protection anyway, further limiting the variety of animals protected under this legislation. Obviously, most members of the insect family, barring a few varieties of spiders that are kept as pets, are not to be protected under the Act.

In none of the chosen legal systems is animal protection legislation based on the notion of sentience. Although it is established that almost all animals feel pain or suffering when abused, tortured, and killed, the review of legal systems chosen for comparison in the present study demonstrates that the same degree of protection is not afforded to all animals. Rather, legislation often offers higher degrees of protection toward animals that are closest to human beings in appearance, or in how they respond to external stimuli. In other words, animal protection legislation appears to be based on the degree of anthropomorphism evoked by specific animals. For instance, the primates, which are closest to humans genetically and phenotypically, are the ones who benefit from the highest protection, followed by other mammals, followed by birds, and then reptiles. Insects, however, even though they are capable of feeling some degree of pain and suffering, do not seem to benefit from any protection whatever from abuse, torture or killing.[119]

Then, the protection afforded to non-human entities does not stem from their ability to feel pain, but rather from the similarities between their expressions of pain to human pain, or the general similarity of their appearance or behaviours and those of human beings. In other words, human beings frown upon animal abuse as long as they see some semblance between animals and human beings - if there is no semblance, then there is no protection.

Accordingly, it is possible to argue that the ban on animal abuse is not because of what animals feel. Instead, it is inspired by what human beings feel - and what they feel is closely linked to how much animals look like human beings and to what extent they can be anthropomorphised. Ostensibly, that can provide a formidable legal precedent to sanctioning robot abuse as well through criminal law.

---

[119] The research shows that *drasophilia,*a species of fruit fly, can sense when it is injured, can learn to avoid noxious stimuli, and can differentiate between the degrees of adverse experience. Still, it cannot have appropriate responses to different degrees of injury, or employ mitigating behaviours, such as rubbing, limping, or guarding the injury. Nonetheless, *tobacco hornworm* can perform all these behaviours. Not only are insects not protected, some bird species seem to have been given protected status due to their tendencies to prey on insects and thereby keep the insect population in control. Lynne U Sneddon and others, 'Defining And Assessing Animal Pain' (2014) 97 Animal Behaviour 201, 204.

Regarding robot abuse, it is clear that robot anthropomorphism creates new social interests that demand recognition by the law. With that perception shift of robots, offensive behaviours toward robots draw public attention.

As the social dimensions of human-robot interactions get stronger, such behaviours toward robots can be expected to become more likely to effect social interests in general morality.[120] The present study suggests that criminal punishments can be used to protect social interests affected by offensive behaviours toward robots. Sanctions for these behaviours can be modelled after those foreseen for the cruel treatment of animals.[121] Criminalisation of offensive behaviours toward robots cannot be precluded because no robot is capable of feeling any pain or suffering since not all criminal offences intend to uphold social interests in the protection of individual life. Indeed, there are plenty of victimless crimes in criminal law systems, such as recreational drug use, prostitution, and gambling, that are criminalised to vindicate social interests in general morals.[122] It is worth mentioning that some proposals in the literature advocate prohibition of certain groups of robots that are designed to be subjected to offensive behaviours, such as 'child sex robots and sex robots designed to refuse advances'.[123]

## Summary

This chapter first explored anthropomorphism and examined the influence of robots' distinctive characteristics on the occurrence of the phenomena. Then, undesirable outcomes of robot anthropomorphism were discussed and whether there is any justification for the human intuition which dictates it is morally wrong to 'torture' social robots, even though they do not 'experience' anything. Finally, the selected legal systems' animal protection legislation was briefly reviewed to determine whether any parallels between animal protection legislation and potential protection for robots can be established.

---

[120]; Roscoe Pound, 'A Survey of Social Interests' (1943) 57 Harvard Law Review 1, 25-26; D D Allan, Christoph Bartneck, and Andrew J Vonasch, 'The Doors of Social Robot Perception: The Influence of Implicit Self-Theories' (2021) 14 International Journal of Social Robotics 127, 128-129;

[121] Kamil Mamak, 'Whether to Save a Robot or a Human: On the Ethical and Legal Limits of Protections For Robots' (2021) 8 Frontiers in Robotics and AI.

[122] Louis Veneziano and Carol Veneziano, 'Are Victimless Crimes Actually Harmful?' (1993) 9 Journal of Contemporary Criminal Justice 1, 2.

[123] Danaher, 'Robotic Rape and Robotic Child Sexual Abuse (n 91) 35; Lily Frank and Sven Nyholm, 'Robot Sex and Consent: Is Consent to Sex Between a Robot and a Human Conceivable, Possible, and Desirable?' (2017) 25 Artificial Intelligence and Law 305, 320-32; Delphine DiTecco, 'New Technology, Same Old Stigma: An Analysis of Feminist Discourses and Sex Work Stigma in Sex Robot Media' (MA, Carleton University 2020) 22.

Through the comparison of animal protection legislation of chosen legal systems and different degrees of legal protection afforded to various animals, it is shown that when protection is afforded to animals, it is generally done in accordance with their mental complexity and similarity of the behaviours they demonstrate to those demonstrated by human beings - rather than just their ability to feel pain. The display of mental complexity seems to be why animals are protected, and the display of mental complexity also equates to some form of closeness to human beings.

It submitted that, in conclusion, animal sentience and or harm to animals is not a necessary prerequisite for their protection by the criminal justice system. By extension, the absence of harm to robots should not be thought to preclude the criminalisation of robot abuse.

# PART C

# *Free-ish*

## Implications of Robot Autonomy

In human beings, voluntary action is much slower than reflex action. But that's not the case with robots; with them it is merely a question of freedom of choice, otherwise the speeds of free and forced action are much the same.

—Isaac Asimov, *I, Robot*[*]

[*] Isaac Asimov, *I, Robot* (first published 1967, HarperVoyager 2018) 147.

# Chapter V: Challenges of Criminal Liability

> Those who think to master the industry are themselves mastered by it; robots must be produced although they are, or rather because they are, a war industry. The product of the human brain has at last escaped from the control of human hands. This is the comedy of science.
>
> —Karel Čapek, quoted by Ignatius F Clarke, *British Future Fiction*[1]

## Introduction

In January 2015, the Random Darknet Shopper ['RDS'] -an online shopping robot- become the first robot to be *arrested*.[2] RDS was a non-physical software entity, deployed as part of some interactive art exhibition. It was programmed to make random purchases on the dark web and then to have these items delivered to the gallery where some art exhibition was held. The RDS, along with the computer it was installed on, was confiscated by the Swiss police after it ordered ecstasy pills.[3] Eventually, the RDS was released without any charges being brought against itself or its masters, as law enforcement authorities considered its purchases 'reasonable means for sparking public debate about questions related to the exhibition'.[4] Still, it can be speculated that had the RDS not been deployed for artistic purposes, the legal persons involved with the design and operations, might have been prosecuted.[5] After all, even though the manufacturers and masters of the robot had neither designed nor intended the robot to purchase illicit drugs, they had nonetheless been reckless in releasing it into the dark web since they could have foreseen those unlawful items would eventually be purchased.

More recently, in October 2021, Ai-Da -the world's first robot artist- was detained by Egyptian authorities for ten days on suspicion of espionage. Ai-Da's eyes featured integrated modems and cameras that allowed it to create artworks.[6] According to the Egyptian authorities, these tools could also be used for spying and collecting sensitive information that could have utilised to threaten the national security of Egypt. Like the Random Dark Net Shopper, Ai-Da was finally released and allowed to appear in the exhibition that she contributed.

---

[1] Ignatius F Clarke, *British Future Fiction, 1700-1914,* Vol 8 (Routledge 2001) .
[2] Mike Power, 'What Happens When a Software Bot Goes on A Darknet Shopping Spree?' (*The Guardian*, 5 December 2014)
[3] Jana Kasperkevic, 'Swiss Police Release Robot That Bought Ecstasy Online' (*The Guardian*, 22 April 2015)
[4] ibid.
[5] Ryan Abbott, *The Reasonable Robot* (CUP 2020) 111.
[6] Nadia Khomami, 'Egypt Detains Artist Robot Ai-Da Before Historic Pyramid Show' (*The Guardian*, 20 October 2021).

Robots involved in these incidents were under human beings' control, as their behaviours were determined, or their operations were supervised in real-time by their human users. However, robots can also make and carry out decisions without the supervision of any human being. [7] Moreover, operating through the world wide web, some robots can engage in criminally offensive behaviours through decentralised networks. They are impossible to control once they are up and running,[8] and they can engage in independent activities that produce outcomes associated with specific criminal offences, unforeseeable to their manufacturers and masters.

As the technology develops, even more sophisticated robots that are more difficult to control or supervise will be introduced. [9] If and when robots engage in behaviours that would have been criminal had they been committed by natural or juridical (legal)persons, and there are no identifiable natural or juridical persons to whom culpability could be attributed, it is unclear on whom criminal liability might be imposed, and on what basis. As robots become more involved in everyday life and take on different roles, including as interaction partners, existing criminal law instruments will inevitably fall short.[10]

The following chapter explores whether robots themselves may be held criminally liable and convicted for the criminal behaviours in which they engage. To that end, the chapter consists of three sections. The first section of the chapter defines the term *robot crime* and introduces a framework for assessing the justifiability of robotic criminal liability. The second section, in turn examines the potential benefits of holding robots criminally liable and explores whether robotic criminal liability would fulfil the functions of criminal law systems, such as the rehabilitation of offenders, the deterrence of crime or the expression of society's condemnation of certain behaviours. The third section assesses the costs associated with imposing criminal liability on robots and, investigates whether punishing robots would have benefits that outweigh the drawbacks, and whether it remains within limitations as determined by the general principles of criminal law.

---

[7] David Sanderson, 'Ai-Da the Art Robot Released from Custody in Egypt' (*The Times*, 21 October 2021).

[8] Samer Hassan and Primavera De Filippi, 'Decentralized Autonomous Organization' (2021) 10 (2) Internet Policy Review 1, 4.

[9] According to the forecasts of the Future Laboratory, with the move towards automation of the workforce, approximately 35 percent of jobs done today by human beings is going to be performed by robots, and the ongoing developments in artificial intelligence and machine learning could enable robots to self-programme criminal activity, leading them the conclude that by 2040 more crime will be committed by robots than by human beings. Jeffrey Wale and David Yuratich, 'Robot Law: What Happens If Intelligent Machines Commit Crimes?' (*The Conversation*, 1 July 2015).
See also Davey Winder, 'Is the Future of Cyber Crime a Nightmare Scenario? - Raconteur' (*Raconteur*, 7 September 2016)

[10] Simon Chesterman, 'Artificial Intelligence and the Limits of Legal Personality' (2020) 69 (4) International and Comparative Law Quarterly 819, 829.

The idea of holding robots criminally liable is beginning to receive sustained attention in the literature. One leading advocate, Gabriel Hallevy, holds that robots are capable of satisfying the awareness criteria for criminal liability, namely, 'the mental element requirements for both intent offences and recklessness offenses.'[11] On the assumption that the presence of the essential physical element of criminal liability is self-evident, Hallevy argues that there is no reason not to hold robots liable as direct perpetrators of criminal offences should they fulfil the mental element requirement of criminal liability.[12]

Ying Hu, another proponent of the idea, draws parallels with corporate criminal liability and suggests that 'if there is reason to treat corporations as moral agents, there is reason to treat sophisticated robots as moral agents as well.'[13] Hu, in a similar manner with Hallevy, reasons that imposing criminal liability on robots makes sense as long as they can make, communicate, and carry out 'nontrivial morally relevant decisions.'[14] Others agree that robots can be viewed as possessing the cognitive states essential for fulfilling the mental element of criminal liability but emphasise that they must be equipped with a 'level of reason-responsiveness' in order to be properly punishable, that is, they should 'at least be prudentially rational, being able to anticipate the negative effects of punishment on the achievement of its goals. '[15]

Those who oppose the idea are often dismissive and quick to label it as an example of 'conceptual confusion'.[16] The authors in this camp, in general, see robots as mere functional artefacts regardless of however autonomous or sophisticated they may be, and thus believe that the attribution of any cognitive or mental state to robots, let alone *mens rea*, would be inconsistent with the factual reality.[17] Nonetheless, when it is considered that criminal law applies to legal entities that do not possess any mental capacity or any independent ability to perform voluntary acts, the present study holds that the idea of imposing criminal liability on robots is not easily dismissed.

---

[11] Gabriel Hallevy, *Liability for Crimes Involving Artificial Intelligence Systems* (Springer 2016) 99.
[12] Hallevy reasons that robots can be held responsible for crimes of negligence, or on a strict liability basis in certain specified circumstances, so long as the mental element of criminal liability is present. ibid 119; 135.
[13] Ying Hu, 'Robot Criminal Liability Revisited' in Soo Yoon, Sang Hoon Han, and Seong Jo Ahn (eds.), *Dangerous Ideas in Law* (Bobmunsa 2018) 498.
[14] Ying Hu, 'Robot Criminals' (2019) 52 University of Michigan Journal of Law Reform 487, 490.
[15] Francesca Lagioia and Giovanni Sartor, 'AI Systems Under Criminal Law: A Legal Analysis and a Regulatory Perspective' (2019) 33 Philosophy & Technology 433, 445, 462.
[16] Alexander Sarch, 'Who Cares What You Think? Criminal Culpability and the Irrelevance of Unmanifested Mental States' (2017) 36 Law and Philosophy 707, 708.
[17] ibid 709.

Criminal law systems are strictly anthropocentric, even more so than other legal domains, since they often focus on suppressing individual human beings' behaviours that disrupt the social order. Even human associations are not accepted as responsible subjects of criminal law in some of the criminal law systems; most notably, German criminal law refuses to recognise corporate criminal liability.[18]

On that account, it appears that non-human entities cannot possibly be considered as responsible subjects under current criminal laws, no matter how perceptive, independent, or intelligent they are.[19] Therefore, the present study submits that under current criminal laws of the national legal systems examined, robots can only be treated as criminal tools, putting them into same category with knives and guns-the category of instruments that are used to facilitate *actus reus* (guilty act) of their manufacturers or masters.[20] Unlike any other criminal tool, the capabilities associated with the characteristic of relative autonomy allow robots to perform criminally offensive behaviours unbeknownst to their manufacturers or masters. Considering that robots can make and carry out independent decisions thanks to the capabilities associated with the characteristic of relative autonomy, it must be admitted that they can also bring about undesirable outcomes that are associated with specific criminal offences by carrying out their independent decisions, which is something no gun or knife can do.[21]

## V.1 Robot Crime

The present study uses the term *robot crime* to refer to the independent activities of robots, that would have been considered as criminal offenses had they been performed by natural or legal persons. The manner in which the term *robot crime* is used purposefully excludes the situations where robots are intentionally or carelessly utilised by human beings to cause undesirable outcomes that are associated with specific criminal offences.

[18] V S Khanna, 'Corporate Criminal Liability: What Purpose Does It Serve?' (1996) 109 Harvard Law Review 1477, 1478; Susanne Beck, 'Intelligent Agents and Criminal Law—Negligence, Diffusion of Liability and Electronic Personhood' (2016) 86 Robotics and Autonomous Systems 138, 140; Sijad Allahverdiyev and Marvin Othman, '"Verbandssanktionengesetz"-Corporate Liability for Germany?' (2022) 23 German Law Journal 637, 639.
[19] Alice Ristroph, 'Criminal Law for Human Beings' in David Dyzenhaus and Thomas Poole (eds), *Hobbes and the Law* (CUP 2012) 97-99.
[20] Ugo Pagallo, *The Laws of Robots* (Springer 2013) 53.
[21] Gabriel Hallevy, *When Robots Kill: Artificial Intelligence Under Criminal Law* (Northeastern UP 2013) 52-57.

In the commission of such criminal offences, robots play roles identical to those of other criminal tools, and when determining criminal liability in cases involving the use of normal criminal tools, the existing instruments of criminal law can be utilised. [22]

Consequently, the manner in which the term *robot crime* is used allows the present study to focus on crimes where the characteristic of relative autonomy is central to the undesirable outcomes of crimes. The implications of the characteristic of relative autonomy in relation to the imposition of criminal liability are first analysed in terms of the reducibility of robots' criminally offensive behaviours to that of their manufacturers or masters. Ryan Abbott explains the concept of reducibility in his work as follows:

> Reducibility, which here refers to [the robot's] act being identified with individual criminal behavior, is also critical because if [the robot] engages in an act that would be criminal for a person and the act is reducible, then there will be a person that can be criminally liable. If [the robot's] act is not effectively reducible, there will not be another party that is aptly punished, in which case intuitively criminal activity can occur without the possibility of punishment.[23]

According to Abbott, most crimes in which robots are involved are reducible to legal persons' behaviours, including most of what would be categorised as robot crimes for the purposes of the present study.

For example, if some robot is designed in order to break into premises and steal goods, regardless of the degree of autonomy it is equipped, the criminal offence corresponding to outcomes of that robot's activity -breaking and entering- would be reducible to the to the act of the legal person who designed the robot. From the legal perspective, the legal person who designed that robot would be regarded as having committed the crime, and the robot would be viewed as the tool or weapon that is used to commit the crime. The situation there is not dissimilar to the cases where picklocks or similar tools are used for robbery since the natural person uses a tool -albeit a very advanced one- to achieve some prohibited consequence.[24]

There is no reason why crimes involving robots, as long as the undesirable outcomes associated with specific criminal offences are not caused by their distinctive characteristics, should not be treated in like manner.

---

[22] This statement also pertains to cases where or somebody's culpable carelessness in coding or using the robot results in harm, so long as the mental element requirement of a specific offence requires nothing more than the defendant to have behaved negligently.

[23] Abbott (n 5) 113.

[24] For example, Julius Caesar was killed after being stabbed twenty-three times, and it is universally accepted that daggers that inflicted fatal wounds did not commit the murder of the dictator, the conspirators who were wielding the daggers did. Antony Kamm, *Julius Caesar* (Routledge 2006) 148.

Furthermore, the present study holds that even when the undesirable outcomes associated with specific criminal offences are unforeseeable, certain robot crimes may remain reducible, so long as they may be construed as constructive liability crimes. Constructive criminal liability originates in the *versari in re illicita* doctrine. This doctrine, which goes back to medieval canon law, provides that any person engaging in unlawful conduct would be criminally liable for all consequences the conduct notwithstanding that they may lack *mens rea* in relation to these consequences.[25]

One popular example of constructive liability offences is the American crime of felony murder. Binder points to several cases to as examples of felony murder: [26]

> New York burglar William Ingram broke into a home, only to be met at the door by the homeowner, brandishing a pistol. The homeowner forced Ingram to lie down, bound him, and called the police. After police took Ingram away, the homeowner suffered a fatal heart attack. Ingram was convicted of felony murder.[27]
> John William Malaske, a young Oklahoma man, got a bottle of vodka for his underage sister and her two friends. One of the friends died of alcohol poisoning. Malaske was convicted of felony murder predicated on the felony of supplying alcohol to a minor. [28]

Though offences of constructive criminal liability persist as part of American law, such offences have been abolished in most of the other legal systems, including the jurisdictions selected for examination in the present study. The reason is that the *versari in re illicita of versari* doctrine violates the principle of *nulla poena sine culpa*.[29]

In fact, the offences of constructive criminal liability never found much favour in Ireland, and some authors go as far as to say the concept has never been applied in here.[30] However, the present study submits that one constructive liability offence did exist in Irish law, in the common law offence of assault occasioning actual bodily harm. Assault, in its non-aggravated form, is not a crime of violence.

---

[25] John R Spencer and Antje Pedain, 'Approaches to Strict and Constructive Liability in Continental Criminal Law', in Andrew Simester (ed), *Appraising Strict Liability* (OUP 2005)
Gardner cites the full maxim in Latin as: *versari in re illicita omnia imputantur quae ex delicto sequuntur,* which can be roughly translated as '[One] is [criminally] liable for all consequences of his unlawful act.' Martin R Gardner, 'The Mens Rea Enigma: Observations on The Role of Motive in the Criminal Law, Past and Present' [1993] Utah Law Review (3) 633, 705.

[26] Guyora Binder, 'Making the Best of Felony Murder' (2000) 91 Boston University Law Review 403, 406.

[27] *People v Ingram,* 492 NE2d 1220, 1220-21 (NY 1986).

[28] *Malaske v State,* 89 P3d 1116, 1117 n 1 (Okla Crim App 2004).

[29] *Nulla poena sine culpa*, or the culpability principle, is recognized as one of the human right by the Court of Justice of the European Union and the European Court of Human Rights.
See Giuseppina Panebianco, 'The Nulla Poena Sine Culpa Principle in European Courts Case Law', Human Rights in European Criminal Law (Springer 2015) 49-52.

[30] Peter Charleton and others, *Criminal Law and Evidence* (2nd edn, Bloomsbury 2020) 383.

Though the primary characteristic of assault is the invasion of another's body space, that may very well happen without any bodily contact, but instead with 'the apprehension of its imminence'.[31] Admittedly, those who commit the offence of non-aggravated assault were considered to have taken the risk that their conduct may occasionally cause bodily harm, and if that happened, they would have been deemed to have committed a more serious crime, assault occasioning actual bodily harm.[32] In other words, the common law offence required no *mens rea* apart from the *mens rea* for the non-aggravated assault and might have resulted in criminal liability even when 'actual bodily harm [came] about by a very indirect route, perhaps involving successive coincidences'.[33] The offence in question was one of constructive liability until it was abolished by Non-Fatal Offences Against the Person Act 1997.[34] Since the abolition of the offence of aggravated assault, no constructive liability offence exists in Irish law, and the introduction of new constructive liability offences do not seem tenable, as the Supreme Court of Ireland has connected the requirement of *mens rea* with the specific safeguards in the Constitution. [35] Though various forms of manslaughter in Irish law resemble the doctrine, even in these offences, the mens rea sought is specific to the criminal outcome.

In Italy, the *versari in re illicita* doctrine corresponds to the concept of *delitto preterintenzionale* or unintentional offences. CP groups criminal offences in three categories on the level of mental element required. Accordingly, criminal offences may be committed wilfully (according to the intention), through negligence (against the intention), or unintentionally (beyond the intention). [36]

In wilful offences, perpetrators know that their behaviours would bring about undesirable consequences and they want these consequences.[37] Negligent offences do no not require perpetrators to be aware of possible consequences; and though they could be aware of the possibility, they need not want such consequences.

---

[31] John Gardner, 'Rationality and thw Rule of Law in Offences Against the Person' (1994) 53 (3) The Cambridge Law Journal 502, 508.
[32] ibid.
[33] ibid 510. Gardner gives the case where person A lunges towards person B; person B is tackled backwards, is tripped by some third party's umbrella trips another, falls backwards into rail tracks and is run over by a train as an example. In that case, the assault would occasion harm even though it would not cause it, and the assault occasioning bodily harm offence would have been committed.
[34] Non-Fatal Offences Against the Person Act 1997, Art 28 (1) b.
[35] *People (DPP) v D'OT* [2003] IR 286; *CC v Ireland* [2006] 4 IR 66, 80.
[36] CP, Art 43.
[37] CP, Arts 42, 43 (1).

In unintentional offences, the perpetrators' acts or omissions result in unlawful consequences that are more severe than they intended. [38] As such, unintentional offences in Italian law, by definition, also similar to constructive liability offences in common law systems. In CP, there is only one unintentional offence: *omicidio preterintenzionale* or involuntary manslaughter. In Art 584, CP provides that 'whoever, by acts aimed at committing [assault and battery] causes the death of a human person is punished'. [39]

However, soon after the Constitution of the Italian Republic entered into force, the constitutionality of unintentional offences became questionable. Some legal scholars have consistently argued that Cost Art 27 (1) -which states that 'Criminal responsibility is personal'- actually safeguards the principle of culpability, that criminal responsibility must correspond to degrees of individual culpability, based on intent or negligence, and this view finally prevailed when the Constitutional Court of the Italian Republic explicitly adopted the same interpretation. [40] The Court thereby effectively declared that the imposition of strict and constructive liability on criminal law defendants incompatible with the Constitution, making it necessary for the legislator to rethink such offences in the form of culpability-based offences.

In German law, offences that are most similar to constructive liability offences are *erfolgsqualifizierte delikte* or result-qualified offences, referring to the criminal offences 'where the basic offence has a further, extended consequence that is not an element of that basic offence'.[41] Even though *erfolgsqualifizierte delikte* resemble constructive criminal liability offences, there is one major difference between the two about the requirement of *mens rea* for the extended consequence. Constructive criminal liability, in Irish and Italian laws, did not require any specific *mens rea* for the extended consequence, whereas under Section 18 of the German Criminal Code [StGB] at least *negligence* towards the extended consequence is required for *erfolgsqualifizierte delikte*:

> If the law imposes a more severe penalty based on a specific result of an offence, the offender or the participant is only liable to the more severe penalty in the event of being charged with at least negligence with respect to that result. [42]

StGB Section 18, therefore, explicitly affirms that German law does not recognise constructive liability offences.

---

[38] CP, art 43.
[39] CP, art 584.
[40] Corte Cost, sentenza 23-24 marzo 1988, n 364; Corte Cost, ordinanza 22-30 novembre 1988, n 1055.
[41] Michael Bohlander, *Principles of German Criminal Law* (Hart 2009) 31.
[42] StGB Section 18.

Moreover, the German Federal Constitutional Court, similar to its Italian counterpart, ascribes constitutional rank to the culpability principle, by establishing that this principle follows intrinsically from the substantive dimension of the *Rechtsstaatsprinzip* or the rule of law[43] and that it is 'rooted in human dignity and in a recognition of the human being's capacity for responsible agency, which the Constitution protects'.[44] To put it differently, criminal punishment of constructive liability offences would be unconstitutional in German law.[45]

The trend in the legal systems compared in the present study is toward abandoning the *versari in re illicita* doctrine and abolishing constructive liability offences that stemmed from the doctrine. Given that the highest courts in these legal systems accept that the culpability principle is protected by their respective constitutions, it is thought making the imposition of strict and constructive criminal liability unequivocally unconstitutional in all three countries, it does not appear likely new offences that are based on the culpability principle can be introduced. Even if the *versari in re illicita* doctrine is applied to robot crimes, it would still not be possible to impose criminal liability for all robot crimes. The reason is that the requirement for constructive criminal liability, related to the commission of some original offence for which the agent does exhibit *mens rea* (which has unintended unlawful consequences) cannot be satisfied. Due to the capabilities associated with the characteristic of relative autonomy, it is possible that, despite flawless care and diligence by the manufacturers or masters of robots, robot crimes might still occur. Abbott constructs the following scenario as an example of such eventualities:

> Experienced and expert programmers may separately contribute code for the software of an autonomous vehicle that unforeseeably results in the vehicle's attempt to collide with individuals wearing white-collared shirts and red-ties. The programmers here would not be criminally or even civilly liable-even though they have created a self-driving car that runs over investment bankers.[46]

As there will not be an original offence committed with *mens rea* in that event, constructive criminal liability rules cannot be utilised to reduce the robot crime in question to the action of some natural or juridical person.

---

[43] BVerfG October 25, 1966, 20 BVerfGE 323, 331.
[44] BVerfG February 26, 1968, 25 BVerfGE 269, 285.
[45] Bohlander (n 41) 31; Johan D Vyver, 'The International Criminal Court and the Concept of Mens Rea in International Criminal Law' (2004) 12 (1) University of Miami International and Comparative Law Review 57, 73.
[46] Ryan Abbott and Alex Sarch, 'Punishing Artificial Intelligence: Legal Fiction or Science Fiction' (2019) 53 UC Davis Law Review 323, 335.

## V.2 Irreducibility of Robot Crimes

On the assumption that the imposition of criminal liability should always be dependent on the free will of the defendant, be it intent[47] or negligence,[48] it is evaluated there are two possible grounds on which robot crimes may be deemed irreducible to individual criminal behaviours. The first of these grounds is concerned with the problems regarding the enforcement of criminal laws, that it may not always be possible or feasible to identify the legal person behind the robot crime, even if there is the one. The second ground for irreducibility is related to the criminal law policy, that it might not always be in the best interests of the general public, or be fair, just, and reasonable to impose liability on the human actors behind the robot crime.

The enforcement of criminal law rules with perfect efficiency -with no guilty unpunished and no innocent sanctioned- is virtually impossible, as is the case for the legal norms that belong to other legal domains. Indeed, no matter how coherent, consistent, and detailed the legal norms themselves may be, the degree that they can be enforced ultimately relies on the availability of material and personnel sources that can be dedicated for that purpose-after all, legal norms do not enforce themselves. In other words, problems related to the enforcement of legal norms are economic – they are concerned with the distribution of limited resources to meet the needs of legal security and certainty, and these problems did not arise with the robots and their characteristic of relative autonomy in the context of criminal law. For criminal liability to be imposed, it should be clear beyond reasonable doubt that the perpetrator has violated the criminal law rule that prohibits the offensive course of action. If the perpetrator cannot be identified, then no criminal liability can be imposed; history is rich with cases where the perpetrators could not be identified and therefore no enforcement was possible.[49] It is submitted that enforcement problems are likely to arise in two scenarios in relation to robot crimes:

---

[47] Here, the term *intent* refers to knowingly and wilfully bringing about the prohibited consequence.

[48] Here, the term *negligence* refers to willingly taking an initial action that according to a reasonable person would lead to the prohibited consequence or failing to exercise due diligence to prevent the prohibited consequence.

[49] Examples of such cases before include the Brussels Airport diamond heist or the 'Tamam Shud' murder case, among many others.
See, e.g., Angelique Chrisafis, 'Diamond Heist at Brussels Airport Nets Gang Up To £30M In Gems' *The Guardian* (19 Feb 2013); Daniel Keane, 'The Six Clues That Have Failed to Solve the Somerton Man Mystery' (*Abc.net.au*, 2021)

**(1) when it is impossible to identify who are instructing robots to deliver prohibited outcomes:**

The first scenario is expected to occur more frequently since the growing autonomy of robots and their increasingly complex algorithmic processes make their use for the commission of criminal offences easier while concurrently making it harder to identify those who are using them to produce prohibited consequences. One example of the first scenario would be where some computer programmer builds autonomous software agents to hack into government computer networks and manages to remain anonymous.[50]

**(2) when it is possible to identify perpetrators but doing so would not be in the public interest.**

The second scenario stems from the fact that law enforcement is a public service, funded by taxpayers and therefore with public resources, tasked with serving all members of society.[51] If the law enforcement spends disproportionate amounts of effort, money, and time to investigate relatively unimportant offences, then it can be seen as wasting public resources. Wasting public resources in that manner is harmful to society at large, and if it occurs often, there may be insufficient resources left to provide other public services. Due to the increasing sophistication of robots, identification of actual perpetrators that provide instructions to deliver prohibited outcomes is expected to become prohibitively expensive and remarkably demanding,[52] having the potential to drain the resources allocated to law enforcement.

It is submitted that it may become counter-productive for law enforcement to try and 'reduce' all robot crimes to actions of natural or legal persons. Especially when the damages caused are minimal and the enforcement costs are likely to outweigh the benefits of imposing criminal liability, the robot crime in question can be deemed to be irreducible for all practical purposes. [53]

The second potential ground for the irreducibility of robot crimes is related to the policy of criminal liability. There are some robot crimes that are expected to be traced back to the culpability o legal persons quite effortlessly, thus making the actual perpetrators controlling robots being easily identifiable.

---

[50] Abbott and Sarch (n 46) 336.
[51] ibid 337.
[52] Ann Thanaraj, 'Can Robots Be Prosecuted For A Crime?' (*University of Cumbria*, 14 September 2018).
[53] *Principles to Underpin Management by Public Entities of Funding (*Office of the Auditor-General 2006) 17-23.

However, even then, it may still be bad criminal law policy to reduce the robot crime to the conduct of some natural or legal person, as the culpability of the individual behaviour that resulted in the undesirable consequences that are associated with specific criminal offences can be regarded too trivial to criminalise.[54]

For example, suppose there are multiple human beings who work to develop some robot, and they sometimes act carelessly in rather insubstantial ways while working (e.g., forgetting to stub their cigarettes) and over the time, such acts result in the robot teaching itself how to play pranks. Finally, as part of some very elaborate plan for a prank, the robot sets some nearby buildings on fire, causing several hundred human beings to die. Here, the risk posed by each individual human behaviour is so insubstantial that it cannot reasonably warrant imposing any criminal liability, and in spite of the grim outcome, the robot crime must be in question regarded as irreducible for policy reasons.

The discussion above regarding the irreducibility of robot crimes indicate that, for the near future, large portions of robot crimes are no longer going to be reducible to behaviours of natural or legal persons. In addition, though some robot crimes are likely to remain reducible to conducts of natural or legal persons, that exercise is not always going to result in criminal liability for natural or legal persons. These findings lead to the conclusion that the irreducibility of robot crimes is inevitable within the framework of existing criminal law instruments, however, they do not prove that robots themselves should be held criminally liable.

## V.3 Justification of Criminal Liability

Theoretical and philosophical debates about criminal punishments focus on either the definitions of criminal punishment or the justifications for imposing it.[55] The present study adopts the broadly accepted Flew-Benn-Hart definition of criminal punishment, formulated by Hart in terms of five characteristics:

1. It must involve pain or other consequences normally considered unpleasant.
2. It must be for an offence against legal rules.
3. It must be of an actual or supposed offender for his offence.
4. It must be intentionally administered by human beings other than the offender.
5. It must be imposed and administered by an authority constituted by a legal system against which the offence is committed. [56]

[54] H L A Hart, *Punishment and Responsibility: Essays in the Philosophy of Law* (OUP 2008) 23-28.
[55] Thomas McPherson, 'Punishment: Definition and Justification' (1967) 28 Analysis 21, 22.
[56] Hart (n 50) 24-27.

According to that definition, 'criminal punishments can only be administered after the conviction of the defendant for a legally recognised criminal offence before the courts of law, provided that the conviction is decided through the established legal procedure and within due process of law. When it comes to the justifiability of criminal punishment present study assumes that the criminal punishment is just if the social utility expected by the imposition of punishment outweighs its costs, provided that the punishment breach does not any fundamental limitations by criminal systems . Furthermore, as criminal law instruments should always be utilised as the last resort to exert social control over undesirable human behaviour, no criminal punishment is justified unless there are no better, feasible alternatives that can deliver comparable outcomes.

The affirmative justifications for criminal punishments are the social benefits that area intended by the imposition of punishments may bring about, such as prevention and reduction of future undesirable conducts, increased protection afforded to the interest in safety, improved levels of general peace and enhanced sense of order, or expression of commitment to certain moral standards central to society.[57] The negative limitations on criminal punishments are based on the social commitment that 'punishment should not violate deeply held normative commitments, such as justice and fairness'.[58] negative limitations focus on the blameworthiness of the defendant,[59] not regarding whether they committed the prohibited act but rather if they could have been expected to apprehend the moral wrongness of using their free will toward undesirable outcome. The present study assesses as the most fundamental of these limitation: requirement for culpability and the desert constraint.

The desert constraint stipulates that no one should be punished beyond what they deserve, even if doing so would have produced broader social benefits.[60] Other limitations on imposing criminal punishment include the capacity for culpability and the requirement for a voluntary act.[61]

---

[57] Antony Duff and Zachary Hoskins, 'Legal Punishment' *The Stanford Encyclopaedia of Philosophy* (Summer edn, 2021).

[58] Abbott and Sarch (n 46) 341.

[59] Sofia M I Jeppsson, 'Retributivism, Justification and Credence: The Epistemic Argument Revisited' [2020] 14 Neuroethics 177.

[60] If a 'pure' consequentialist account is adopted for justification of punishment, one has to deem blatantly unjust punishments justified so long as they reduce crime or otherwise benefit society. See Antony Duff and Zachary Hoskins, 'Legal Punishment' *The Stanford Encyclopaedia of Philosophy* (Summer edn, 2021)

[61] Antony Duff, *The Realm of Criminal Law* (OUP 2018) 20; Nicola Lacey and Hanna Pickard, 'To Blame or To Forgive? Reconciling Punishment and Forgiveness in Criminal Justice' (2015) 35 Oxford Journal of Legal Studies 665, 666.

### V.3.1 Affirmative Justifications

Since the criminal punishment often involves denial of the individual interest in liberty and deliberate infliction of pain and suffering by the State, the imposition of criminal punishment can only be read regarded as legitimate, or excusable, only if the positive outcomes for society outweigh negative outcomes on to punished parties. Regarding the positive outcomes expected from the imposition of criminal punishment, theories of criminal punishment can be divided into two camps, retributivist, and consequentialist. Retributivist theories argue that criminal punishments are legitimate so long as these punishments are deserved according to the principle of justice, even in any cases where that would have produced negative outcomes on society. As such, it can be argued that retributivist theories attach the greatest value is to the social interest in the satisfaction feeling of justice. Consequentialist accounts, on the other hand, hold that criminal punishment cannot be legitimate when doing so would have negative outcomes for society.

#### V.3.1.a Retributivist Justifications

Retribution is possibly the oldest justification for the imposition of criminal punishments, since it essentially enables the transformation of victims' desires for revenge into demonstrations of the fact that any wrongdoing is subject to punishment that it deserves.[62] Perhaps due to the long history of the concept, there are multiple retributivist accounts on the justification of criminal punishment.

Still, all of these accounts hold that the wrongness of the conduct that is being punished justifies the suffering inflicted by criminal punishment.[63] In other words, according to the retributivists, the fact that one offender has committed some legally established criminal offence is the necessary and sufficient condition for the imposition of criminal punishment.

Moreover, retributivist justifications of criminal punishment incorporate the principle of proportionality and hold that the severity of punishment should be, to the extent possible, in proportion to the culpability of the behaviour that is punished.

---

[62] Lex talionis (law of retribution) was the core concept in the laws of Hammurabi, introduced back in 2000 BC. See George E Vincent, 'The Laws of Hammurabi' (1904) 9 American Journal of Sociology 744.

[63] Victor Tadros, *The Ends of Harm* (OUP 2011) 26; John Cottingham, 'Varieties of Retribution' (1979) 29 (116) The Philosophical Quarterly 238, 239; Douglas Husak, 'Retributivism in Extremis' (2012) 32 Law and Philosophy 3, 5.

One account of retributivist justifications associate retribution with revenge or retaliation. Accordingly, this account argues that wrongdoers should be made to suffer through the imposition and execution of criminal punishments since they inflicted suffering on others through their conduct. The ancient principle of *lex talionis*, developed in Babylonian law and present in both biblical and early Roman law, is often mentioned in relation to tat account of retributivist justifications. Biblical expression of the principle is as follows:

> And if any mischief follows, then thou shalt give life for life, eye for eye, tooth for tooth.[64] Breach for breach, eye for eye, tooth for tooth: as he hath caused a blemish in a man, so shall it be done to him again. [65]

Another account of retributivist justifications argues that the imposition of criminal punishment is justified as it expresses society's condemnation of offensive conduct.[66] The expression of condemnation brings society together in common feeling when certain boundaries are crossed. [67] Moreover, expressing condemnation of offensive conducts provides victims of such conducts with 'a sense of vindication and satisfaction' and increases the sense of security among not only victims but also society in general.

That expressivist account of retributivist justifications calls for imposing criminal liability on robots upon consideration of the phenomenon of robot anthropomorphism. It is submitted that, since criminal law instruments reflect society's perceptions of praiseworthy and blameworthy human conduct, criminal law should somehow give effect to robot anthropomorphism, as well.[68]

Indeed, if criminal law fails to express condemnation of robot crimes despite robots themselves being popularly perceived as blameworthy, it may erode the confidence in criminal law as it applies to human beings. On this point, one author, Christina Mulligan, goes even further and states that punishing wrongdoing robots may be necessary to provide psychological satisfaction to those whom they harm.[69]

---

[64] Exodus 21: 23-35.
[65] Leviticus 24:19-21
[66] Paul H Robinson, *Intuitions of Justice and the Utility of Desert* (OUP 2013) 162.
[67] Abbott and Sarch (n 46) 346.
[68] Indeed, human beings' innate cognitive responses even extend to legal fictions, i.e., corporations see Mihailis E Diamantis, 'Corporate Criminal Minds' (2015) 91 Notre Dame Law Review 2049, 2089.
[69] Christina Mulligan, 'Revenge Against Robots' (2018) 69 South Carolina Law Review 579, 580.

However, expressing condemnation of robots' offensive conducts may also communicate that the law views robots as actors equal to natural persons, at least in terms of being held accountable.[70] In the present day, there are no natural or juridical persons held accountable without having any rights; as such, communication of the message mentioned above might propagate the idea that robots have rights to certain kinds of benefits, protections, and dignities. [71] There, scientific research which establishes that intuitions of justice are broadly consistent with the principle of retribution. [72] However, even if it is assumed that such accounts provide ample justification for the imposition of criminal punishment, since the criminal punishment is expected to be perceived as undesirable by the offender the application of that justification is not very straightforward in relation to robots. After all, robots are functional artefacts that can independently make increasingly complicated decisions, but they cannot actually care about anything, or any sensibilities that can be considered equivalent the human ability to attribute value to experiences. John Danaher defines tithe discord between the human beings' innate desire for retribution and the absence of appropriate subjects of retributive punishment as the retribution gap. Accordingly, human beings look for a culpable wrongdoer deserving of criminal punishment upon and justify their intuitions with retributive motives, but robots cannot be not appropriate subjects of these retributive attitudes as they lack the required mental states to experience punishments as 'unpleasant'. [73] Most retributivist accounts rely on the society's judgement of the moral value of criminally offensive behaviours. As such, retributive justification of criminal punishment is reliant on the moral agency of the punished party.

The possibility of direct imposition of criminal liability on robots is rejected categorically by some authors as 'criminal liability mirrors moral blameworthiness'.[74] These authors argue that as robots do not have moral agency, they cannot be attributed any moral blame, and supposing that criminal liability reflects moral accountability for the commission of the criminal offence, they could not be imposed criminal liability.[75]

---

[70] Michael S Moore, 'Justifying Retributivism' (1993) 27 Israel Law Review 15, 18.

[71] Markus Kneer and Michael T Stuart, 'Playing The Blame Game With Robots' [2021] Companion of the 2021 ACM/IEEE International Conference on Human-Robot Interaction 405, 410.

[72] Kevin M Carlsmith and John M Darley, 'Psychological Aspects of Retributive Justice' (2008) 40 Advances in Experimental Social Psychology 193, 194; Keith Jensen, 'Punishment And Spite, The Dark Side Of Cooperation' (2010) 365 Philosophical Transactions of the Royal Society: Biological Sciences 2635, 2638-2640.

[73] John Danaher, 'Robots, Law and the Retribution Gap' (2016) 18 Ethics and Information Technology 299, 302-305.

[74] David Lefkowitz , 'Blame and The Criminal Law' (2015) 6 Jurisprudence 451, 452; Alexander Sarch, 'Should Criminal Law Mirror Moral Blameworthiness or Criminal Culpability? A Reply to Husak' (2021) 41 Law and Philosophy 305, 331.

[75] Moral agency is defined as the ability to make choices based on some notions of right and wrong.

However, even though the legal values that are upheld by criminal law systems and enforced through liberty-depriving criminal punishments might have been inspired by moral values, it is observed that in national legal systems chosen. Criminal law systems exclusively focus on what behaviours are impermissible and whether such behaviours are products of the defendant's use of their free will.

In general, criminal law systems do not question whether any subjective moral justifications exist for the defendant's behaviours, not at least until the sentencing stage.[76] When defendants perform criminally offensive behaviours, so long as these behaviours are performed because of the use of their free will, they are imposed criminal liability, regardless of their subjective moral perspectives. Consequently, the present study notes these authors are not accurate in their claims that the fact that robots cannot make any moral judgments about their behaviours does not preclude the possibility of direct imposition of criminal liability on them if criminally offensive behaviours stem from their autonomy.[77] In other words, for there to be any retributive justification for the criminally offensive behaviour of any given entity, the entity in question must be able to be held morally blameworthy for that behaviour, and moral blameworthiness requires the entity to be able to make moral value judgements about the potential courses of conduct that it can follow. Robots, regardless of whatever characteristics they have and to what degree they can exercise the capabilities that facilitate these characteristics, cannot make value judgements about the moral worth of their behaviours. Since robots cannot attribute any moral value -good or bad- to the possible courses of conduct, it follows that they cannot refrain from performing some behaviours because that would be morally wrong to do so; nor can they decide to perform any behaviour because of its moral worth.

In short, robots have no moral decision-making capabilities or moral agency, and since they have no moral agency, they cannot be ascribed any moral blameworthiness, and therefore imposing criminal liability on robots would not achieve any retributive outcome. The present study agrees with the last argument and submits that robot anthropomorphism does not provide sufficient grounds for retributive justification for criminal punishment.

---

See Vinit Haksar, 'Moral Agents' *Routledge Encyclopaedia of Philosophy* (1998).

[76] Richard C Fuller, 'Morals and the Criminal Law' (1942) 32 Journal of Criminal Law and Criminology 624, 625.

[77] Hallevy, *When Robots Kill* (n 21) 69.

However, the imposition of criminal liability on non-human entities, such as animals that injured or killed human beings the common practice across the criminal systems of the European continent. In these trials, there was deemed to be some retributivist justification. Even though animals that were imposed criminal punishments in these trials did not do have any capability to make moral value judgements, they showed independence (by definition, living entities are self-operating biological mechanisms,) and some of them (pigs) displayed basic levels of reasoning and planning. Along with the similar reactions of pain and suffering in animals, it appears that the phenomenon of anthropomorphism invoked by animals in medieval society strong enough to attribute the human capacity to make moral value judgements on animals [78] The practice of animal trials, in the present day, is often dismissed as absurd and inhumane. The reason for the prevalent dismissal of animal trials must be that the attribution of moral agency cannot substitute the presence of actual moral evaluation ability. [79]

In relation to animal trials, however, though the moral agency attributed to animals was illusory, criminal punishments imposed on them did have real negative outcomes perceived such- most of the animals brought before the courts were subjected cruel treatment and painful deaths. However, as robots are not capable of feeling pain or suffering punishments, the achievement of the objective of retribution would be as illusory as the pain or suffering brought about by punishment. The imposition of criminal punishments on robots can therefore satisfy the collective demand in society for having the criminally offensive behaviours punished, closing the retribution gap, at least in appearance.

### V.3.1.b Consequentialist Accounts

Consequentialist accounts, in contrast with retributivist justifications, are forward-looking.  in that they are based on the understanding that the imposition of criminal punishment is justified provided that its future benefits outweigh its potential costs to society in general and to the individuals that are being punished.

[78] Hampton Carson, 'The Trial of Animals and Insects. A Little-Known Chapter of Medieval Jurisprudence' (1917) 56 Proceedings of the American Philosophical Society 410
[79] Larry Alexander, Kimberly Kessler-Ferzan, and Stephen J Morse, *Crime and Culpability* (CUP 2009) 145.

Consequentialist accounts on justification of criminal punishment treats these punishments as instruments that help improve the general conditions in social life by reducing the occurrence of harmful, criminally offensive behaviours. Unlike the instruments offered by private law systems, neither the criminal law nor the punishments contained therein are concerned with repairing the damages caused by offensive behaviours, thus, criminal punishments can only reduce harmful conducts by preventing the occurrence of future crimes to the degree that is possible.[80] It is surmised that the imposition of criminal punishments can help reduce the future crimes three ways. The most straightforward way is incapacitation: When the perpetrators are deprived of liberty and personal freedom of movement, they are also prevented from committing further crimes.[81] The second way through which the reduction of future crimes is deterrence. By punishing some specific offender, criminal law discourages that offender from committing any more crimes in the future.[82] In addition, punishment of that specific offender discourages other potential offenders from committing similar crimes, since they know that they will likely face similar negative consequences. Finally, the last way through which commission of future crimes can be prevented is rehabilitation.[83] If perpetrators are provided with assistance to see errors of the conduct during incarceration and allowed to receive some training, it is more likely that they will not commit crimes after being released. [84] When the criminal execution codes of legal systems chosen for the comparison in the present study is examined, we find that all three legal systems place emphasis on first deterrence, second incapacitation, and third rehabilitation of the offenders.

In the present day, Common Law system (Ireland) is reluctant to establish the main purpose of criminal punishment whereas Civil Law legal systems have already elaborated on the primary objective to be achieved by executing criminal penalties either by legislation or by doctrine, strong traditions in rejecting retributivist approaches in favour of either the pursuit of the social interest in strengthening general peace and security (Germany) or the pursuit of the social interest in maintaining highest possible level of socio-cultural, economical and too much lesser degree, political equality (Italy).

---

[80] Per Duff and Hoskins, the foremost good that any punishment can bring is the reduction of crime. That is because the proponents of consequentialist accounts defend only criminalising harmful behaviours. Accordingly, in reducing crime, they aim to reduce the harm caused by crimes. Duff and Hoskins (n 53).
[81] Antony Duff, *The Realm of Criminal Law* (OUP 2018) 20.
[82] Abbott (n 5) 116.
[83] Lisa Forsberg and Thomas Douglas, 'What Is Criminal Rehabilitation?' [2020] Criminal Law and Philosophy.
[84] ibid.

In German Law, legislation identifies the objective of punishment. Indeed, the Prison Act determines the objective of executing the most essential criminal law punishment, imprisonment, as follows.

> By serving his prison sentence the prisoner shall be enabled in future to lead a life in social responsibility without committing criminal offences (objective of treatment). The execution of the prison sentence shall also serve to protect the general public from further criminal offences. .[85]

In the first sentence of the cited paragraph, it is observed that the treatment of the wrongdoer, or the of rehabilitation of the prisoner is determined as the ultimate goal of the incarceration punishments. That sentence, however, is written in the passive tense and does not explicitly assign the task of the achievement of that purpose to the criminal law system. the sentence is formulated using somewhat nebulous language that one can even argue that's it assigns that task to the prisoner.

Moreover, one can also argue that the sentence does not bring any command of law but instead merely observes that serving the prison sentence is the precondition for any prisoner to rejoint social life. In the German original of the text the sentence ends with the determinative term v*ollzugsziel* between parentheses. Though in the federal government commissioned translation of the Prison Act, that term is translated as 'objective of treatment'; the present study defends that the term can equally, or perhaps even more accurately, be translated as 'purpose of enforcement/execution'. Since the German original is the binding version of the Prison Act, it should be admitted that the objective of rehabilitation is not as explicitly given the status of the main and primary objective of the incarceration punishment as suggested by the English version. the second sentence is clearer regarding its imposition of the task of 'protecting general public from further offences' to the criminal law system.

In Italian law, the purpose of punishment is defined by Cost Art 27 (3):

> Punishment cannot consist in treatment contrary to human dignity and must aim at rehabilitating the offender.

However, Cost Art 27 has been only acknowledged and implemented by the Italian legislature since the penitentiary reforms of 1975. The main and primary objective of the Italian criminal law system since then must be regarded as the rehabilitation of the offender.

[85] Strafvollzugsgesetz [StVollzG][Prison Act], March 16, 1976, BGB1 at 581, last amended by Art 27 of the Law of October 5, 202, BGBl I at 4607, Section 2.

Irish legislation does not directly express what the purpose of criminal punishment is. However, the Ministry of Justice describes the purpose of Irish penal system as to ensure that crimes are punished while offenders are rehabilitated and reintegrated in society. Though the purpose of imposing criminal punishments is not clearly defined in Irish criminal law system as it is in German and Italian Criminal law systems, but nonetheless it as well seems to be focused on utilitarian justifications of criminal punishments and no longer on the retributivist justifications. Accordingly, the tendency of legal systems chosen for comparison for the present study is towards giving less and less weight to retributivist justifications when assessing the imposition of criminal punishment is justified or not.

Opponents of the direct imposition of criminal liability on robots argue that punishing robots cannot deliver the outcome of reducing the rate of occurrence of criminally offensive behaviours of robots since robots cannot understand the wrongfulness of their offensive behaviours and therefore cannot be deterred or rehabilitated. Accordingly, the objective of deterrence can only be achieved in relation to entities that can establish similarities between their potential courses of conduct and behaviours of entities who are punished. [86] Unless the entity in question can recognise such similarities, punishing other entities cannot possibly deter that entity from performing the same criminal behaviour.[87] At present, there are no robots that are programmed to respond to others' punishment in ways that can be interpreted as 'being deterred', but robots that are responsive to punishment imposed on themselves and others are considered to be theoretically possible, especially with rapid advances in machine learning technologies.

Though the imposition of criminal punishment on robots cannot bring about specific deterrence, it can nonetheless bring about some rather indirect benefits, for instance, through general deterrence.[88] The imposition of criminal punishment on the robot could also deprive the manufacturers and masters of the financial gains that they would otherwise have made and 'deter' both them and other manufacturers from creating robots that tend to perform criminally offensive behaviours. [89]

---

[86] Hallevy, *When Robots Kill* (n 21) 157-168; Gabriel Lima and others, 'The Conflict Between People's Urge to Punish AI and Legal Systems' (2021) 8 Frontiers in Robotics and AI 6.
[87] Peter Asaro, 'A Body to Kick, but Still No Soul to Damn: Legal Perspectives on Robotics' in Patrick Lin, Keith Abney and George A Bekey (eds), *Robot Ethics: The Ethical and Social Implications of Robotics* (MIT 2014) 181.
[88] Hart (n 50) 19.
[89] Abbott and Sarch (n 42) 346.

Opponents of robotic criminal liability also overlook the possibility that imposing criminal liability on robots can bring about some preventive benefits incapacitation or rehabilitation. Preventive benefits, through incapacitation, can be achieved if the associated punishment is limiting the functioning of the robot or in most severe cases shutting down the 'criminal' robot in question - the robot death penalty.[90] Similarly, with punishments that involve measures such as reprogramming the software of the robot to eliminate the decision-making pathways that produced the behaviour that resulted in criminally offensive outcome, rehabilitation can also be achieved.

### V.3.2 Negative Limitations

The imposition of criminal punishment is one of the most effective tools for upholding the values that determine the purposes of criminal law systems. However, criminal punishment is also quite impactful on interests of individuals, such as the interest in personal freedom and liberty. In that sense, legal systems accept that the imposition of any criminal punishment should strive the balance of social interests that are going to be achieved and the implications on individual interests. It is widely accepted that no criminal punishment should transgress core commitments that are common to all modern legal orders, such as justice and fairness. These commitments, in the context of criminal law, can be expressed as negative limitations on punishment – if they are breached the imposition criminal punishment, especially imprisonment cannot be regarded as justified. For the purposes of the present study, we find two of the negative limitations relevant, the desert constraint and the requirement of capacity for culpability.

#### V.3.2.a Desert Constraint

Perhaps the most fundamental limitation on punishments is the desert constraint. The desert constraint stipulates that one cannot be imposed any criminal punishment that goes beyond what is deserved by the criminal. The desert constraint focuses on the blameworthiness of conduct, provides that one cannot be punished more than one deserves.[91] The blameworthiness of the conduct here refers to the society's judgement on the moral value (on the degree of wrongness of the conduct) expressed by the criminal law systems, but it is not concerned with the moral agency of the criminal offender.

[90] Mark A Lemley and Bryan Casey, 'Remedies for Robots' (2019) 86 University of Chicago Law Review 1313, 1389.
[91] Robinson (n 62) 162.

It can be argued that the desert constraint is another restatement of the principle of proportionality, as it holds that punishments that go beyond what is proportionate to society's collective judgement of the gravity of wrongness of the criminally offensive behaviours. However, that does not necessarily mean that the offender must have the capability to make evaluations of the moral worth of the potential courses of action.

The present study espouses the view that no social utility outweighs the failure to recognise one's inherent dignity, therefore no violations of the limitations on the desert constraint is deemed to be justified. That said, since robots do not feel pain or have any vested interests, it is not possible to argue that they would benefit from the desert constraint.

### V.3.2.b Capacity of Culpability

The desert constraint is not the only negative limitation on punishment. Equally important are the characteristics the wrongdoers must possess for them to be properly punished. The default position in criminal law doctrine is that punishment can only be properly imposed in response to culpable wrongdoing. Wrongdoers who cannot deliberate and lack the ability to take decisions based on their conscious beliefs have no capacity for culpability and therefore cannot be appropriately subjected to punishment. This limitation is assessed to be best evidenced by the formulations of incapacity or insanity defences across criminal liability across selected legal systems. [92]

The prevailing opinion on robot crimes stems from the idea that robots, like any other artificially intelligent agent, cannot be regarded as in possession of the capacity for culpability. Three different shortcomings of robots are pointed to in support of this view: that they lack (1) mental states needed for culpability, (2) ability to engage in voluntary acts, (3) consciousness and therefore the ability to be truly punished.

[92] In Ireland, the defence of insanity is regulated under The Criminal Law (Insanity) Act 2006. The insanity defence under Section 5 of the Act allows the defendant to seek an exemption from criminal liability based on the absence of mens rea. The mere presence of mental disorder when the crime was committed does not guarantee the success of the defence of insanity at the trial. Per Section 5(1) (b), there are three further requirements: not knowing the nature and quality of the act, not knowing that what they were doing was wrong or was unable to refrain from committing the act. In Germany, according to the StGB Section 20, those who perform offensive behaviours because a mental disorder makes them unable to see the wrongness of the behaviour or act on such insight are considered not guilty. Likewise, PC Art 88 provides that anyone who was (at the moment in which they perform the wrong behaviour) unable to understand the nature of their behaviour or want the prohibited outcome due to some mental disorder cannot be held liable.

It is argued that, because of these incompetences of robots, no robot actions can satisfy the elements of liability that are common to all criminal offences, and punishing them regardless would be in breach of the principle of legality. [93]

The principle of legality stems from the rule of law, and, in the context of criminal law, it is understood to indicate that only the law can define crimes and prescribe punishments.[94] In other words, unless a defendant satisfies all elements of a crime, it would be contrary to the rule of law to convict the defendant of the crime. Thus, it can be inferred that the eligibility challenge implies that punishing robots would threaten the rule of law and undermine society's trust in criminal law instruments.

### *V.3.3* Possible Alternatives

*Strict Liability Crimes*

New forms of strict liability crimes may allow the punishment of robot crimes through by-passing the requirement of the capacity for culpability. In strict liability crimes, the *actus reus* (guilty act) element is sufficient for the imposition of criminal liability. The opinion in the literature is that strict liability offences are often unjustified since such offences allow blameless persons to be imposed criminal liability. There is, consequently, strong resistance against new strict liability crimes.[95]

Be that as it may, since strict liability crimes do not require the element of *mens rea.* they can be used to impose criminal liability for robot crimes. Indeed, it is becoming increasingly probable that robots may cause undesirable outcomes in the absence of any form of *mens rea* attributable to legal persons-with no intent to cause harm, no knowledge of any inculpatory facts, no reckless disregard of risks, and no negligent unawareness of risks on the part of robots' manufacturers, owners and users. Nonetheless, when any of these legal persons are prosecuted for strict liability crimes, they must still meet the voluntary act requirement.[96]

---

[93] Abbott (n 5) 118.
[94] Peter Westen, 'Two Rules of Legality in Criminal Law' (2007) 26 Law and Philosophy 229, 232.
[95] Andrew P Simester, 'Is Strict Liability Always Wrong?' in Andrew P Simester (ed), *Appraising Strict Liability* (OUP 2005) 21-23; Andrew Ashworth, 'Should Strict Criminal Liability Be Removed from All Imprisonable Offences?' (2010) 45 Irish Jurist 1, 3.
[96] Wayne R LaFave, *Substantive Criminal Law (Vol 1)* (3rd edn, Thomson Reuters 2018), s 6, sub-s 6.1(c).

In the context of criminal law, it is found that only the 'bodily movements guided by conscious mental representations'[97] qualify as voluntary acts; with reflexes, convulsions, or movements that happen unconsciously explicitly ruled out as involuntary human acts or reflexes. Should the same requirement be sought for robot crimes and robots are regarded as having no awareness of their actions, no robot conduct can be deemed as voluntary and therefore punishable.

However, the reason for the present study's opposition to strict criminal liability for robot crimes is not the aforementioned argument regarding the lack of awareness of robots. Instead, the present study subscribes to the principle which stipulates that criminal liability should be based on culpability and, in the face of almost unanimous condemnation of strict criminal liability offences, the present study also holds that does not seem likely for the robot crimes to be addressed as new forms of strict liability crimes, even if robots were regarded to be 'aware'. [98]

*Corporate Criminal Liability*

Corporations do not have any physical or mental presence that is independent of human beings forming them. Nonetheless, it is acknowledged by most legal systems law that corporations have the legal capacity to have rights and obligations, as well as the ability to perform legally meaningful acts in their names. Though they do not exist apart from the collective imagination upheld by the law, corporations do have separate legal status- personhood- that is independent from those of human beings that form them.

The corporate person is held responsible for the activities (acts and omissions) that are performed on its behalf. Though in private law domains holding the corporate person liable often requires accessing its assets, holding that person criminally liable is often seen as problematic.

---

[97] Gideon Yaffe, 'The Voluntary Act Requirement', in Andrei Marmor (ed), *The Routledge Companion to Philosophy of Law* (Routledge 2012) 175.

[98] Gerber summarises the prevalent disdain towards strict liability offences as follows:
The imposition of strict or absolute liability on criminal defendants is condemned by an impressive consensus of jurists and deemed unjust by nearly all. The primary locus of discussion of strict liability settles among disputants in the perennial controversy concerning the justification of legal punishment: both utilitarians and retributivists have argued against strict liability, the former on the ground that it cannot be expected to deter or prevent commission of further offences, the latter on the ground that it is morally objectionable. David Gerber, 'Strict Criminal Liability and Justice' (1974) 60 (4) Archives for Philosophy of Law and Social Philosophy 513, 513.

So much so that, German law still does not recognise corporations as responsible subjects of criminal law, though Italian and Irish legal systems have put in place domestic legislation that regulate the imposition of criminal liability of corporations.[99]

Corporations, since they have no mind on their own, cannot satisfy the *mens rea* element or voluntary act requirement of criminal liability on their own. Generally speaking, corporate liability is based on the link between the corporations and the acts of their agents. The said link is often postulated on the principle of *respondeat superior*. The principle of *respondeat superior* provides for the imputation of agent's mental states to corporation as long as 'the agent was acting within the scope of her employment and in furtherance of corporate interests'.[100] In other words, under this principle, the corporation's liability is induced from the individual liability of its agents. The liability of a company, therefore, does not represent an independent wrongdoing, but arises due the unique relationship between corporations and their agents. [101]

Thus, it can be argued that all types of criminal punishment, unlike obligations in private law, require the limitation of the agency of the punished party. Imprisonment limits the agency of the prisoners directly, by restricting their personal freedom and liberty. In respect of incapacitation measures such as the revocation of the driving license of the punished party, the limitation of agency is achieved by depriving the punished party from being able to perform certain activities that they would otherwise have been allowed to engage in.

The present study acknowledges that, in the present day, most robots still have easily identifiable human manufacturers, masters and most robot crimes can be brought within the remit of justifiable criminal punishment through the principle of *respondeat superior*. However, in such cases where mental states of identified legal persons could be imputed to robots, criminal law systems already have tools to punish culpable human actors, and there is no need to impose direct criminal liability on robots by utilising the principle as justification.

---

[99] Albert W Alschuler, 'Two Ways to Think About the Punishment of Corporations' (2009) 46 American Criminal Law Review 1359, 1367; Ashley S Kircher, 'Corporate Criminal Liability Versus Corporate Securities Fraud Liability: Analyzing the Divergence in Standards Of Culpability' (2009) 46 American Criminal Law Review 157.

[100] Eli Lederman, 'Models for Imposing Corporate Criminal Liability: From Adaptation and Imitation Toward Aggregation and the Search for Self-Identity' (2000) 4 Buffalo Criminal Law Review 641, 655.

[101] W Robert Thomas, 'Of Systems and Persons: The Ability and Responsibility of Corporate Law to Improve Criminal Punishment' (2016) 78 Ohio State Law Journal 601, 613.

The principle of *respondeat superior*, on the surface, appears to be useful for filling in the gap of criminal accountability for the independent behaviours of robots (i.e., that the culpable mental states of manufacturers, owners and users of robots can be imputed to their robots), but considering the unpredictability of the robots' behaviours, establishing individual intent of human actors would still remain a challenge. Unlike corporations that have human agent, robots are not always accompanied by human actors.

Moreover, the situation of corporations and robots are very different from one another. Corporations have legal agency but since they have no separate mental or physical existence to exercise that agency, it is established from the principle that corporate agency is exercised by human beings. Robots, on the other hand, have some physical and mental agency and can act independently. What they do not have is any legal agency, or the recognition of the agency by the law. As such, though the principle can help achieve legal recognition of the agency of robots, it does quite relate to the challenges at hand.

## V.4 Intentional Stance

It is submitted that viewing robots as intentional agents can provide the conceptual background for the determination of liability for robot crimes. The concept of intentional agency is derived from Dennett's theory of 'intentional stance'. One of Dennett's three level of abstractions for interpreting and understanding behaviours of any given entity, the intentional stance is adopted when the entity's behaviours can be most accurately predicted by treating it as some rational agent whose behaviours are governed by its intentional states.[102]

The term *intentional* refers to the assumption that the entity in question will always act in accordance with its objectives and in order to achieve the goals determined these objectives. According to Dennett, '[a]nything that is usefully and voluminously predictable from the intentional stance is, by definition, an intentional system.'[103] Human beings or some other animals meet Dennett's criteria and therefore are international agents.

[102] Dennett C Dennett, *The Intentional Stance* (MIT 2006).
[103] Dennett C Dennett, "Intentional Systems Theory' in ABeckermann , P McLaughlin , and S Walter S (eds), *The Oxford Handbook of Philosophy of Mind* (OUP 2009) 339.

Similarly, it is submitted by List and Pettit that corporations, along other organised collectives are 'usefully and voluminously predictable from the intentional stance.'[104] The present study agrees with the submission of List and Pettit: Though corporations or other human associations do not have their own decision-making and implementation abilities separate from that of human beings who are acting on their behalf, they do have a set of interests t definitive  from the human agents, and human agents' actions on behalf of the corporations are determined by the objective of satisfying these corporate interests. Consequently, the activities that are performed on behalf of corporate agencies can be predicted on the basis of such separate sets of intentions – that are distinct from the interests of corporations' human agents.

At this point, it is necessary to draw attention to some similarities between the characteristics of said group agents and robots. Group agents and robots both have internal architectures, with the former having social architectures and the latter having computational architectures. [105] Consequently, both the group agents and robot systems both have a tendency to gather and process inputs from a variety of sources. They each must meet specific rationality standards at the level of the entity as a whole in order to perform their purposes adequately. They must each alter or update their views in response to the changes in their environments.[106]

The difference here is that different group agents have had economic and social power and importance for a long time. Indeed, states can turn the lives of millions upside down overnight through wars and laws, or corporations can have a significant impact on human beings' well-being as a result of the decisions they make, or even create huge disasters. This means that the issue of legal liability for group agents has already been addressed. Indeed, it is submitted that that the determination of one of more individuals to bear the responsibility in the instance of corporate damages is challenging, at least in some cases, since it is possible that the entirety of individual responsibility may fall short of adequately covering the full extent of the harm inflicted. In other words, if the individuals are always the only ones to blame, there is a deficiency in the amount of responsibility that is given to them.

[104] Christian List and Philip Pettit, Group Agency (OUP 2011) 185.
[105] Christian List, 'Group Agency and Artificial Intelligence' (2021) 34 Philosophy & Technology 1213, 1215.
[106] ibid 1222.

Philip Pettit gives the example of the Herald of Free Enterprise, a passenger ferry which sank in the English Channel in 1987- killing nearly two hundred people. According to the findings of the investigation about the accident, the ferry firm was extremely sloppy and had poor safety protocols in place. Nonetheless, despite the lengthy lawsuits that followed, no individual was found legally liable to a degree that seemed intuitively appropriate. [107]

However, corporate harms, as in the case in the example, have a source beyond the individuals themselves. These harms stem from the actions of a group agent. For this reason, it is argued that corporate entities are 'loci of responsibility themselves, over and above their members'. [108] In the example of the ferry firm, the responsibility gap might have been avoided by holding the company itself responsible for its poor safety protocols. Similarly, it can be argued that robots above a certain threshold of autonomy constitute new source of agency, distinct from the agency of any human makers and users, since, by definition, they are expected to function with little or no human input.

### V.4.1 Robotic Mens Rea

Both of the two requirements for the imposition of criminal liability (the requirement for some conscious and uncoerced behaviour, and the *mens rea* element) can be interpreted for robots in a manner analogous with how they are defined concerning natural persons. The prevailing theory, based on the culpability of subjects holds that one is criminally culpable for the unlawful conduct action to the degree that the conduct exposes their insufficient regard for legally protected interests: These protected interests values provide legally recognised values bearing on how one should behave. Insufficient regard to these legally protected interests either stems from ill will or indifference that produces mistakes in the way one recognises, weighs, and responds to the applicable legal reasons for action. [109]

Criminal law norms are based on the implicit assumption that human beings cannot always be expected guide their behaviours in the society on the basis of the respect they have for one another.[110]

[107] Philip Pettit, 'Responsibility Incorporated' (2007) 117 Ethics 171.
[108] List (n 97) 1223.
[109] Abbott and Sarch (n 42) 355.
[110] Peter Westen, 'An Attitudinal Theory of Excuse' (2006) 25 Law and Philosophy 289, 375.

In fact, the rules of criminal laws are designed to discourage or deter their subjects from demonstrating disrespect in the society. That goal is achieved through the broad range of criminal sanctions that are imposed in response to the behaviours that may negatively impact certain interests and values.[111]

That formulation of the notion of culpability accounts for corporate criminal liability. By means of the hierarchies of employees they consist of, corporations have information-gathering, reasoning and decision-making skills.[112]

Then, through their employees, corporations may perform behaviours that expose their insufficient regard for the legally protected interests, and therefore be regarded as criminally culpable. For example, suppose a corporation learns (through its employees) that its manufacturing activities generate toxic emissions that are fatally harmful for the children and elderly population of a neighbouring town. Here, the corporation has a legally recognized reason to alter its conduct. If the corporation continues its harmful operations, its failure (through its employees) to attach sufficient weight to the legally recognized reason would be revealed and the corporation would potentially qualify as criminally culpable.

Similarly, through their sensors, processors, and actuators, robots may gather and process information, and determine the most efficient way to accomplish their pre-set objectives. As such, it is submitted that the law ought to acknowledge that some robots must have sufficient reasoning and decision-making abilities to display insufficient regard for legal norms – just as it recognizes the decision-making instruments of corporations as such. For example, suppose a robot is designed with the ability to take account of the interests of human beings and follow legal requirements. If the ends up performing behaviours that are inconsistent with taking account of these legally protected interests and values, the impacts of its behaviours would not be dissimilar to that in corporate criminal liability offences.

To determine when robots may be deemed to possess some functional equivalent of human beings' *mens rea,* Michael Bratman's work in philosophy of action that characterizes the functional roles of intentions can be extended to robots.

---

[111] Victor Tadros, *Criminal Responsibility* (OUP 2010) 250.
[112] List and Pettit (n 85) 160-65.

According to Bratman, actors who act (with the purpose) to bring about an outcome 'guide [their] conduct in the direction of causing'[113] that outcome. In other words, 'in the normal case [of an intentional agent], one [who intends an outcome] is prepared to make adjustments in what one is doing in response to indications of one's success or failure in promoting'[114] that outcome. For example, suppose a natural person is driving, intending to hit and slightly injure a pedestrian. Should that person discern that conditions have changed, and behavioural adjustments are needed to make this outcome more probable, he will be predisposed toward making the necessary adjustments.

Equally compelling is the application of Bratman's conception of intention to robots. Suppose a robot is monitoring its surroundings to identify new courses of action that would increase the probability of the outcome (harm an elderly tourist), and it is then more willing to adopt these new courses of action to make the outcome more likely. Regardless of whether the outcome is the end goal, the robot can be deemed to have the purpose of causing that outcome, or mens rea. The same strategy may be employed to argue that a robot possessed some other mens rea, such as recklessness.

### V.4.2 Robotic Criminal Punishment

Under the definition provided by Hart, criminal punishment 'must involve pain or other consequences normally considered unpleasant'.[115] Even if robots are convicted of crimes and subjected to negative treatments such as being reprogrammed or terminated, such treatments cannot be considered as punishment according to , since robots cannot experience 'pain' or 'unpleasantness'. In response, it can be argued that Hart's definition of punishment only requires the treatment in question should be normally considered unpleasant – not necessarily unpleasant to the offender. Hart's definition includes punishments imposed on those who do not experience their sentence as unpleasant for any reason.

Indeed, because the convicted party overtly wants to be imprisoned, for example to further their political agenda, it does not mean that imprisonment after conviction ceases to be punishment.[116]

---

[113] Michael Bratman, 'What is Intention?', in Philip R Cohen, Jerry Morgan, and Martha E Pollack (eds), *Intentions in Communication* (MIT 2012) 157.
[114] ibid 158.
[115] Hart (n 50) 5.
[116] Abbott and Sarch (n 42) 364.

It is submitted that defendants have objective interests set back by their punishments, even when they do not all form negative personal experiences about these setbacks.[117] It can be contended that it is objectively bad for any human being to be deprived of their autonomy liberty, regardless of whether they perceive such deprivations as painful. Should that be the case, Hart's definition would then require punishment to objectively set back defendant's interests, since negative subjective experiences are only one way of doing so.

When it comes to whether robots can have objective interests, there are two views in the scholarly literature. Some authors maintain that only living organisms who are at least capable of phenomenal experiences such as pleasure and pain can be said to have normatively important interests. For instance, Joel Feinberg takes the capacity for cognition as the prerequisite for experiencing things be good or bad.[118] Interests, he says, 'are compounded out of desires and aims, both of which presuppose something like belief, or cognitive awareness.'[119] On this view, since robots are incapable of cognitive awareness, they cannot possess normatively important interests.

In contrast, Philippa Foot contends that situations can be assessed as good or bad for an entity as long as the entity has identifiable functions. [120] For example, developments such as nourishment and reproduction are good for living organisms as these events advance the organisms' biological functions; and events like physical injury or death are bad since such events hinder the organisms' physiological functions. If living organisms can be said to have interests in survival and reproduction, ultimately following their biological programming; then it seems that robots who follow their digital programming can also be assumed to have interests in this sense. After all, robots also have a wide range of identifiable functions and need to have characteristics patterns of behaviours to stay in good working order.

The present study follows the understanding of interests suggested by Jhering who suggested that from the legal perspective, interests must be understood as objectives, needs, personal goals, desires, and wants that motivate the behaviours of any given entity.

---

[117] Guy Fletcher, 'A Fresh Start for the Objective-List Theory of Well-Being' (2013) 25 Utilitas 206, 206.
[118] Joel Feinberg, 'The Rights of Animals and Unborn Generations' in William T Blackstone (ed), *Philosophy and Environmental Crisis* (University of Georgia P 1974) 51.
[119] ibid 52.
[120] Philippa Foot, *Natural Goodness* (Clarendon 2010) 26.

Using the term *agency* in the legal sense, 'agent' refers to entities that are acting in the pursuit of interests of another entity often in exchange for some positive developments regarding their own interests.

Under the existing law, only entities that can be recognised as agents are legal persons. Considering that legal personhood denotes that entities in question have the capacity to have legally protected interests (rights) along with the capability of being subjected to legally enforceable tasks (obligations) on their own behalf, robots cannot be regarded as legal agents- since they cannot have interests of their own. That does not change the fact, however, that we have new entities that, by design, cannot have any interest that do not stem from their specified functions. Robots can act autonomously on other entities' interests, but current levels of technological development do not yet allow the creation of decision-making capacities that would allow them to operate without having any specified functions.

Robots may be programmed to preserve their functionality for a long period of time, but that does not mean they have the drive to preserve their functionality themselves – they have that capability because they are designed in that way, and they are designed in that way to serve another entity's interests more effectively. In short, to state that robots that have the capability to self-preserve themselves have interests would not be conceptually wrong, but it would be legally irrelevant.

If the imposition of criminal punishment must include negative outcomes for the punished parties' interests, then it must be admitted that robots cannot be criminally punished. For the last few decades, especially in the countries that are members of civil-law group, retribution is being replaced by more consequentialist approaches to criminal punishment. For that reason, the evolutionary direction of criminal law is expected to allow for the relaxation of that element for robot crimes, especially as anthropomorphism would create the illusion of retribution being achieved.

## Summary

In the previous chapters of the present study, it was established that robots are expected to undertake increasingly significant roles in the ordinary course of everyday life, facilitated by their continuously developing capabilities that are associated with the characteristics of relative autonomy and social agency. As robots operate more and more independently of continuous human control in everyday life, it becomes more probable that they may autonomously engage in negative behaviours where it may not always be possible to identify any legal person to whose conduct robots' behaviours can be reduced. Some of these negative behaviours are bound to have outcomes that would have attracted criminal sanctions if they were committed by human beings or non-human juridical persons.

Nonetheless, under current criminal laws, robots are not regarded as responsible subjects of criminal law and have no status beyond that of property, as such, it is submitted that the criminal law systems are facing accountability and responsibility gaps in relation to robots' independent behaviours. However imposing punishments on natural or juridical persons for crimes they did not commit or facilitate would be contrary to core principles of criminal law. chapter also assessed current exceptions to these core principles such as the strict liability criminal offences and corporate criminal liability. However even under these exceptions, the punished party is required to have some form of control over the criminally offensive behaviours.

The company is held criminally liable for what was done on its behalf and with its sanctioning – since companies have independent agency but lack the independent existence to exercise it. Equally, in strict liability crimes, the punished party uses their own agency to undertake legally risky activities. In either case, no-one is punished for actions that they did not facilitate in reasonably foreseeable manner.

Accordingly, because of the relative autonomy of robots, it will not always be fair and consistent with the general principles of criminal law that the manufacturers or masters of the robots are held responsible and punished for robot crimes that they had no control over the commission of. Penalising manufacturers and masters of robots for robot crimes, which robots decide to perform as a result of their independent decision-making procedures, free from human control, will mean imposing criminal liability without establishing objective and subjective elements of the crime first, and this is not legally acceptable.

The suggested imposition of criminal liability on robots is inspired by the criminal liability of another non-human entity, corporations. It has been observed that the theoretical background of the liability of corporations is that the corporations have an intentional agency that is different from, and exceeds, the sum of the agencies of the individuals working for them. The reason is that the corporations are not just mere collectives of their members but can produce results different from the intentions of their members since they use their own decision-making and implementation systems. This study acknowledges that this potential of independent decision-making is more evident in robots. Thus, in a system where the agency of groups is recognised, and legal consequences are attributed to it, it is structurally possible to recognise the agency of robots as well. That said, the possible punishment of robots brings about a myriad of practical challenges to modern criminal law. It may be expected that some robots, especially those without a physical embodiment, would be almost impossible to effectively punish, apart from switching them off. Even if such practical issues were resolved, punishing robots would still require major changes to criminal law.

# Chapter VI: Challenges of Civil Liability

## Introduction

The personal, social, and economic implications of the distinctive characteristics of social robots are not limited to one legal domain. First, thanks to the specific capabilities associated with the characteristic of social agency, social robots pose threats to the legal values protected by existing legal instruments. As explored earlier in the present study, the use of such robots risks undermining the values that legal systems seek to safeguard even without violating the legal norms themselves.[1] Equally, however, the growing numbers of robots working side-by-side with human beings increase the importance of clarifying the uncertainty of determining the legal liability for the consequences undesirable independent activities of robots.[2] Unlike other functional artefacts, robots can cause damages through behaviours that are not prompted by any external guidance, to the surprise of even their manufacturers thanks to the capabilities associated with relative autonomy. In the last chapter, we focused on implications of the robotic characteristic of relative autonomy in relation to determining liability for the criminally offensive behaviours of robots. In the following chapter, the focus is on the private law liability for remedying undesirable outcomes of robots' behaviours

Uncertainties regarding the private law liability for robot acts are expected to arise in every legal system; and they therefore should be analysed with a view of the diversity of their manifestation in different social, economic, and cultural contexts. The core principles and the demarcations of these principles, even though they are assessed to be close to one another, are not established in the same manner in chosen legal systems of Germany, Italy, and Ireland. The application of existing private law liability instruments to robots' behaviours is assessed to have certain common shortcomings that are attributable to the capabilities associated with both of the characteristic of relative autonomy.

[1] For example, the anthropomorphisation of robots may desensitise society to violence and sexual misconduct, or cause people to become anti-social. It also can undercut long-standing privacy values, without violating any existing privacy laws, through the emergence of new types of personal data. See Chapter IV, Section IV.2.2, 'Diminution of Privacy and Solitude'.

[2] Trevor N White and Seth D Baum, 'Liability for Present and Future Robotics Technology' in Patrick Lin, Ryan Jenkins and Keith Abney (eds) *Robot Ethics 2.0: From Autonomous Cars to Artificial Intelligence* (OUP 2017) 66.

These shortcomings are especially significant in relation to social robots in view of their capacity to enter into profound and personal relationships with human beings.[3] The questions regarding the private law liability for social robots refers to accountability for these robots' behaviours and seeks to establish who will pay for any possible damages. That said, it is not the only kind of liability that can arise from the behaviours of robots. As explored in the first chapter of the present study, it may be possible for social robots to autonomously perform behaviours that are regarded as criminal offences if performed by humans, such as bodily injury, sexual assault or rape, some types of invasion of privacy and so on. Then, it must be noted that private law and criminal law, fields of law that address liability for robots are based on different principles and have separate structures. As such, this chapter only deals with liability arising from private law for social robots. In order to reveal what is expressed by compensation law, it is deemed fitting to use the public law-private law distinction, which is characteristic of Continental or Civil law systems.

## VI.1 Contract Law

For any free-market economy that is based on the notion of voluntary exchange, as in most economies of the Western world, the most important social interactions are promises. Roscoe Pound expressed the role of promises for capitalist economies decades ago as follows:

> In a commercial age wealth is now largely made up of promises (…) important part of one's substance is made up of advantages which others have promised to provide or render him.[4]

The healthy operation of liberal economic systems depends on the collective presupposition that promises will be fulfilled one way or another.[5] As it happens, when promisors fail to honour their promises, contract law systems provide for the enforcement of promises, making the legal instruments provided by these systems essential legal tools for any properly functioning competitive free-market economy.[6]

[3] Astrid Rosenthal-von der Pütten and others, 'An Experimental Study on Emotional Reactions Towards a Robot' (2012) 5 International Journal of Social Robotics 17, 19; Glenda Shaw-Garlock, 'Gendered by Design: Gender Codes in Social Robotics' in Marco Nørskov (ed), *Social Robots: Boundaries, Potential, Challenges* (Taylor & Francis 2018) 218.
[4] Roscoe Pound, 'Individual Interests of Substance: Promised Advantages' (1945) 59 Harvard Law Review 1, 2.
[5] Larry M Preston, 'Freedom, Markets, and Voluntary Exchange' (1984) 78 American Political Science Review 959, 960.
[6] Stephen Michael Waddams, 'The Modern Role of Contract Law' (1983) 8 Canadian Business Law Journal 2, 4.

Contracting parties do not always negotiate and conclude contracts themselves; they can empower others to create rights and obligations on their behalf, to act as their agents.

Some entities, such as human associations and other juridical persons, can only enter into contracts through their agents as they have no separate physical or mental presence, in fact, no existence apart from the legal fiction of juridical personhood.[7] Since contractual agents' promises are binding on legal persons that granted them the representative authority, the contract law systems provide safeguards for both contracting parties, the principal of the army and counterparty, to protect their interests -at least partially- from undesirable outcomes arising from the conduct of contractual agents who exceed their authority.[8]

In the present time, one of the roles that robots can fulfil in the ordinary course of everyday life is that of agents in contract law, and the inadequacies of existing contract law instruments are related to the consequences associated with the characteristic of relative autonomy in such roles. To elaborate, some robots, often deployed in online environments, can independently negotiate terms and conditions of contracts, and create binding rights and obligations on their masters.[9] Contracts concluded by such robots are shaped by their independent decision-making abilities, meaning that their masters cannot always know, let alone consent, to the contents of such contracts. However, no contract law system recognises either the relative autonomy or the contractual agency of robots. Instead, such robots are viewed as instruments that provide for communication of their masters' intentions.[10] For the purposes of the present study, the term *contract* is understood as the exchange of promises between two or more competent parties that are legally enforceable, therefore bearing impacts on parties' respective rights and obligations.[11] Contracts are regarded to enable the expression of relative autonomy and therefore are considered as the necessary reflections of the freedom of choice.[12]

---

[7] Boudewijn Bouckaert, 'Corporate Personality: Myth, Fiction Or Reality?' (1991) 25 Israel Law Review 156, 156-157.
[8] Gabriel Rauterberg, 'The Essential Roles of Agency Law' (2020) 118 Michigan Law Review 609, 650.
[9] Examples include *smart contracts*, 'pieces of software which regulate the exchange of assets and services between participants'.
See Massimo Bartoletti, 'Smart Contracts' (2020) 3 Frontiers in Blockchain 1, 1; Tilen Čuk and Arnaud van Waeyenberge, 'European Legal Framework for Algorithmic and High Frequency Trading (Mifid 2 And MAR)' (2018) 9 European Journal of Risk Regulation 146, 146-147.
[10] This approach is only accurate for cases where circumstances for robots to conclude contracts and contents of these contracts are pre-determined by their masters.
[11] Ewan McKendrick, 'Contract: in General', in Andrew Burrows (ed), *English Private Law* (3rd edn, OUP 2013) para 8.01.
[12] Jacob Turner, *Robot Rules* (Kindle edn, Palgrave Macmillan 2018) ch 2, s 2.5, sub-s 2.5.2.

The following section first establishes the recognition of the basic principles of contract law in legal systems that are chosen for comparison in the present study and then moves on to discuss their application to contracts that relate to robots.

Basic principles of contract law are close to one another in all three national legal systems examined in the present study. There are three basic principles that are common to all three selected jurisdictions:

(1) Freedom of contract,
(2) Privity of contract, and
(3) Sanctity of contract.

Moreover, despite the differences in how selected common law legal systems (Ireland) and civil law legal systems (Germany, Italy) provide for these basic principles, it is also evaluated that these principles are established with exceptions similar to one another. The examination of how these principles are reflected in legal systems are relevant to robotics when it comes to the determination of who is responsible for the undesirable consequences of robots -who can negotiate and enter into contracts by their own independent decision-making capacities- exceed or violate the authority given to them.

### VI.1.1 Core Principles of Contract Law

#### VI.1.1.a Freedom of Contract

The freedom of contract is perhaps the most fundamental principle of contract law, since it equips legal persons with the power to choose with whom to contract, whether to contract or not, and on which terms to contract.[13] Whilst the freedom of contract is protected by law in all selected legal systems, the relevant legal instruments through which these protections are provided differ, as well as the exceptions to the principle.

[13] Roscoe Pound, 'Liberty of Contract' (1909) 18 Yale Law Journal 454, 457.

In Germany, the freedom of contract is understood to be protected within the framework of Art 2(1) of the Basic Law of the Federal Republic of Germany:[14]

> Every person shall have the right to free development of his personality insofar as he does not violate the rights of others or offend against the constitutional order or the moral law.

GG Art 2 (1) does not explicitly refer to the freedom of contract and merely provides for the right of free development of personality. For that reason, the constitutional protection afforded to of the freedom of contract is inferred from case-law. [15] In 1957, the German Federal Constitutional Court (BVerfG) interpreted the right to free development of personality under GG Art 2(1) as an all arondright to freedom of action, including all manifestations of freedom that are not protected by specific fundamental rights.[16] Though the freedom of contract is not explicitly mentioned in *Elfes* decision, it was possible to infer that the freedom of contract -which is not protected by any other specified fundamental right- is guaranteed under GG Art 2 (1). Thereafter, the BVerfG recognized the substantive freedom of contract in a 1990 decision, [17] but it did not expressly devise this substantive freedom as a constitutional right until the *Bürgschaft* decision of 1993.[18] Though it is recognised as a constitutional right, the freedom of contract in German law is not unbridled. The parties do not have complete freedom as to the contents of a contract, nor is the freedom to conclude or not to conclude a contract unlimited.[19] These limitations are directed at preventing violations of the rights of others, offences against the constitutional order, and transgressions of the moral law in GG Art 2 (1).

Of course, the nature of these limitations has been clarified by legislation and case-law. The limitations to the freedom of contract included in the German Civil Code [BGB][20] are the breach of statutory prohibitions, the violation of public policy, and the creation of a grossly inadequate and unduly obtained bargain.

---

[14] Grundgesetz für die Bundesrepublik Deutschland (GG) [English translation: Basic Law for the Federal Republic of Germany (GG).

[15] Aurelia Colombi Ciacchi, 'Party Autonomy As a Fundamental Right in the European Union' (2010) 6 European Review of Contract Law 303, 305; Manfred Pieck, 'A Study of the Significant Aspects of German Contract Law' (1996) 3 Annual Survey of International & Comparative Law 111, 112.

[16] 16.01.1957 BVerfGE 6, 32 - 1BvR 253/56 *Elfes* decision

[17] 07.02.1990 BVerfGE 81, 242 - 1 BvR 26/84 *Handelsvertreter* decision, see Ciacchi (n 15) 306.

[18] 19.10.1993 BVerfGE 89, 214 - 1 BvR 567/89 *Bürgschaft* decision.

[19] For example, whenever there is a monopoly, the monopoly is legally obligated to enter into contracts concerning whatever it monopolizes, see Pieck (n 15) 112.

[20] Bürgerliches Gesetzbuch [BGB] [Civil Code], August 18, 1896, RGBl at 195, revised January 2, 2002, BGBl I at 42, last amended by Art 4 of the Law of July 15, 2021, BGBl I at 1146

***Breach of Statutory Prohibitions:*** BGB Section 134 establishes that 'A legal transaction that violates a statutory prohibition is void unless the statute leads to a different conclusion'. According to this provision, contracts that violate a statutory prohibition are void, unless it appears from the statute that the prohibition was not intended to nullify contravening contracts. In the absence of explicit language in the statute, it becomes a matter of interpretation whether or not the prohibitions were intended to nullify contravening contracts.

Accordingly, Pieck argues that 'it is thus possible that a contract which violates a statutory prohibition is valid when the prohibiting statute contains no express language invalidating a contravening contract'.[21] Pieck offers the example of a contract entered into on a Sunday in violation of a statute forbidding transactions on Sunday being held valid as the statute did not spell out that the violating contracts are void.[22] However, it must be emphasised that, if a statute prescribes that a contract requires a government license, a contract entered into without the required license would be void. Likewise, if a contract violates any of the fundamental rights contained in the Basic Law, such contract would also be void.[23]

***Violation of Public Policy:*** BGB Section 138 (1) states that 'a legal transaction which is contrary to public policy is void'. This provision establishes that certain contracts can be invalidated for public policy reasons. According to Youngs, such reasons 'principally protect the interests of the state and society, but they seem sometimes to be protecting fundamental rights of the individual.'[24]

To determine whether or not a contract is immoral as a practical matter requires a reference to the relevant case-law, which has held that 'contrary to public policy' includes sexual immorality.[25] Contracts in restraint of trade and contracts which oppressively restrict a person's independence or economic freedom have also been held to be contrary to public policy and thus void.[26]

---

[21] Pieck (n 15) 112.
[22] ibid 112.
[23] Basil S Markesinis, Hannes Unberath, and Angus Johnston, *The German Law of Contract* (Bloomsbury Publishing 2006) 43.
[24] Raymond Youngs, 'Constitutional Limitations on Freedom of Contract: What Can the Germans Teach Us' (2000) 29 Anglo-American Law Review 498, 503.
[25] OLG Hamm 2 December 1985, NJW 1986, 781.
[26] BGH, 31 March 1982, BHGZ 83, 313 *Solus agreement* decision.

***Undue Advantage:*** In German Law, contracts that are grossly to the disadvantage of one of the parties cannot be enforced and are therefore void. BGB Section 138 (2) provides that, in particular,

> a legal transaction is void by which a person, by exploiting the predicament, inexperience, lack of sound judgement or considerable weakness of will of another, causes himself or a third party, in exchange for an act of performance, to be promised or granted pecuniary advantages which are clearly disproportionate to the performance.

However, in practice, BGB Section 138 (2) is regarded to protect the exploited party, and a transaction may be upheld despite being in breach of this provision.[27] In short, in German law, the freedom of contract is protected by the constitutional legislation, namely, GG Art 2 (2). However, the freedom of contract does not extend to the breach of statutory prohibitions, violations of public policy (including sexual immorality), and contracts that provide and unfair advantage to one of the parties. Contracts that are not considered to be within the freedom of contract are regarded to be unenforceable and therefore void.

### *Freedom of Contract in Italy*

In Italy, just as in Germany, the text of the Constitution of the Italian Republic [Cost] does not refer to the freedom of contract. However, the case-law that the Italian Constitutional Court has been developing since 1962[28] establishes that the freedom of contract is indirectly protected due to its functional relation with Cost Art 41:[29]

> Private-sector economic initiative is freely exercised.
>
> It cannot be conducted in conflict with social usefulness or in such a manner that could damage safety, liberty and human dignity.

Cost 41 is not placed under in the section that addresses civil rights and duties but under the title 'Economic Relations'. Thus, it seems that 'private sector economic initiative' and, indirectly, the freedom of contract are provided somehow lesser degrees of protection than personal rights identified by Cost. Affirmingly, Cost Art 41 itself indicates that the exercise of that freedom may be restricted on the grounds of 'conflict with social usefulness'.

---

[27] BGHZ 89, 316.

[28] Corte costituzionale, decision n. 7 of 20 February 1962; Corte costituzionale, decision n. 30 of 8 April 1965; Corte costituzionale, decision n. 37 of 13 March 1969; Corte costituzionale, decision n. 241 of 15 May 1990; Corte costituzionale, decision n. 268 of 30 June 1994; Corte costituzionale, decision n. 70 of 17 March 2000.

[29] Costituzione della Repubblica Italiana [English translation: 'The Constitution of the Italian Republic' (*Documents, The Italian Constitutional Court*, 2012)

Ciacchi notes that although the Italian Constitutional Court's approach is widely accepted by Italian legal literature, some scholars still categorise the freedom of contract as part of the self-determination rights of individuals and define it as one 'manifestation of the freedom to decide on one's own personal and patrimonial legal sphere.'[30] These researchers, according to Ciacchi, rely on Cost Art 2, which is also the legal basis of personality rights in Italy:[31]

> The Republic recognises and guarantees the inviolable rights of the person, both as an individual and in the social groups where human personality is expressed. The Republic expects that the fundamental duties of political, economic and social solidarity be fulfilled.

Though the wording in that provision indicates the presence a catch-all fundamental right, such a constitutional right to general freedom of action has not been recognized in Italian case-law. Moreover, the Italian Constitutional Court explicitly declined to consider the freedom of contract under aegis of Cost Art 2.[32] Nonetheless, the general freedom of contract is fully recognized in Art 1322 of the Italian Civil Code [CC]:[33]

> The parties can freely determine the content of the contract within the limits imposed by the law. The parties may also conclude contracts that are not regulated, provided they are aimed at achieving interests worthy of protection according to the legal order.

In Italian law, similarly to German law, freedom of contract is not unqualified. CC Art 1322, right after the sentence that establishes the freedom of contract, states that this freedom is subject to the 'limits imposed by the law'. Explaining these limits, CC Art 1343 stipulates that a contract is unlawful if it is contrary to 'mandatory rules, to the public order, or good morals'.[34]

***Mandatory Rules:*** Mandatory rules are statutory rights and obligations that cannot be derogated from by contracting parties. Unlike default rules of contract law that confer alienable rights, mandatory rules confer inalienable (immutable) rights on the concerned parties.[35] This limitation can also be understood as 'the breach of imperative statutory provisions'.[36]

---

[30] Ciacchi (n 15) 314.
[31] ibid.
[32] Corte costituzionale, decision n. 16 of 21 March 1968.
[33] Codice Civile [CC] [Civil Code], Regio Decreto 16 marzo 1942, n 262, GU 4 aprile 1942, n 79
[34] CC Art 1344 further provides that a contract is also unlawful when 'it constitutes a means for evading the application of a mandatory rule'. Moreover, CC Art 1345 establishes that a contract is unlawful 'when the parties are led to conclude it solely by an unlawful motive, common to both'.
[35] Stephen J. Ware, 'Default Rules from Mandatory Rules: Privatizing Law Through Arbitration' (1999) 83 Minnesota Law Review 703, 706.
[36] Ole Lando, *Principles of European Contract Law - Part III* (Kluwer 2003) 218.

***Public Order/Good Morals:*** 'Public order' refers to the external, collective order of the society; as well as the body of principles that underpin the operation of the legal order. Accordingly, this limitation addresses the social, moral, and economic values that hold a society together and applies to contracts that would offend some overriding public interest. Good morals include not only the principles of sexual immorality, 'but rather the collection of social and moral principles that form [….] a society and are respected by majority of the participants in that society.'[37]

***Undue Advantage:*** Under CC Art 1448, when there is a disproportion between the performance of one party and that of the other, and the disproportion has depended on the state of need of one party which the other has taken advantage of, the injured party can request termination of the contract. [38] Contracts that are established by taking advantage of ignorance or necessity of the other party, can be terminated – but only by the party who is taken advantage of. The provision regarding undue advantage appears to be similar in effect in both German and Italian laws.

It should be noted that, under CC Art 1322 (2), the parties' freedom to create 'new' contracts is recognized insofar as they pursue an interest 'worthy of protection'. Interestingly, Antoniolli and Veneziano note that this provision is interpreted as prohibiting judges from invalidating contracts merely because they are not regulated in CC, and it also serves the purpose of 'recognizing the validity of contractual agreement deriving from the international practice'.[39]

In Italian law, the Constitution does not directly refer to the freedom of contract. However, the well-developed case law on the topic points out that it is implicitly protected thanks to its functional relation Cost Art 41, which addresses the private sector-public sector cooperation and the freedom of economic enterprise. As is the case in German law, in Italian law, the freedom of contract does not extend to the violation of statutory prohibits (expressed as 'mandatory rules' in the context of Italian law), and the violation of public policy (expressed as 'public policy' and 'good morals' in Italian law').

[37] Chantal Mak, *Fundamental Rights in European Contract Law* (Kluwer 2008) 32.
[38] Art 1448 CC.
[39] Luisa Antoniolli and Anna Veneziano, *Principles of European Contract Law and Italian Law* (Kluwer 2005) 32.

In Ireland, contract law forms part of the common law system. In other words, it is based on custom and judicial precedent rather than statutes. Nonetheless, it can be assumed that the Irish Constitution -albeit implicitly - protects the freedom of contract. This assumption is based on Ciacchi's proposition that suggests the constitutional dimension of freedom of contract can be acknowledged in all jurisdictions 'where the principles of self-determination, free development of personality and/or general freedom of the person are enshrined in the Constitution.'[40] In the Irish Constitution, the general personal rights are guaranteed in Art 40.3.1 which provides that 'the State guarantees in its laws to respect, and, as far as practicable, by its laws to defend and vindicate the personal rights of the citizen.' Indeed, the term 'personal rights', as interpreted by the courts, has led to the recognition and vindication of many rights not expressly provided for in the text of the Constitution. These 'unenumerated rights' include the right to bodily integrity, the right to marry and the right to earn a living, as well as the freedom of contract, among others. It is questionable, however, whether the Irish Constitution, guarantees the freedom of contract to the extent that it is protected in some other common law jurisdictions; for example, in the United States, the US Constitution limits the possibility of frustration of contract through legislative change.[41] Since the Constitution bars the effective voidance of contractual obligations through statutory instruments, it can be said that party autonomy is protected as a fundamental right in the US law.

Still, as is the case in other common law countries, freedom of contract is an essential doctrine that underpins the contract law of Ireland. In this context, it is necessary to consider the evolution of freedom of contract throughout the history of common law. The period between c. 1770 and c. 1870, called the 'Classical Period' of the Common Law by Patrick Atiyah, in his seminal book, was the pinnacle of the theories of natural law and the philosophy of free enterprise.[42]

---

[40] Aurelia Colombi Ciacchi, 'Freedom of Contract as Freedom From Unconscionable Contracts' in Mel Kenny, James Devenney and Lorna Fox O'Mahony (eds), *Unconscionability in European Private Financial Transactions* (CUP 2010) 24.

[41] US Constitution, Article 1, Section 10, Clause I expresses that 'no State shall […] pass any […] law impairing the obligation of contracts'.

[42] Patrick S Atiyah and Stephen A Smith, *Atiyah's Introduction to the Law of Contract* (6th edn, Clarendon 2005) 36.

To the judges of that period, natural law meant that 'individuals had inalienable rights to own property, and therefore to make their own arrangements to deal with that property, and hence to make contracts for themselves.'[43] In other words, the philosophy of *laissez-faire* was understood to mean that the State and its courts should refrain from interfering with transactions between persons as much as possible. The importance given to freedom of contract during this period was clearly best expressed in the 1874 case of *Printing and Numerical Registering Company v Sampson*:[44]

> If there is one thing which more than another public policy requires, it is that men of full age and competent understanding shall have the utmost liberty of contracting, and that their contracts, when entered into freely and voluntarily, shall be held sacred and shall be enforced by Courts of justice.

This libertarian approach to freedom of contract assumed that all contracting parties were negotiating on an equal footing. Nonetheless, with economies becoming increasingly complex and inequality of bargaining powers coming to be the rule rather than the exception, it became apparent that this approach was not workable and could serve to legitimate of the exploitation of the weaker sections of the society. The 1905 case of *Lochner v New York* illustrates one such instance. The dispute in that case arose from a statute enacted by the State of New York that restricted the number of hours that bakers could work, based on evidence that working for over sixty hours per week may be harmful to the health of the bakers.[45] Lochner was prosecuted twice for violating this statute, and on his conviction, he appealed. In its decision, the US Supreme Court struck down the statute and held that the restriction brought by the State of New York was not 'a legitimate exercise of the police power of the State', but 'an unreasonable, unnecessary and arbitrary interference with the right and liberty of the individual to contract in relation to labour.'[46] The decision in *Lochner v New York* was overturned in 1937.[47]

According to former Justice John Paul Stevens, today 'there is virtually universal agreement among judges and scholars that it was incorrectly decided'.[48] Accordingly, it can be argued that one of the important restrictions imposed on freedom of contract is related to unfair contract terms and unequal bargaining powers.

---

[43] ibid.
[44] *Printing and Numerical Registering Co v Sampson* (1874-75) LR 19 Eq 462, 465.
[45] *Lochner v New York,* 198 U.S. 45, 69 (1905) (Harlan, J., dissenting).
[46] *Lochner v New York,* 198 U.S. 45, 56 (1905) (Peckham, J., plurality opinion).
[47] *West Coast Hotel Co. v Parrish*, 300 U.S. 379, 391 (1937) (Hughes, C.J., plurality opinion).
[48] John Paul Stevens, *Five Chiefs* (Little, Brown and Company 2011) 25.

***Unfair Contract Terms:*** In the reality of everyday life, most parties do not have the bargaining power required to choose the terms they like best. Among others, tenants, employees, and smaller businesses often have to accept the terms dictated to them by their stronger counter-parties. These terms 'come in the form of general conditions that the other party is indifferent about and simply accepts without bothering to read them.'[49] Accordingly, in Ireland, like elsewhere, legislators and courts tend to intervene to protect the interests of the weaker party.

In Ireland, the intervention to protect the weaker party is not carried out under a catch-all provision; instead, contracts like employment, residential lease, consumer credit and consumer sale are specifically governed by vast number of mandatory rules that set boundaries for stronger parties.[50] For example, in employment contracts which are characterised by the more advantageous bargaining position of employers, the legislation lays down detailed rules aiming to ensure employees are afforded minimum wage protections, unfair dismissal rights, paid sick leave, safety at the workplace.[51]

Here, the EU Directive on Unfair Terms in Consumer Contracts (93/13/EEC)[52] and the Unfair Terms Regulations[53] that transposed the Directive into Irish national law are relevant. The Directive protects consumers against grossly one-sided standard contracts. Accordingly, if a consumer contract contains an unfair term, that term will not be binding on the consumer and the rest of the contract will continue to be valid, as long as it continues to make sense.

**Public Policy:** In common law jurisdictions, contracts contrary to public policy cannot be enforced. Still, 'public policy' is regarded as an elusive term and the understanding of it continually changes, and therefore it is not possible to precisely determine which contracts oppose the public policy beforehand.[54] It is noted that apart judges have closely followed the principles laid down in the English courts on the subject of public policy and therefore a separate doctrine of Irish public policy has not evolved.

---

[49] Jan M. Smits, *Contract Law: A Comparative Introduction* (Edward Elgar 2014) 10.
[50] ibid.
[51] e.g., Organisation of Working Time Act 1997 regulates a number of employment conditions including maximum working hours, night work, annual leave and public holiday leave. National Minimum Wage Act 2000 provides for an enforceable national minimum wage.
[52] Council Directive 93/13/EEC of 5 April 1993 on Unfair Terms in Consumer Contracts [1993] OJ L95/29.
[53] European Communities (Unfair Terms in Consumer Contracts) Regulations 1995 SI 27/1995.
[54] Farshad Ghodoosi, 'The Concept of Public Policy In Law: Revisiting the Role of the Public Policy Doctrine in The Enforcement Of Private Legal Arrangements' (2016) 94 Nebraska Law Review 685, 687: William Binchy, 'Tort Law In Ireland: A Half-Century Review' (2016) 56 Irish Jurist 199, 206.

There are three kinds of contracts that may contradict public policy:[55] contracts which subvert the status of marriage, agreements in restraint of trade,[56] agreements to oust the jurisdiction of courts.

**Illegal Contracts**: Irish legal system accepts that courts should refuse to enforce contracts that involve illegal acts. Here, it must be noted that 'illegal' does not necessarily mean criminal. Contracts that involve criminal acts -e.g., agreements to sell illicit drugs- are of course illegal; however, some contracts can be regarded as illegal even if the acts they contemplate are not criminal, for example, agreements to pay for sex.

In Irish law, both the statutory law and common law identify illegal contracts. Illegal contracts at common law include contracts to commit crimes or torts,[57] contracts prejudicial to the administration of justice,[58] agreements which serve to defraud the Revenue or to corrupt public officials,[59] contracts tending to encourage immorality,[60] contracts to trade with enemies of State, and contracts that breach foreign law.[61]

---

[55] Robert Clark, *Contract Law in Ireland* (7th edn, Kindle edn, Round Hall 2013) pt 5, ch 15.

[56] The Irish doctrine of public policy regarding the agreements in restraint of trade is identical with the English doctrine. The governing principle has been affirmed in the 1966 case of *Petrofina (Great Britain) Ltd. v* Martin by the Court of Appeals. The Irish courts adopted that approach, most distinctly in the 1987 case of *John Orr Ltd v Orr* :

> All restraints of trade in the absence of special justifying circumstances are contrary to public policy and are therefore void. A restraint may be justified if it is reasonable in the interests of the contracting parties and in the interests of the public.

*Petrofina (Great Britain) Ltd v Martin* [1966] Ch 146, 180. (Diplock L.J.); *John Orr Ltd. and Vescom B.V. v John Orr* [1986] IEHC 1, [1987] ILRM 702, para 2. (Costello J.)

[57] The principle of *'ex turpi causa non oritur actio*' (no action can be based on a disreputable cause) illustrates this kind of illegal contracts. The principle indicates that parties who have knowingly entered into an illegal contract may not be able to enforce it Bruce MacDougall, 'Ex Turpi Causa: Should A Defence Arise from A Base Cause' (1991) 55 Saskatchewan Law Review 1, 2-3.

[58] Contracts that prevent the cause of justice, while not always of themselves criminal, are illegal and cannot give rise to enforceable obligations. For example, In *Brady v Flood* the defendant paid a sum of money to get criminal charges of conspiracy dropped, but the court refused to hear litigation arising from this transaction as it was regarded as illegal. *Brady v Flood (*1841) 6 Circuit Cases 309.

[59] In the case of *Lord Mayor of Dublin v Hayes*, the defendant was appointed Marshall of the City of Dublin, the position also gave him the post of Registrar. This entitled the defendant to collect some fees which he agreed to transfer to the City Treasurer. The appointment was made in exchange for the defendant's promise. The promise was held unenforceable because such a contract would tend to encourage corrupt practices amongst public officials. *Lord Mayor, Alderman & Burgess of Dublin v Hayes* (1876) 10 IRCL 226; Clark (n 55) pt 5, ch 16, s 2, sub-s 3.

[60] Even if the conduct contemplated by the contract is not itself illegal, contracts that encourage immoral behaviours are illegal. Courts have often interpreted 'immoral behaviours' as illicit sexual behaviours, meaning 'intercourse which takes place outside the confines of marriage'. In the recent past, contracts that were seen as immoral included those related to property rights, made between cohabiting couples, denying individuals the right to enter into enforceable contracts because of sexual orientation and lifestyle or marital status. Nevertheless, this approach is changing rapidly. On a different note, contracts designed to promote unlawful gambling is just as illegal as one aimed at furthering illicit sexual activity.
See Atiyah and Smith (n 42) -227; Clark (n 55) pt 5.

[61] Clark (n 55) pt 5, ch 16, s 2, sub-s 6. The courts often refuse to enforce a contract that is illegal according to the law of the place of performance. In *Stanhope v Hospitals Trust Ltd*. the plaintiff in Natal posted Irish sweepstake tickets to the Dublin office where the draw was to take place. Sweepstakes were illegal in Natal; the contract was illegal according to the law of the place where the contract was formed; it was not illegal under Irish law and could be enforceable in Ireland. *Stanhope v. Hospitals Trust Ltd.* (1936) Ir Jur Rep 25; Clark (n 55) pt 5, ch 16, s 2, sub-s 7.

In addition to illegal contracts at common law, contracts that fall foul of Acts of the Oireachtas or Statutory Instruments might also be invalidated by legislation, either expressly or impliedly. For instance, the now repealed Moneylenders Acts 1900–1989 made contracts entered into by unlicensed moneylenders unenforceable; and charging of compound interest is also unlawful or the Family Home Protection Act 1976 Section 3 (1), provides: '[W]here a spouse, without the prior consent in writing of the other spouse, purports to convey any interest in the family home to any person except the other spouse, then, [subject to legislative exceptions] the purported conveyance shall be void.'

In conclusion, in all three selected legal systems, the freedom to contract is implicitly protected by the jurisdictions' respective constitutions. The protection afforded to the liberty of contract in these jurisdictions' respective constitutions is often inferred by the courts. In none of the selected legal systems is the freedom of contract unbridled. All three jurisdictions appear to subject the liberty of contract to the limitations imposed by the statutory prohibitions, unfair advantage, and public policy violations (flagrant moral breaches).

In the present day, in none of the legal systems are there provisions that prohibit the use of robots for the conclusion of contracts, or conclusion of contracts regarding robots themselves. However, it remains contentious whether the use of robots as agents to negotiate a contract constitutes an unfair advantage since the robots, by definition, have better computational and analytical abilities, and stronger will compared to an ordinary human being – and that may be regarded as 'exploiting the predicament, inexperience, lack of sound judgement or considerable weakness of will of another' during the negotiations with a view to the formation of a contract. Nonetheless, only the German law appears to attribute any consequence to the unfairness during the negotiations while Italian law and the Irish law focus on the unfair terms. When it comes to public policy or good morals prohibitions, there appears to be no specific prohibition addressed to robots. That said, the child pornography laws in all three countries, for instance, appear to prohibit the use of child sex robots.

### VI.1.1.b Privity of Contract

Like the secondary obligations that arise in tort law systems, contracts only take effect between contracting parties and do not bind any third party.[62] The contracting parties are free to assign the benefits, but not the burden, of the contract to a third party.

*Privity of Contract in Germany*

In Germany, the principle of privity of contract is not explicitly included in any legislation. However, BGB Section 328 regulates contracts for the benefit of third parties:

> (1) Performance to a third party may be agreed by contract with the effect that the third party acquires the right to demand the performance directly.
> (2) In the absence of a specific provision it is to be inferred from the circumstances, in particular from the purpose of the contract, whether the third party is to acquire the right, whether the right of the third party is to come into existence immediately or only under certain conditions, and whether the power is to be reserved for the parties to the contract to terminate or alter the right of the third party without his approval.

According to the seminal work of Markesinis, Unberath and Johnston, when the provision was being drafted, the legislator had insurance and annuity contracts in mind but the flexibility of BGB Section 328 (2) allows courts to apply this provision to new situations.[63] Even in such cases, however, the privity of contract appears to still create obligations that can be claimed between parties that are pre-determined and agreed upon.

*Privity of Contract in Italy*

In Italy, the principle of privity of contract is expressed in CC Art 1372 which establishes that 'a contract has no effect with respect to third parties except in cases provided for by law.' The provision emphasises that third parties cannot be bound by an agreement concluded between contracting parties but allows the contract to produce effects on others (legal effects, such as rights) on others only when the law so provides. The Italian law allows the contracting parties to extend the remit of the contract to third parties in specific situations. Similar to German law, contracts for the benefit of third parties are the most obvious exception to the principle of privity of contract in Italian law.[64]

---

[62] Tudor Vlad Rădulescu, 'The Principle of Relative Effect of Contracts: A Historical View and Aspects of Comparative Law', *CKS 2018: Challenges of the Knowledge Society* ("Nicolae Titulescu" University Publishing House 2018) 292.
[63] Basil S Markesinis, Hannes Unberath, and Angus Johnston, *The German Law of Contract* (Bloomsbury Publishing 2006) 187.
[64] Guido Alpa and Vincenzo Zeno-Zencovich, *Italian Private Law* (Routledge-Cavendish 2007) 225.

*Privity of Contract in Ireland*

In Ireland, the privity of contract is a controversial doctrine: Closely related to the requirement of consideration, it incorporates two ideas. First, only a party to a contract can have any burdens from the contract enforced on them. A third party cannot be sued for breach of the contract. This first aspect of the doctrine of privity is generally accepted as 'good law', as it would be not only unfair but also contrary the notion of freedom of contract if two parties could impose contractual obligations on a third party without the latter's consent.[65] The second aspect of the doctrine is the idea that only a party to a contract can enforce the contract or a term of the contract, and a third party cannot sue to enforce the contract, even if the contract was consciously made for their benefit.[66]

Exceptions to the second aspect of the privity rule include agency, collateral contracts, and assignments. Still, the Law Reform Commission found that 'it is clear that the current exceptions do not, and will not, cover every situation where an unjust or illogical result is caused by the privity rule'[67] and published a report recommending reform in this area with the introduction of the Contract Law (Privity of Contract and Third-Party Rights) Bill 2008. The Bill stipulates that a third party should be entitled to enforce the contract where the parties intended to give the third party a right to enforce; where the contract expressly states that the third party has a right of enforcement; and where the contract permits a third party to rely on exclusions or limitations on liability.[68] Such a bill is yet to be enacted.

### VI.1.1.c Sanctity of Contracts

Equally important is the principle *pacta sunt servanda*, the sanctity of contract. Accordingly, once the contract is formed, the parties are bound by it – they must strictly observe and comply with this contract.[69] The principle indicates that the contract has the force of law between the parties, and therefore symbolises the importance of a promise and the force of the mutual and accordant wills of the parties in contract law.

---

[65] ibid; Ewan McKendrick, *Contract Law* (5th edn, OUP 2012) 952-954; Law Reform Commission, *Privity of Contract and Third Party Rights* (LRC 88-2008) 5.

[66] Law Reform Commission, *Privity of Contract and Third Party Rights* (LRC 88-2008) 6; Cliona Kelly, 'Privity of Contract- The Benefits of Reform' (2008) 1 Irish Judicial Studies Institute Journal 145, 146. In the antiquated case of *Murphy v Brower*, this principle is summarised as 'no stranger to the consideration can take advantage of a contract, although made for his benefit. *Murphy v Brower* (1868) 2 IRCL 506.

[67] Law Reform Commission, *Privity of Contract and Third Party Rights* (LRC 88-2008) 26

[68] ibid 103, Contract Law (Privity of Contract and Third Party Rights) Bill 2008, sec 3.

[69] Hans Wehberg, 'Pacta Sunt Servanda' (1959) 53 The American Journal of International Law 775.

The principle of *pacta sund servanda* embodies the collective expectation that promises made are going to be upheld, and as such, its enforcement by contract law systems their fundamental role in the healthy operation of economic systems.

### *Sanctity of Contracts in Germany*

In Germany, the principle of sanctity of contract is expressed in the BGB Section 242 that states 'an obligor has a duty to perform according to the requirements of good faith, taking the customary practice into consideration.' In German law, the theory of *Wegfall der Geschäftsgrundlage* (disappearance of the foundation of the transaction) addresses the consequences of altered circumstances on the contract. When the circumstances surrounding the performance of an obligation are unforeseeably and extensively altered, the foundations of the transaction that created the aforementioned obligation would disappear.[70] Accordingly, to demand the performance of that obligation would constitute *bad faith*, and therefore the contracting parties can no longer bound to their original contractual commitments.[71]

### *Sanctity of Contracts in Italy*

In Italy, the principle of *pacta sunt servanda* is reiterated in CC Art 1372, which affirms that 'a contract is binding between the parties'. CC Art 1467, however, provides relief when the performance of the contractual obligations has become unreasonably difficult due to extraordinary and unforeseeable events for a contracting party, so long as this party has not assumed the risk for such changes.

In contracts with continuous or periodical execution or adjourned execution and in cases where the obligation of one of the parties has become excessively onerous due to extraordinary and unpredictable events, the party who is obliged to such performance can demand the dissolution of the contract with the effects laid down in CC Art 1458. The dissolution cannot be demanded if the supervening burden is part of the normal risk of the contract. The party against whom the dissolution is demanded can prevent this by offering to modify equitably the conditions of the contract.

---

[70] Peter Hay, 'Frustration and its Solution in German Law' (1961) 10 The American Journal of Comparative Law 345, 361.

[71] Hans van Houttie, 'Changed Circumstances and Pacta Sunt Servanda' in Emmanuel Gaillard (ed) *Transnational Rules in International Commercial Arbitration* (ICC 1993) 113.

In Ireland, just as in other common law jurisdictions, the principle of sanctity of contract complements the principle of freedom of contract. This principle establishes that a contract is 'sacred' and should be enforced according to its terms, as agreed by the parties. In particular, the principle of sanctity of contract stipulates a contract should not be rewritten by a court because it proves to be improvident for one of the parties or is regarded to be unsatisfactory in some way.[72] The most significant limitation to the principle of sanctity of contracts in civil law countries, the doctrine of *rebus sic stantibus*, does not exist in common law jurisdictions, including Ireland.[73] Nevertheless, courts in common law countries have managed to achieve similar results through three concepts; impossibility, impracticability, and frustration.[74] According to Eisenberg, these three concepts underpin the following precepts that govern unexpected circumstances and cases:[75]

> (1) A contract consists not only of the writing in which it is partly embodied, but also includes, among other things, certain kinds of tacit assumptions.
> (2) These assumptions may be either event-centred or magnitude-centred.
> (3) The problems presented by unexpected-circumstances cases should be viewed in significant part through a remedial lens.

### VI.1.2 Shortcomings of Contract Law

In terms of the civil liability for robots, contract law may be relevant in two basic circumstances: 'contracts for robots', where robots are the subjects of a transaction, and 'contracts by robots', where robots have a role in negotiation.[76] 'Contracts for robots' refers to contracts about the purchase, sale, rent, lease, and use of robots. In this context, robots' characteristics are not of any particular significance:[77] the robot is no different to any other potentially hazardous object of exchange, product, or commodity.[78]

---

[72] ibid.
[73] Aziz T Taliba, 'Rebus Sic Stantibus : A Comparative Survey' (2001) 8 Murdoch University Electronic Journal of Law.
[74] ibid.
[75] Melvin A Eisenberg, 'Impossibility, Impracticability, and Frustration' (2009) 1 Journal of Legal Analysis 258.
[76] Andreas Lober, Tim Caesar and Wojtek Ropel, 'Germany Chapter', in Jeremy Bensoussan and Alain Bensoussan (eds), *Comparative Handbook: Robotic Technologies Law* (Kindle edn, Larcier 2016) ch 6, s 6.6.
[77] Christophe Leroux and Roberto Labruto, 'Suggestion For a Green Paper on Legal Issues in Robotics' (Report, euRobotics 2012) 54.
[78] A survey on the legal regulations on robotic technologies of seventeen countries (including China, Portugal, Spain, the UK, and the US) concluded that none of these countries had a separate regulation on contracts for robots, see Alain Bensoussan and Jérémy Bensoussan (eds), *Comparative Handbook: Robotic Technologies Law* (Kindle edn, Larcier 2016).

Accordingly, in the present study, the term *robot contract* is used to refer solely to contracts by robots. These are contracts that are formed through interactions between robots and legal persons or through interactions between robots alone. (Such contracts are concluded by robots on behalf of principals who are recognised as legal persons. In some areas, such contracts are the norm; on Wall Street, '70 % of transactions on the stock market are made by trading robots'.[79] At international level, the most significant regulation regarding such contracts is found in the UN Convention on the Use of Electronic Communications in International Contracts:

> A contract formed by the interaction of an automated message system and a natural person, or by the interaction of automated message systems, shall not be denied validity or enforceability on the sole ground that no natural person reviewed or intervened in each of the individual actions carried out by the automated message systems or the resulting contract.[80]

However, it is not unmistakably clear whether this article can in fact be applied in the context of contracts formed by robots, as it was not drafted with that specific intention. The article merely provides that computer-generated contracts are not to be denied validity because of the lack of review.[81]

Unlike automated message systems that are used to generate contracts, autonomous robots can learn from previous decisions they have made, and make decisions that go beyond the pre-determined parameters, making the application of this article problematic even if it is accepted to impose liability for robot contracts to the recognized legal persons on whose behalf robots have contracted.

There are two roles that robots can play in the establishment of contracts. Evidently, robots can be used as tools that facilitate the conclusion of contracts, in accordance with the strictly pre-determined requirements of the parties. In such cases, their roles are not much different from those of vending machines or ticket machines, and there are no shortcomings in the existing legal instruments in determining liability insofar as they regulate computer-generated contracts. The second possibility refers to cases where contracts are concluded by autonomous robots that go beyond the pre-determined requirements of the parties.

---

[79] Franck Dedieu and Béatrice Mathieu, 'Les Robots Vont-Ils Prendre Nos Jobs?' (*L'Expansion.fr*, 3 June 2014).
[80] United Nations Convention on the Use of Electronic Communications in International Contracts (adopted 23 November 2005, entered into force 23 March 1976) 2898 UNTS 3, art 12.
[81] Turner (n 12 ) ch 3, s 2.5, sub-s 2.5.1.

Though not entirely clear, in the selected legal systems, there is a tendency to treat the decisions of robots as those of their users; the latter are held directly responsible for the obligations arising from contracts formed by robots. In Germany, not being recognised as legal persons, robots are not capable of making genuine declarations of will, but as they have become increasingly autonomous, it has been debated whether declarations made by robots should be regarded as such. In the legal literature, so far, the most common approach is to decide whether robots' declarations should be ascribed to their users since users often establish some core conditions for the decisions and declarations made by their robots. That approach argues that users should be held directly accountable for their robots' decisions, and these decisions should always be regarded as those of the users.[82] However, that approach is not workable in the long-term, as robots that are able to learn from their experiences can also modify their core conditions on their own, hence responsibility of for declarations made by such robots cannot be treated as if they had been given pre-acceptance by their users.

In Italian law and legal literature, no notable references address robot contracts. However, Law n.12/2019, passed in February 2019, defines a 'smart contract', as 'a computer program operating on technologies based on distributed registers and whose execution automatically constrains two or more parties based on the effects predefined by them.'[83] In other words, smart contract is 'the conversion of an agreement between two or more parties into a computer program which is capable of verifying that certain conditions/events are triggered and, thus, automatically execute certain actions defined.'[84] That statute can also apply to robot contracts by analogy, since robots can make autonomous decisions and these decisions are based on users' previously declared authorisations to govern the acts of their robots' behaviours. Accordingly, in Italy, just as in Germany, robots would be deemed to act in the name of and on behalf of the users, regardless of the degree of autonomy they display.

In Ireland, just as in Germany and Italy, any legally enforceable contract requires the intention to create legal relations. When robots are used as tools that simply facilitate the conclusion of contracts in accordance with pre-determined requirements stipulated by the parties.

---

[82] Lober, Caesar and Ropel (n 76) ch 6, s 6.6.
[83] Legge 12 febbraio 2019, n 12, GU 12 febbraiao 2019, n 36.
[84] Francesco Maruffi, 'Distributed Ledger Technologies and Smart Contracts In Italy' (*Blockchain - A Blog by Baker McKenzie*, 2019)

In the aforementioned first possibility, some form of binding agreement would indubitably be formed. However, where robots with higher degrees of autonomy are enabled to conclude contracts that go beyond the pre-determined core requirements identified by the parties, is not clear whether the intention to create legal relations would be regarded to be present and therefore whether binding agreements would be formed. In such cases, robots cannot be considered as simple artefacts to communicate the parties' intentions. Instead, it is necessary to regard them as agents that can make decisions independently from the contracting parties on whose behalf they are acting. Indeed, the contracting parties may very well be unaware that any purported contract has been concluded and it is not clarified whether intention to create legal relations would be accepted to be present.

## VI.2 Tort Law

Tort law systems are based on society's collective expectation that 'those who are engaged in some course of conduct will act with due care not to cast an unreasonable risk of injury upon others'.[85] The law protects that expectation through primary obligations imposed by tort law systems. Owed by everyone to everyone else, these primary obligations are best summarised in the precept *alterum non laedere* (not to do harm to others).[86] Failure to perform these primary obligations gives rise to secondary obligations. owed by tortfeasors to those they harmed:

> Primary obligations are the ground of tortfeasors' secondary responsibilities to repair the harms wrought by their torts. Repairing harm wrongly done is the next, or second-best way of discharging an obligation not to do the harm wrongly in the first place.[87]

In tort law systems, obligations can only be imposed on entities with the legal capacity to hold rights and obligations. Since it is just human beings (natural persons) and human associations (juridical persons) who possess that legal capacity under the law, it is assessed that robots can only be treated as instruments under the tort law systems, as is the case for other major domains of positive law.

[85] Roscoe Pound, *Social Control Through Law* (Routledge 2017) 14.
[86] Francesco Parisi, 'Alterum Non Laedere: An Intellectual History of Civil Liability' (1994) 39 The American Journal of Jurisprudence 317, 318.
[87] Gregory C Keating, 'The Priority of Respect over Repair' (2012) 18 Legal Theory 293, 294.

Then, the tort can be defined as an act that causes harm to another person's person and or property in contravention of the rules of law. Like crimes, torts are wrongdoings, but unless identified as such by the State, torts are not *ipso facto* crimes.[88] In the words of Winfield,

> Tortious liability arises from the breach of a duty primarily fixed by the law: such duty is towards persons generally and its breach is redressable by an action for damages[89]

In that regard, primary aims of tort law are to provide relief for the injuries suffered by parties whose rights have been infringed and to impose liability on the parties who caused the injuries, similar to those of contract law. However, contrary to contracts, the rights infringed in tort law do not arise from certain legal relations agreed between the parties. On the contrary, these rights are safeguarded against the 'whole' world.[90]

In terms of responsibility for the independent activities of robots, three types of tortious liability may be relevant: strict liability, fault-based liability, and vicarious liability. In the following section, how to impose liability when the tortfeasor is a robot is discussed. Notwithstanding the distinctions between civil law and common law legal systems, tortious liability can be addressed under three main categories: fault-based liability, strict liability, and vicarious liability.

The general, fault-based liability rule offers little-to-no remedy for victims of the tortious conduct of robots. It is not disputed that robots' manufacturers and masters still have primary obligations; they must act with reasonable due when exercising control over their robots. However, neither manufacturers nor masters are in any position to reasonably foresee and prevent harm that robots' independent activities may cause since they cannot be aware of all external and internal circumstances that influence robots' activities, thanks to the capacities of robots that are associated with the characteristic of relative autonomy.[91]

---

[88] Donal Nolan and John Davies, 'Torts and Equitable Wrongs', in Andrew Burrows (ed), *English Private Law* (3rd edn, OUP 2013) para 17.01. The principle of *nulla poena sine lege scripta* -there is to be no penalty without written law- requiring that one cannot be punished for doing something that is not prohibited by law. According to Scalia J., this principle is accepted and codified in most modern democratic states as a basic requirement of the rule of law. *Rogers v. Tennessee* 532 U.S. 451, 468 (2001) (Scalia, J., dissenting).

[89] Percy H Winfield, *The Province of the Law of Tort* (CUP 1931) 32.

[90] e.g., rights of bodily safety and freedom, rights of reputation, and rights of property Robert H Stevens, *Torts and Rights* (OUP 2009) 5; *Allen v Flood* [1898] 1 AC 1.

[91] External factors are incidental in uncontrolled environments and internal factors can be self-modified by robots. See Alice Guerra, Francesco Parisi and Daniel Pi, 'Liability For Robots I: Legal Challenges' (2021) 18 Journal of Institutional Economics 332, 334.

The present study does not mean to suggest that robots' tortious conduct can never stem from the negligence or intent of their manufacturers, masters, or other legal persons. Even when that is the case, however, complicated algorithms of robots will likely make it near-impossible to trace damages back to the fault of that legal person.[92] Where the general rule cannot provide remedy, tort law systems can still impose secondary obligations through alternative, no-fault liability regimes. No-fault liability regimes, as the name suggests, do not require the defendant to be at fault; it is sufficient that the defendant's activities caused the plaintiff's injuries. However, the failure to establish fault-based liability, on its own, does not provide sufficient grounds for resorting to these regimes.[93] As no-fault liability regimes are exceptions to the general rule, each regime is individually justified by some specific social interest.[94] In that regard, under current tort laws, the most prominent alternative liability regimes are strict product liability, liability for abnormally dangerous activities, and possession of wild animals.

Since robots can self-modify thanks to their autonomy, the damages they cause cannot always be traced to their manufacturers. Thus, strict product liability is assessed to be inapplicable to robots' tortious conduct.[95] Likewise, no abnormal dangers are associated with robots. On the contrary, interactions with them are safer than those with other human beings or animals; making the latter regimes of no-fault liability inapplicable as well.[96]

### VI.2.1 General Rule: Fault-Based Liability

The fault-based tortious liability is based on the defendant's intentional behaviour or negligence. In intentional torts, the responsibility for tortious behaviour falls on the party who has caused the damage on the other party by voluntarily performing the prohibited wrongful action. In negligence-based torts, the negligence refers to the concept of a reasonable person who fails to take precautions against a foreseeable harm, and thereby causes damages due to some unintentional fault, such as lack of due care.

---

[92] Caroline Cauffman, 'Robo-Liability: The European Union in Search of the Best Way to Deal With Liability for Damage Caused By Artificial Intelligence' (2018) 25 Maastricht Journal of European and Comparative Law 527, 529-530.
[93] F Patrick Hubbard, 'Allocating the risk of physical injury from "sophisticated robots": Efficiency, Fairness, and Innovation' in Ryan Calo, A Michael Froomkin and Ian Kerr (eds), *Robot Law* (Edward Elgar 2016) 37.
[94] Jules L Coleman, 'The Morality of Strict Tort Liability' (1976) 18 William & Mary Law Review 259, 267.
[95] Aysegul Bugra, 'Room for Compulsory Product Liability Insurance in the European Union for Smart Robots? Reflections on the Compelling Challenges', in Pierpaolo Marano and Kyriaki Noussia (eds), *InsurTech: A Legal and Regulatory View* (Springer 2019) 167.
[96] Hubbard (n 93) 38.

In German law, fault-based tortious liability arises when a person 'knowingly or negligently injures the life, body, health, liberty, property, or any right of another person illegally'.[97] This legal system regulates act (or omission) and fault as the elements of tort; but limits the general duty of care to not breaching another party's protected rights.

The most prominent provision regarding fault-based liability can be found in BGB Section 823 (1):

> A person who, intentionally or negligently, unlawfully injures the life, body, health, freedom, property or another right of another person is liable to make compensation to the other party for the damage arising from this.

According to BGB Section 276 (2), 'a person acts negligently if he fails to exercise reasonable care'. Of course, there has to be a chain of causality between the breach and damages, as well as a degree of predictability.

It could be very difficult to retrace the chain of causality from damage caused by robot to a person that is responsible. Considering the relative autonomy of robots, it might be even more difficult to determine whether such person has acted intentionally or negligently, therefore making fault-based liability challenging to establish. In that regard, the limitation on types of damages that can be compensated under fault-based liability in German law can prove to be problematic as well, as not all behaviours of robots are expected to lead to catastrophic consequences and injure protected rights of humans. Robots –especially social robots– operate in uncontrolled environments, and, as a result of their programming or due to random events, they may perform behaviours that have no consequence on for human health or property rights, but which cause annoyance.

For instance, in 2016, Tay, an AI bot was released by via Twitter caused controversy when it started to post offensive tweets, leading Microsoft to shut it down within only 16 hours after its launch. Tweets posted by Tay were very offensive, but they did not infringe any of the protected rights.[98] Later on, in 2019, vision algorithms trained on unbalanced data sets failed to recognize women or people of colour, thereby causing offence but not actually injuring the life, body, health, freedom of anyone, or infringing any of anyone's protected rights.[99]

---

[97] BGB Section 823 (1); Ralph Dornfeld Owen, 'Tort Liability in German School Law' (1955) 20 (1) Law and Contemporary Problems 72, 73.
[98] Jane Wakefield, 'Microsoft Chatbot Goes Rogue On Twitter' (*BBC News*, 2016).
[99] Will Knight and Karen Hay, 'Never Mind Killer Robots—Here Are Six Real AI Dangers To Watch Out for in 2019' (*MIT Technology Review*, 1 July 2019).

In Italy, the core of fault-based tort liability can be found in two provisions. The first of these provisions is CC Art 1173 which establishes that 'obligations arise from contracts, unlawful acts, or any other acts or facts which are capable of producing obligations under the law.' The other provision is CC Art 2043 which provides that 'any intentional or negligent act that causes an unjustified injury to another obliges the person who has committed the act to pay damages.' These provisions do not explain what negligence is, but the definition from CP Art 43 (3) can be resorted to by analogy with the meaning of negligence:

> negligent, i.e., contrary to intentional, when the event, even though reasonably foreseeable, is not desired by the actor and occurs because of carelessness, imprudence, or lack of skill, or failure to observe laws, regulations, orders or instructions.

This definition encompasses both 'the reasonable man standard' and categorical negligence. Though one might be able to claim larger damages in Italian law compared to German law, the criticisms about the shortcomings of fault-based liability instruments in German law apply here as well.

In Ireland, fault-based liability addresses 'the failure to exercise the care of an ordinarily prudent and careful man'.[100] Per Bryan McMahon and William Binchy, there are four elements in fault-based liability:

1) A duty of care, that is, the existence of a legally recognized obligation requiring the defendant to conform to a certain standard of behaviour for the protection of others against unreasonable risks;
2) A failure to conform to the required standard;
3) Actual loss or damage to recognized interests of the plaintiff; and
4) A sufficiently close causal connection between the conduct and resulting injury.[101]

In general, fault-based tort liability arises when there is a breach of a duty of care owed to the claimant that has caused damages. The principle which requires that the damages should be reasonably foreseeable is of central concern.[102] In applying these principles to damages caused by robots' behaviours, the first question is to identify persons who may have duties of care. The present study submits that the manufacturers, programmers, and users of robots, can all be regarded to have different duties of care, subject to varying standards of care regarding robot behaviours.

[100] Richard A Posner, 'A Theory of Negligence' (1972) 1 Journal of Legal Studies 29.
[101] Bryan McMahon and William Binchy, *Law of Torts* (4th edn, Bloomsbury 2013) para 5.02.
[102] ibid para 5.06.

To begin with, robots' manufacturers and programmers are under the duties to warn consumers of the risks, and to exercise reasonable care to prevent defects arising from manufacturing and programming processes, respectively.[103] Robots' manufacturers and programmers perform their duty to warn the consumers with warning labels placed on their end-products, and breach of this duty occurs when they fail to notify users of the risks or dangers of which they were aware of.

Whether the manufacturers and programmers have fulfilled their duties to exercise reasonable care to prevent defects is determined according to the standards of the industry they are participating in.[104] Examples of breach of this duty include bugs in the robot's programming that could have been detected by the programmer, or an incorrect or inadequate knowledge base. Also, robots' users are under the duty to take proper care and avoid foreseeable risks. Their duty includes obligations such as keeping the knowledge bases of robots up-to-date, or not using the robots for purposes beyond their specified functions.[105] It should be noted that if robots' manufacturers, programmers, and users show reasonable care and conform to the required standards, there shall be no breach of their duties of care, even if the robots' behaviours cause damages.[106]

Another question in applying the fault-based liability rules to robots' behaviours is whether there was a 'sufficiently close causal connection' between the breach of duty and damages. Here, 'close causal connection' not only refers to the factual link between the breach of duty and damages but also 'inquires into whether this factual link was proximate rather than remote'.[107]

The concept of foreseeability is the cornerstone of 'proximate' cause. Accordingly, the liability of a tortfeasor for the damages caused by his wrongful conduct is limited by the principles of reasonable foreseeability. In Wagon Mound No. 1, this approach is elaborated as follows:

> (…) if it is asked why a man should be responsible for the natural or necessary or probable consequences of his act (or any other similar description of them) the answer is that it is not because they are natural or necessary or probable, but because, since they have this quality, it is judged by the standard of the reasonable man that he ought to have foreseen them.[108]

---

[103] Peter M Asaro, 'Robots And Responsibility From A Legal Perspective', (2007 IEEE International Conference on Robotics and Automation, Workshop on Roboethics, Rome, April 2007)

[104] ibid.

[105] Maruerite E Gerstner, 'Liability Issues with Artificial Intelligence Software' (1993) 33 Santa Clara Law Review 239, 241-242

[106] Turner (n 12 ) ch 3, s 2.1

[107] David G Owen, 'The Five Elements of Negligence' (2007) 35 Hofstra Law Review 1671, 1674.

[108] *Overseas Tankship (UK) Ltd v Morts Dock & Engineering Company Ltd* [1961] AC 388, 27.

Indeed, foreseeability is intricately connected with negligence, bringing the doctrine *novus actus interveniens* to the forefront of the debate over the possibility of imposing fault-based liability for damages caused by robots' behaviours. *Novus actus interveniens* (i.e., new intervening act) is any act or event that breaks the causal connection between a wrong committed by the defendant and the subsequent damages suffered by the plaintiff, with the effect of relieving the defendant of liability for these damages.[109] The chain of causation may be broken because of defendant's actions,[110] natural events,[111] or due to actions of a third party.[112]

However, neither manufacturers nor masters are in any position to reasonably foresee and prevent harm that robots' independent activities may cause since they cannot be aware of all external and internal circumstances that influence robots' activities, thanks to the capacities of robots that are associated with the characteristic of relative autonomy.[113] The present study does not mean to suggest that robots' tortious conduct can never stem from the negligence or intent of their manufacturers, masters, or other legal persons. Even when that is the case, however, complicated algorithms of robots will likely make it near-impossible to trace damages back to the fault of that legal person.[114]

In the past, robots were perceived as automatons that 'do what they are told to do, in the way that they are told to do it, or the variation of means itself is programmed and utterly predictable.'[115] However, at the present time, in addition to automatons, there are also robots that are equipped with learning skills, and which contain algorithms that are capable of generating unpredictable behaviours. These characteristics are perhaps most evident in robots that are used as social interaction partners. Indeed, there are increasing numbers of social robots that can acquire knowledge or develop skills through their interactions with the living entities who inhabit their environment.

---

[109] *Home Office v Dorset Yacht Co Ltd* [1970] A.C. 1004; *Haynes v Harwood* [1935] 1 KB 146, 156.
For Ireland, see Ailbhe O'Neill, 'Rescuing the Law of Tort? The Decision of The Supreme Court In O'Neill v. Dunnes Stores' (2010) 45 Irish Jurist 240, 241-243; *Breslin v MIBI* [2003] 2 IR 203.
[110] *McKew v Holland* [1969] 3 All ER 1621; affirmed in Ireland in *Conole v Redbank Oyster Co* [1976] IR 191 (SC).
[111] *Steamship Co Ltd v Royal Norwegian Government* [1952] AC 292.
[112] *Knightley v Johns and Others* [1982] 1 WLR 349.
[113] External factors are incidental in uncontrolled environments and internal factors can be self-modified by robots.
See Alice Guerra, Francesco Parisi, and Daniel Pi, 'Liability For Robots I: Legal Challenges' (2021) 18 Journal of Institutional Economics 332, 334.
[114] Caroline Cauffman, 'Robo-Liability: The European Union in Search of the Best Way to Deal With Liability for Damage Caused By Artificial Intelligence' (2018) 25 Maastricht Journal of European and Comparative Law 527, 529-530.
[115] Curtis E A Karnow, 'The Application of Traditional Tort Theory to Embodied Machine Intelligence' in Ryan Calo, A Michael Froomkin and Ian Kerr (eds), *Robot Law* (Edward Elgar 2016) 74.

Besides, by their very nature, social robots can respond to external stimuli by modifying their inner states, and they can, on their own, adjust the rules that govern their functions to improve their efficiency. Consequently, in near future, it is submitted that even units of the same model will behave in quite different ways. Behaviours of will thus be unpredictable,[116] making increasingly difficult to apply the causation to the liability for robots and attribute liability to their owners. Accordingly, the doctrine of *novus actus interveniens* can be expected to consistently prevent imposition of fault-based liability for robots' behaviours. In the framework of fault-based tortious liability, the emergence of robots appears to challenge the basic notion of a reasonable person guarding against foreseeable harm. Taking the different approaches in all three legal systems into account, it is submitted that the reasonable person standard has to be modified in order to be able to address the behaviours of robots. Indeed, as explored above, complex decision-making mechanisms may be developed by autonomous artificially intelligent agents-including robots- that are able to gain knowledge or skills by interacting with the entities in their environments, and even when it comes to their programmers, it can be difficult for the courts to determine when they acted unreasonably.

### VI.2.2 Strict Liability

Where the general fault-based liability rule cannot provide remedy, tort law systems can still impose secondary obligations through alternative, no-fault liability regimes. No-fault liability regimes, as the name suggests, do not require the defendant to be at fault; it is sufficient that the defendant's activities caused the plaintiff's injuries. However, the failure to establish fault-based liability, on its own, does not provide sufficient grounds for resorting to these regimes.[117] As no-fault liability regimes are exceptions to the general rule, each regime is individually justified by some specific social interest.[118] The law ascribes strict liability to 'those activities it considers useful and necessary but that create abnormally dangerous risks to society'.[119] In other words. when it comes to the cases of strict liability, the law imposes liability regardless of the person's intention or use of ordinary care.

---

[116] Ugo Pagallo, 'Even Angels Need the Rules: AI, Roboethics, and the Law', *European Conference on Artificial Intelligence (ECAI)* (IOS 2016) 209, 210.
[117] F Patrick Hubbard, 'Allocating the risk of physical injury from "sophisticated robots": Efficiency, Fairness, and Innovation' in Ryan Calo, A Michael Froomkin and Ian Kerr (eds), *Robot Law* (Edward Elgar 2016) 37.
[118] Coleman (n 94) 267.
[119] Woodrow Barfield, 'Liability for Autonomous and Artificially Intelligent Robots' (2018) 9 Paladyn Journal of Behavioral Robotics 193, 197.

### VI.2.2.a Strict Product Liability

Strict product liability focuses on the defective condition of the product itself, and away from an examination of the defendant's conduct in making the product, since 'the innocent victim of a dangerous product should be compensated, even if the defendants were not negligent in making it'.[120] This study examines 'strict product liability' for robots based on the Product Liability Directive of the European Union which provides for one of the most developed 'strict product liability" regimes in the world.[121] The EU Directive holds the 'producer' liable for the damages caused by a defective product.[122] Under the EU Directive, a product is defective when,[123]

> it does not provide the safety which a person is entitled to expect, taking all circumstances into account, including: (a) the presentation of the product; (b) the use to which it could reasonably be expected that the product would be put; (c) the time when the product was put into circulation

For the purposes of the EU Directive, "damage" means damage caused by death or by personal injuries; damage to an item of property intended for private use or consumption other than the defective product itself, with a lower threshold of €500.[124] The injured person must prove the actual damage, the defect in the product and the causal relationship between damage and defect.[125]

Fault on the part of the producer does not need to be proven. The producer is, however, freed from liability if he proves that, "he did not put the product into circulation, the defect was due to the compliance of the product with mandatory regulations issued by public authorities,' or 'the state of scientific or technical knowledge at the time the product was put into circulation could not detect the defect.'[126]

To begin with, it must be noted that it is debated whether robots fall under the scope of the EU Directive as 'products'. For its own purposes, the EU Directive indicates that the term 'product' includes all movables ('even when incorporated into another movable or into an immovable object') and electricity.[127]

---

[120] Karnow (n 115) 66.
[121] Turner (n 12 ) ch 2, s 2.2.
[122] Council Directive 85/374/EEC of 25 July 1985 on the Approximation of the Laws, Regulations and Administrative Provisions of the Member States Concerning Liability for Defective Products [1985] OJ L210/29, art 1.
[123] ibid, art 6.
[124] ibid, art 9.
[125] ibid, art 4.
[126] ibid, art 7.
[127] ibid, art 2.

The Directive does not directly address software, which brings into existence all the characteristics of robots and maintains them. However, it is determined that 'most commentators are in agreement that the software and its medium constitute a tangible product' even though 'the information contained within the software medium is intangible and cannot always be regarded as a product'.[128] According to this consensus, robots can be regarded as 'products' in the context of the EU Directive, regardless of whether they are embodied or not and their programmers must be regarded as 'producers'. Even if it is accepted that the EU Directive applies to robots, the said Directive is designed on the assumption that products are static. This is most obvious when the defences provided in the Directive against liability claims are examined:

> (…) having regard to the circumstances, it is probable that the defect which caused the damage did not exist at the time when the product was put into circulation by him or that this defect came into being afterwards; or (…) that the state of scientific and technical knowledge at the time when he put the product into circulation was not such as to enable the existence of the defect to be discovered (…) [129]

Robots, thanks to the capabilities associated with the characteristic of relative autonomy can continue changing in unpredictable ways after they leave the production lines. Thus, the defences provided against 'strict product liability' under the EU Directive claims appear to be unduly permissive for robot producers, as robots, by definition, are going to have defects that cannot be determined at the time of initial circulation. It is likely that robot manufacturers will increasingly be able to take advantage of these defences and thereby gradually render the protection of fortify, the EU product-liability directive insufficient. Not only can consumers find it difficult to demonstrate the existence of a defect in their robots, but also the liability of manufacturers mostly depends on when the product is put into circulation. No responsibility for updates or upgrades, seem to have been provided by the Directive.

[128] Karin Alheit, 'The Applicability of the EU Product Liability Directive to Software' (2001) 34 Comparative and International Law Journal of Southern Africa 188, 203-204.
[129] Council Directive 85/374/EEC of 25 July 1985 on the Approximation of the Laws, Regulations and Administrative Provisions of the Member States Concerning liability for Defective Products [1985] OJ L210/29, art 7.

### VI.2.2.b Liability for 'Abnormally-Dangerous Activities'

Liability for 'ultra-hazardous' or 'abnormally dangerous' activities' focuses on activities that are considered so dangerous 'that the law makes those who are engaging in them in effect insurers to others who are hurt without requiring proof of negligence and other types of fault'.[130] In all selected legal systems, an activity may be regarded ultra-hazardous if it is uncommon, poses a high risk of harm, and creates a high degree of injury when injury does occur. Examples of ultra-hazardous activities include blasting and demolitions, or disposing of nuclear and chemical wastes. However, there are differences between legal systems regarding whether abnormally dangerous activities are exhaustively enumerated or not.[131]

In Germany, strict liability is imposed in relation to specific hazards by separate statutes - there is no *general* liability for dangerous activities in German law. To cite a few instances of liability for dangerous activities, Section 7 of Road Traffic Act (StVG),[132] provides strict liability in respect of road traffic accidents; Section 8 of Federal Data Protection Act (BDSG)[133] for breaches of data security, and BGB Sections 833/ and 834 for BGB for injuries caused by animals. There are other statutes that provide for harm caused by aircraft and so on.[134] In German law, for strict liability for ultra-hazardous activities to be enforced, it is sufficient that the defendant possesses the source of danger; and he can be held liable without being at fault.

In Italy, unlike in Germany, CC Art 2050 imposes general liability for damages arising from exercise of dangerous activities:

> Whoever causes injury to another in the performance of an activity dangerous by its nature or by reason of the instrumentalities employed, is liable for damages, unless he proves that he has taken all suitable measures to avoid the injury.

Accordingly, in Italian law, there is no exhaustive enumeration of dangerous activities: If damage is caused in execution of a dangerous activity, operator is deemed liable – without the need for any specific statutory provision. However, the liability of dangerous activities is not absolute in Italian law, if the operator proves that he has taken all suitable measures,

---

[130] Harold G Christensen, "Torts: Strict Liability for Ultra-Hazardous Activities" (1951) 49 Michigan Law Review 919; Karnow (n 125) 67.
[131] Elspeth Reid, 'Liability for Dangerous Activities: A Comparative Analysis' (1999) 48 International and Comparative Law Quarterly 731, 733-735.
[132] [Straßenverkehrsgesetz] [StVG] [Road Traffic Act], May 3, 1909, RGBl at 437, revised March 5, 2003, BGBl I at 310, last amended by Art 1 of the Law of July 12, 2021, BGBl I at 3108.
[133] [Bundesdatenschutzgesetz][BDSG] [Federal Data Protection Act], June 30, 2017, BGBl I at 2097, last amended by Art 10 of the Law of June 23, 2021, BGBl I at 1858.
[134] [Luftverkehrsgesetz] [LuftVG][Aviation Act], August 1, 1922, RGBl at 681, revised May 10, 2007, BGBl I at 698, last amended by Art 131 of the Law of August 10, 2021, BGBl I at 3436.

he is absolved from liability. This provision in Italian law appears to suggest that most challenges posed by robots are covered.

In Irish law, strict liability arising from dangerous activities is based on the rule in *Rylands v Fletcher,* a UK case. The rule imposes strict liability if a proprietor accumulates dangerous substances which escape from his land and cause damage to other properties. The substances subject to this rule include anything 'likely to do mischief if it escapes'.[135] The case involved two neighbouring coal mining operators. The defendant built a reservoir to provide water to his mill, but his contractors failed to discover an underground tunnel that connected to the plaintiff's mine. The water broke from the reservoir and flooded the mine. In the case, the defendant was not negligent; the contractors may have been, but he was not responsible for them, since, at that time, the liability of independent contractors could not be established. In response, the House of Lords formulated the 'rule in Rylands v. Fletcher' which imposes strict liability if a defendant accumulates dangerous substances (anything likely to do mischief if it escapes) which escape from his land and cause damage to properties.

Over the years, the rule was applied to damages caused by the escape of a wide range of substances, including noxious gases,[136] explosives,[137] sewage, poisonous trees,[138] and even persons likely to cause a disturbance if they escape.[139] In *Cambridge Water Company v Eastern Counties Leather*, the House of Lords considered the rule as an extension of the law of nuisance to cases of isolated escape; and as under nuisance law damages are awarded only if injury of the relevant type was foreseeable by the defendant. As such, foreseeability was accepted as one of tbe components of the rule.[140] Also, *Transco plc v. Stockport*[141] explicitly qualifies the foreseeability requirement so that the plaintiff only have to show that the defendant should have foreseen that there could be injury if the object escaped, not that he should have foreseen that there might be an escape.

[135] *Rylands v Fletcher* [1868] UKHL 1.
[136] *Halsey v Esso Petroleum* [1961] 2 All ER 145.
[137] *Read v J Lyons & Co Ltd* [1947] AC 156, [1946] UKHL 2.
[138] *Noble v Harrison* [1926] 2 KB 332.
[139] *A-G v Corke* [1933] Ch 89.
[140] *Cambridge Water Co Ltd v Eastern Counties Leather plc* [1994] 1 All ER 53; Elspeth Reid, 'Liability for Dangerous Activities: A Comparative Analysis' (1999) 48 International and Comparative Law Quarterly 731, 740.
[141] *Transco plc v. Stockport MBC* [2004] 2 AC 1.

It is submitted that 'liability for ultra-hazardous activities' theory is not well-suited to determine liability for robot behaviours, due to several reasons.[142] First of all, the statistical data already indicates that robots are unlikely to regularly pose hazards. According to the incident reports, since 1984, only 41 robot-related accidents occurred in the US, out of 20,890 workplace accidents.[143] That number looks even smaller considering that there were 233,305 industrial robots deployed in the US as of 2017. Moreover, many of the injuries caused by the independent activities of robots are not necessarily fatal or even serious.[144] Then, since 'liability for ultra-hazardous activities', requires most robot behaviours to be foreseeably or stereotypically dangerous, the regime is not useful in determining the liability for the behaviours of robots. Furthermore, if the rule regarding liability for ultra-hazardous activities remains as one of the main mechanisms through which risk and liability are distributed, technological innovation in society may be hindered and a rather vicious circle may be created. Indeed, per Barfield and Pagallo, stricter the liability rules are, 'the less we can test our AI systems, and the more such rules can have a chilling effect'.[145]

### VI.2.2.c Vicarious Liability

Vicarious liability refers to the secondary responsibility of superiors for the acts and omissions of their subordinates, or, in a broader sense, the responsibility of any person that has sufficient control over the actions of another. Vicarious liability is understood as liability imposed upon a person simply because of their relationship to the person who has committed the wrong, even though they have committed no wrong themselves.[146]
For instance, the law may hold the employer liable for the wrongs of an employee, the principal liable for the wrongs of an agent, or the firm liable for the wrongs of its partner, in spite of the fact that the employer, the principal, or the firm may not have been at fault in any way. When the law imposes liability in these circumstances we speak of the employer, principal, or firm being 'vicariously liable'.[147]

---

[142] Karnow (n 125) 67.
[143] "Accident Search Results" (*Occupational Safety and Health Administration,* 1 March 2019)
[144] Sami Haddadin and others, 'Injury Evaluation of Human-Robot Impacts' (2008 IEEE International Conference on Robotics and Automation, Pasedena, May 2008).
[145] Woodrow Barfield and Ugo Pagallo, *Advanced Introduction to Law And Artificial Intelligence* (Edward Elgar 2020) 97.
[146] Nolan and Davies (n 88) para 17.367.
[147] McMahon and Binchy (n 101) para 43.01.

In Germany, vicarious liability is regulated by BGB Sections 831 and 832. Under BGB Section 831 (1):

> A person who uses another person to perform a task is liable to make compensation for the damage that the other unlawfully inflicts on a third party when carrying out the task. Liability in damages does not apply if the principal exercises reasonable care when selecting the person deployed and, to the extent that he is to procure devices or equipment or to manage the business activity, in the procurement or management, or if the damage would have occurred even if this care had been exercised.

According to Merriam-Webster's dictionary, 'a person who uses another person to perform a task' can be called as an employer. Thus, the present study submits that BGB BGB Section 831 can also be approached as employers' liability.

Liability in BGB Section 831 is based on fault, but the employer's fault is presumed, and the burden of proof is reversed since the section requires the employer to demonstrate that care was taken in the appointment, supervision or management of employees or simply that the damage would have occurred in any event. For the vicarious liability of the employer to arise, the employee must have been acting in the course of their employment when the event causing the damage happened. Even if the act was performed in the course of employment, intentional wrongful acts, such as stealing, do not engage the liability of the employer. Moreover, the employer is not responsible for the damages suffered by the employee on their way to and from the place of employment.[148] BGB Section 832 imposes liability for the wrongful acts of minors and persons in need of care, on those who are obliged to supervise them. The liability of parents, guardians, and teachers for those in their care comes under this section, albeit the text of this Section does not explicitly mention any of them.

BGB Section 1626(2) and Section 1631(1) impose on parents the obligation to supervise and guard their minor children, and if there are no parents, guardians are imposed the same obligations under BGB Section 1793 and Section 1800. Liability there is based on the presumption that the parents did not exercise due care in the supervision of their child and this presumption may be rebutted by showing that all necessary care was taken.

[148] 08.01.1926 RGZ 112, 290 III. Civil Senate (III 32/25).
See also Karsini Vanessa Soerjaman, 'Employer's Liability For Damage Caused by the Employee, A Comparative Study: Germany and the Netherlands' (LLM, University of Tilburg 2013) 17.

Under BGB Section 832, no fault is required on the part of the minor or person under supervision: Then the issue is whether the parents were negligent in permitting the person enough scope to cause harm, i.e., whether they did take the necessary and possible steps to prevent the harm occurring.[149] In Italian law, in contrast, vicarious liability is regarded as a form of strict liability, meaning that once the plaintiff establishes the other elements of the case, the defendant will be held strictly liable unless he can exonerate himself according to the formula prescribed by the various provisions of CC. These establish that a person is liable not only for the damages he causes by his own act, but also for that which is caused by the acts of persons for whom he is responsible,[150] or by things which are in his custody.[151] Considering that currently the Italian law addresses robots as things, CC Art 2051 may be particularly relevant for the liability of robot behaviours:

> Everyone is liable for injuries caused by things in his custody, unless he proves that the injuries were the result of a fortuitous event.

CC Art 2051, similarly to other strict liability provisions, excludes any need to prove fault and holds the custodian liable for any damage the thing in custody causes, unless such actions rise to the level of a "fortuitous event". In the case law, a fortuitous event is described 'as the intervention of an element endowed with autonomous causal impetus and with an unpredictable and unavoidable character.'[152]

In Ireland, vicarious liability indicates the person made liable is a joint tortfeasor with the actual wrongdoer, and the victim can sue either or both of them.[153] Moreover, in principle, vicarious liability is secondary and, after having paid the damages, the person who was made vicariously liable has a right to reimbursement from the actual wrongdoer.[154] In other words, for vicarious liability, there has to be some relationship between the parties, and the wrong must be committed within the scope of that relationship.[155]

---

[149] Pieter Pauw, 'The Liability of Parents For Loss Caused By Their Children' (1978) 11 The Comparative and International Law Journal of Southern Africa 305, 314; Gerhard Wagner, 'Children as Tortfeasors Under German Law' in Miquel Martín-Casals (ed), *Children in Tort Law Part I: Children as Tortfeasors* (Springer 2006) 217, 245.
[150] Kwame Opoku, 'Delictual Liability In German Law' (1972) 21 International and Comparative Law Quarterly 230, 239.
[151] CC arts 2051, 2052.
[152] Court of Cassation, 14 January 1992, n. 347, *Bernabei v Palazzari*. It appears that 'fortuitous event' is akin to *novus actus interveniens* in Irish law, which is an unforeseeable event that interrupts the chain of causation.
[153] Nolan and Davies (n 88) para 17.375.
[154] Turner (n 12 ) ch 3, s 2.3.
[155] ibid.

It is submitted that vicarious liability might be utilised in determining the responsibility for robots, but that it would require specific regulations. This is because in the case of the application of vicarious liability, the relative autonomy of robots would be given some effect and legal persons would be liable to pay for the damages arising for the behaviours of robots.[156] As such, the doctrine better accommodates the characteristics of robots which differentiate them from other artefacts.

However, as the possible application of vicarious liability also means that not all behaviours of robots would be ascribable to a legally recognised persons/ as such, it only provides a partial solution.[157] In fact, in most cases, the defendant would still be relieved from any liability as the robot behaviours were 'unpredictable' to them, blurring the line between vicarious liability and other forms of liability in terms of usefulness. It cannot be known whose actions have sown the seeds of independent activities that resulted in undesirable outcomes. Admittedly, the actions of manufacturers or masters are the first ones to come to mind. However, since robots operate in uncontrolled everyday environments, it must be acknowledged that the independent activities that result in undesirable outcomes can also be prompted by their random human interactants' or even bystanders' actions.

[156] Samir Chopra and Laurence F White, *A Legal Theory for Autonomous Artificial Agents* (UM 2011) 130.
[157] Ugo Pagallo, *The Laws of Robots* (Springer 2013) 117.

## Summary

In the context of contract law and tort law systems, recognising robot agency requires giving them the legal capacity to act or 'the power to perform acts with legal effect'.[158] In fact, robots are already capable of performing behaviours that have legal effects: Where robots' independent activities cause damage, remedial obligations will arise, though it is not clear on whom. Similarly, when robots conclude contracts, binding rights and obligations are created to be borne by whoever authorised them.[159] Though robots *de facto* have the power to perform acts with legal effects, such acts cannot be legally considered theirs unless they are given the legal capacity to act. In that respect, the present study identifies the law's lack of distinction between the autonomy of robots and the free will of human beings who develop or deploy them as the root cause of legal issues arising in contract law and tort law systems. Without that distinction, robots can only be treated as machines fulfilling their programmed tasks. Therefore, the liability of those who use them can only be absolute, ultimately hampering the protection of social interests in general progress.

Then, if robots are given the legal capacity to act, it becomes possible to provide robots' manufacturers and masters with some safeguards, such as opportunities to limit their liability. In this chapter, the shortcomings of existing legal instruments in determining the private law liability for the autonomous of robots have been assessed. Across the selected jurisdictions, both contract law and tort law have shortcomings in the determination of the liability for robot behaviour.

However, when robots are utilised as simple tools for human interactions, the existing legal instruments are sufficient, and these shortcomings are not visible. This situation suggests that the existing private law legal instruments have not yet adapted to the characteristics of robots as unique artefacts. These shortcomings in determining private law liability for behaviours of robots are due to the growing autonomy of robots.

Since the capabilities associated with the characteristic of relative autonomy reduce the predictability of robot behaviours, it is not possible to apply the law of negligence, based on foreseeability, to determine the liability for independent activities robots.

---

[158] Andrea Bertolini and Francesca Episcopo, 'Robots and AI as Legal Subjects? Disentangling the Ontological and Functional Perspective' (2022) 9 Frontiers in Robotics and AI 6-7.
[159] Paulius Čerka, Jurgita Grigienė and Gintarė Sirbikytė, 'Liability for Damages Caused by Artificial Intelligence' (2015) 31 Computer Law & Security Review 376, 382.

This situation confines the claims for damages caused by the behaviour of robots to the domains of strict liability and vicarious liability, which are limited instruments that may be resorted to only in exceptional circumstances. It is submitted that the existing legal instrument that is best placed to determine liability for the behaviours of robots is that of vicarious liability, as it would amount to the recognition of robots as agents. However, as vicarious liability can only be used in contexts that are similar to the deployment of robots to complete determinate and specific tasks, it would not apply to all injuries arising from the independent activities of robots, revealing the need for new civil liability instruments.

# PART D

# *Pax Robota*

## Recommendations and Conclusions

> The task of the human brain remains what it has always been; that of discovering new data to be analysed, and of devising new concepts to be tested. A pity the Society for Humanity won't understand that.
>
> — Isaac Asimov, *I, Robot*[1]

[1] Isaac Asimov, *I, Robot* (first published 1967, HarperVoyager 2018) 241.

# Chapter VII: Feasibility of the New Legal Framework

> Robots are not people. They are mechanically more perfect than we are, they have an astounding intellectual capacity, but they have no soul.[1]
>
> —Karel Capek, *Rossum's Universal Robots*

## Introduction

In the previous chapters of the present study, it was established that existing legal instruments cannot adequately address the legal issues surrounding human-robot social interactions. The law does not give effect to the distinctive characteristics that distinguish social robots from other functional artefacts. The characteristic of relative autonomy that all robots have, regardless of whether they are social or non-social, is similar to the power of free will of human beings. Unlike the characteristic of relative autonomy, however, human beings' free will is empowered by legal systems through the recognition of their capacity to act which provides for the legal effect of decisions that are made with free will, and, as such, it can be understood the power to create rights and obligations through one's own actions.

The capacity to act is one of the two elements of legal personhood. The other element is juridical capacity, the susceptibility of the legal person to have rights and obligations. The juridical capacity refers to the fitness to be subject to legal relations in it's own name. Since the capacity to act is closely associated with free agency, it can be acquired and may be lost in accordance with the changes regarding the entity's power of agency. The juridical capacity is linked, however, to the needs, desires, wants, goals, in short, interests of legal persons. As long as the legal person is functional – it will have some individual interests, and regardless of whether the legal person has the capability of agency to pursue these interests, they there maintain their juridical capacity. For that reason, juridical capacity can be regarded as the heart of legal personhood.

It was established earlier that since they are functional artefacts that are devised to fulfil prescribed functions, robots cannot have any needs, desires, wants, or objectives on their own to motivate their independent activities. Considering that robots cannot have any interests on their own, they cannot logically be subject to legal relationships in their own rights.

[1] Karel Čapek, *R.U.R.: Rossum's Universal Robots* (David Wyllie tr, originally published 1920, eBooks@Adelaide 2014).

Therefore, it is evident that robots do not have any characteristics that would warrant them to be equipped with juridical capacity. Considering that juridical capacity is the core of legal personhood and robots have no functional use for that capacity, the new legal status that is outlined to address the shortcomings of existing legal instruments cannot be called legal personhood.

Though the conferral of legal personhood on robots will not solve the legal issues surrounding human-robot interactions, the solution to these problems consists in the creation of some form of legal status for robots that recognises that the characteristic of relative autonomy is the functional equivalent of legal persons' capability to act. The lack of legal recognition of that characteristic prevents the law from addressing the imbalances between legally protected interests that may arise because of the complications that these characteristics result in. To reiterate, in criminal law and private law, existing legal instruments are insufficient because they connect the liability to be imposed with human beings' use of free will even in no fault liability scenarios. The link between the free will and responsibility is often explicit, as is the case in formation of contracts, and crimes and torts where fault of the defendant should be proved. Sometimes however, the link is required through the concepts such as foreseeability or due diligence, albeit at weaker degrees. Since no such link exists between undesirable outcomes of robots' independent activities and the free will of their manufacturers and masters, existing legal instruments cannot quite offer responses to the legal issues surrounding human-robot social interactions. Any comprehensive solution of the legal issues surrounding human-robot social interactions, especially the challenge of individual autonomy requires the creation of a new legal status for robots.

The following chapter assesses whether the introduction of a new legal status for robots is feasible under legal systems chosen for comparison for the purposes of the present study. The chapter explores whether the conferral of some form of legal status on robots has the potential to solve the problems posed by the characteristic of individual autonomy of robots and those posed by the phenomena of robot anthropomorphism. We first consider whether the creation of a legal status beyond property but short of personhood is structurally possible, and secondly, whether it would at least be desirable from the utilitarian perspective.

Though conferral of legal personhood does not address the legal issues at hand, perhaps because of the degree of anthropomorphic responses invoked by humanoid social robots, several countries have given -symbolically- some such status to specific robots.

In October 2017, the Kingdom of Saudi Arabia granted citizenship to a humanoid social robot named Sophia.[2] This was the first time that a country purported to grant a robot legal personhood in its own right, assuming that only legal persons can be citizens. Following the footsteps of the Kingdom of Saudi Arabia, in November 2017, Japan granted residency to an AI chatbot named Mirai.[3] Even though these are the first robots to be granted places on real-life registries, and even though granting these robots citizenship and personhood is not intended to have any legal consequences, they signify the start of the legal trend of recognising robots as more than just functional artefacts. [4]

## VII.1 Conceptual Possibility of New Legal Status

Whether some new legal status, distinct from the existing legal statuses of property-hood and personhood can be introduced for robots, equipping them with the legal capacity to act is ultimately a question of legal ontology: Is it theoretically possible for national legal systems examined for the purposes of the present study, to confer robots with the legal capacity to act? The concept of legal personhood is of critical importance here, as the only two legal statuses that are available are those of property and legal persons, and the proposed framework suggests that robots should be moved beyond the status of the property.

The word 'person' connotes slightly different meanings in everyday use and as a legal term. Colloquially, person refers to an individual human being;[5] and it is imperative to note the distinction between this intuitive understanding and the legal meaning of 'person'. [6] Indeed, the legal notion personhood does not only apply to individual human beings and is not the guaranteed complement of the physical existence of human beings' physical existence.[7] In other words, 'legal person' is one of the abstract notions maintained by law for functional purposes and it refers to any entity that does at least have the juridical capacity.

---

[2] Marie Boran, 'Saudi Arabia Becomes First Country to Grant A Robot Citizenship' (*The Irish Times*, 2 November 2017)
[3] Anthony Cuthbertson, 'An AI Chatbot Just Became A of Japan' (*Newsweek*, 6 November 2017)
[4] Lawrence B Solum, 'Legal Personhood for Artificial Intelligences' (1992) 70 North Carolina Law Review 1231; S M Solaiman, 'Legal Personality of Robots, Corporations, Idols and Chimpanzees: A Quest for Legitimacy' (2016) 25 Artificial Intelligence and Law 155, 172.
[5] John Chipman Gray, *The Nature and Sources of Law* (Adamant Media Corporation 2006) 27.
[6] Samir Chopra and Laurence F White, *A Legal Theory for Autonomous Artificial Agents* (UM 2011) 153.
[7] Here, the reference is made to the institution of slavery, *see below* footnote 19.
Maximilian Koessler, 'The Person In Imagination or Persona Ficta of the Corporation' (1949) 9 Louisiana Law Review 435.

Origins of non-human 'legal persons' can be traced back to Roman law, which was said to be 'systematically ignorant of the biological status of its subjects'.[8] 'Collegium' in ancient Rome expressed the association of several persons being united in any office or in any common purpose. The 'collegia' were viewed as being capable of holding property, contracting and owing obligations, and, in having these capacities, were considered 'purely fictious or artificial personages' acting through the agency of their members.[9] The fiction of 'legal person' saw little use in the aftermath of the disintegration of the Roman Empire, until it was revived during the Middle Ages by the Catholic Church.

The doctrine of *persona ficta*, attributed to Pope Innocent IV, distinguished between 'that which God made, persons with souls, and that which man made, persons without souls'. [10] This thirteenth-century doctrine, was a useful device for handling the property of monasteries whose members, having taken holy orders, were 'property-less and civilly dead'. Accordingly, no monk in a monastery was a legal person, but the monastery itself was and could therefore own its buildings and grounds and enter into contracts to buy and sell land and goods.[11] In the following centuries, when joint stock companies began to emerge, it was not clear how the law would evolve to accommodate them.

The first case where an English court contended with corporate existence was the 1612 *Case of Sutton's Hospital* decided by Sir Edward Coke. Coke, in order to make sense of corporate existence, turned to canon law and the doctrine of *persona ficta*, and concluded that corporations were legal persons, but they were different in kind from human beings and could therefore be treated differently under law: [12]

> [T]he Corporation itself is only *in abstracto,* and rests only in intendment and consideration of the Law; for a Corporation aggregate of many is invisible, immortal and rests only in intendment and consideration of the law. […] They may not commit treason, nor be outlawed, nor be excommunicated, for they have no souls, neither can they appear in person, but by attorney.

Since they cannot functionally exercise any juridical capacity, robots cannot be afforded legal personhood, however, the other capability associated with personhood, the capacity to act, can help overcome the inadequacies of existing legal instruments.

---

[8] Peter A French, *Collective and Corporate Responsibility* (Columbia UP 1984) 135
[9] George Long, 'Collegium', *A Dictionary of Greek and Roman Antiquities* (1875) 310, 311.
[10] John Dewey, 'The Historic Background of Corporate Legal Personality' (1926) 35 The Yale Law Journal 655, 666.
[11] Ladelle McWhorter, 'The Morality of Corporate Persons' (2017) 55 The Southern Journal of Philosophy 126, 129.
[12] ibid. 130; *Case of Sutton's Hospital* (1612) 10 Rep 32, 77 Eng Rep 960, 973.

The conceptual possibility of conferring legal status on robots depends on how flexible the ontology of law is and whether it is possible the divide the components of the notion of personhood. The present study adapts the analysis of Bryson and others that legal persons are 'fictive, divisible, and not necessarily accountable' to the notion of legal status[13] The first observation, states that legal personhood is fictive, and means that legal status accorded to any entity is created by the law.[14] Accordingly, conferring legal status on an entity implies that the legal system addresses its rules to that. Since notion of legal status is fictive, inherent characteristics of the entity in question do not determine whether it can be accorded that legal status or not. [15] Indeed, there is now a long history of conferring legal personhood on corporations whilst recognizing that corporate personhood is fictional. Consider, for instance, the analysis articulated by the United States Supreme Court in the 1819 case of *Trustees of Dartmouth College v Woodward*:

> A corporation is an artificial being, invisible, intangible, and existing only in contemplation of law. Being the mere creature of law, it possesses only those properties which the charter of its creation confers upon it, either expressly, or as incidental to its very existence. (…) Among the most important are immortality, and, if the expression may be allowed, individuality; properties, by which a perpetual succession of many persons are considered as the same, and may act as a single individual. They enable a corporation to manage its own affairs, and to hold property, without the perplexing intricacies, the hazardous and endless necessity, of perpetual conveyances for the purpose of transmitting it from hand to hand.[16]

The Supreme Court, in another case, emphasises that 'the corporate personality is a fiction, although a fiction intended to be acted upon as though it were a fact.'[17]

the recognition of corporations as legal persons, although they do not have any form of real-world capabilities or presence that is distinguishable from those of human beings that constitute the corporations' agents, organs, and employees, makes it possible to confer on corporations some of the rights and obligations given to human legal persons, e.g. the right to bind others by contract and the obligation to satisfy their commitments under contract.[18]

To sum up, it can be submitted that it is not necessary to assess whether an entity satisfies any requirements in a biological, metaphysical, ethical, or even social sense when addressing the question of its legal status.

---

[13] Joanna J Bryson, Mihailis E Diamantis, and Thomas D Grant, 'Of, For, and by the People: The Legal Lacuna of Synthetic Persons' (2017) 25 Artificial Intelligence and Law 273, 279.
[14] ibid 282
[15] Jacob Turner, Robot Rules (Kindle edn, Palgrave Macmillan 2018) ch 5, s 2, sub-s 2.1.
[16] *Trustees of Dartmouth College v. Woodward* 17 U.S. (4 Wheat.) 518 (1819) 636.
[17] *International Shoe Co. v. Washington* 326 U.S. 310 (1945) 316.
[18] ibid.

In fact, legal status accorded to any entity, regardless of whether that entity is real or conceptual simply results from a decision of the legal system to confer legal status on a given entity - nothing more, nothing less.[19] Lawmakers, in that regard, should make the decision to confer legal status on any given entity after considering the legal system's ultimate objectives, and considering whether conferring personality would advance or hinder these objectives. Since the objectives of a legal system might change over time, it must be submitted that so might those on whom and what degree of legal status is conferred. For example, in Roman law, some human beings -slaves- were denied legal personhood due to the legal systems' objective of perpetuating privileges of some smaller group of persons.[20] With the abolition of slavery in all jurisdictions, it is accepted that all natural persons are legal persons. Likewise, granting personhood to corporations (that are not real persons) may be due to the objective of advancing economic progress.[21] In short, it can be said that there is nothing in the nature of legal systems that prevent the recognition of legal personhood to robots, and there is significant precedent to enable it.

Bryson and others' second observation is that legal personhood is divisible to its components. The basis of this observation is that since the legal status is made up of posited rights and obligations of the entity, 'entities can have more, fewer, overlapping, or even disjointed sets of these.'[22] The observation that legal status is divisible does not mean that those that have the legal status, for instance, legal personhood can be divided, it merely states that not all legal persons have the same rights and obligations, and it applies both to the legal personhood of human beings and to juridical persons. In other words, different types of legal persons can be accorded different capabilities. The second observation supports our evaluation of legal personhood as one type of legal status and there can be other types with different combination of the two core legal capabilities that make up legal personhood, even though Bryson and others were referring to differences among the extent and types of rights and obligations that different legal persons can exercise and hold.Typically, to fully enjoy their rights and become subject to legal enforcement obligations as legal persons, human beings should be of the age of majority and of sound mind.

[19] Bryson, Diamantis, and Grant (n 12) 281.
[20] William Smith, 'Servus', *A Dictionary of Greek and Roman Antiquities* (1875) 1036, 1038.
[21] Turner (n 15) ch 5, s 3, sub-s 3.6.2.
[22] Bryson, Diamantis, and Grant (n 12) 282.

For example, minor children, even though they are natural persons, cannot vote, sign contracts, marry or become involved in various sorts of legal arrangements because they do not have the capability to perform legally significant actions.[23] The case of permanently intellectually impaired persons is similar to that of children.[24] Moreover, the struggle for equal rights for women, ethnic, religious, and sexual minorities is ongoing in much of the non-Western world also supports the finding that not all legal persons have the same capacity to act.[25]

When it comes to juridical persons, it can be said that they almost always have fewer rights and obligations than legal persons. Even in the United States, where the corporations have nearly every constitutional right and obligation, the Supreme Court, albeit rarely, is known to rule to the contrary. To illustrate, in *Hale v Henkel*, the Court held that corporate employees cannot assert the privilege against self-incrimination on behalf of their employer.[26]

In 1948, the General Assembly of United Nations have also deliberated on the content of the notion of legal personhood. The Assembly asked the International Court of Justice whether the UN has the capacity to bring lawsuits against any sovereign state (which are like other juridical persons, associations of individual human beings that are brought together for the achievement of some common pursuits). The Court advised in the affirmative and emphasized the to the varying characteristics of legal persons:

> The subjects of law in any legal system are not necessarily identical in their nature or in the extent of their rights, and their nature depends upon the needs of the community. (…) [t]he Court has come to the conclusion that the [United Nations] Organization is an international person. That is not the same thing as saying that it is a State, which it certainly is not, or that its legal personality and rights and duties are the same as those of a State. (…)

The statement of the Court affirms that there are differences among the legal capabilities associated with different types of legal persons.[27] In that ruling, it was found that legal persons do not need to possess all the same rights and obligations.

---

[23] Peter M Asaro, 'Robots and Responsibility from A Legal Perspective', (IEEE International Conference on Robotics and Automation, Workshop on Roboethics, Rome, April 2007) 3.
[24] ibid.
[25] Mahnaz Afkhami, 'Gender Apartheid and the Discourse of Relativity of Rights in Muslim Societies' in Courtney W. Howland (ed), *Religious Fundamentalisms and the Human Rights of Women* (Palgrave Macmillan 1999) 72; Bruce Dunne, 'Power and Sexuality in the Middle East' in Christine Koggel (ed), *Moral Issues in Global Perspective - Volume 2: Human Diversity and Equality* (Broadview 2006) 271-273.
[26] *Hale v Henkel*, 201 U.S. 43 (1906) 43.
[27] *Reparation of Injuries Suffered in Service of the U.N.* (Advisory Opinion) [1949] I.C.J. 174, 178.

The divisibility of legal personhood necessitates the determination of rights to be conferred on a legal person after the legal system decides to recognize an entity as a legal person. Referring to the example given by Yoval Noah Hariri, 'though Toyota or Argentina has neither a body nor a mind, they are subject to international laws, they can own land and money, and they can sue and be sued in court.'[28] Legal persons do not even need to have the ability to exercise any capacity to defend their rights and perform their obligations on their own.

Bryson and others' third observation explains that once legal personhood is conferred to an entity, no matter how detailed the rights and obligations of that personhood are determined, circumstances might make it impossible for the entity interact with the legal system.[29] To put it another way, conferring legal personhood to an entity does not mean anything without the instrument through which the rights and obligations entailed by that legal personhood can be enforced. The instrument that is vital for a legal person seeking to protect its rights is standing – the right to assert or enforce legal rights in a judicial forum.[30]

Before asserting legal rights or enforcing duties, the legal person must exercise the legal standing it possesses. However, for the legal person to have standing, they do not need to have any actual capacity to assert rights. For instance, the Constitution of Ecuador recognises nature's right to integral respect for its existence and for the maintenance and regeneration of its life cycles, structure, functions and evolutionary processes'. [31] Even though nature has been granted that right as a matter of law, it obviously lacks the real capacities needed to protect that right from breaches. To effectuate these rights, the Constitution of Ecuador also states that 'all persons, communities, peoples and nations can call upon public authorities to enforce the rights of nature', and thereby gives standing to everyone in Ecuador to file lawsuits in favour of nature. [32] A 2011 case, was brought before the court by two individuals, involved the Vilcabamba River as the plaintiff. In that case, constitutional injunction was granted in favour of the Vilcabamba River and against the Provincial Government of Loja who attempted to conduct some environmentally-harmful project over the river.[33]

---

[28] Yuval Noah Harari, 'The Rise of The Useless Class' (*TED - Ideas*, 24 February 2017).
[29] Bryson, Diamantis, and Grant (n 12) 284.
[30] ibid.
[31] Constitution of Ecuador (2008), Arts 71-74.
[32] ibid. art 71
[33] Granted Constitutional Injunction 11121-2011-0010 (Provincial Court of Loja).

Just as rights mean nothing unless the legal system provides the standing to protect them, obligations are meaningless if there are no procedures to enforce them. That is, the legal capacity of an entity to bear liability and answer for its own breaches is complementary to its capacity to bring claims.[34] It is not possible to determine in advance how the robots will interact with other legal persons and institutions of the legal system regardless of whether they are granted legal personhood or any similar legal status. Thus, if robots are conferred personhood-like legal status, it is inevitable that every rule invoked on behalf of or against them will require novel developments in the law. This does not only mean that it is necessary to determine how the existing instruments apply to robots - both the standing of the robots against other entities and the standing of the other entities against the robots must also be taken into account when addressing the subject of new legal personhood.

## VII.2 Pragmatic Desirability of New Legal Status

The question of conferring some new legal status on robots is ultimately a pragmatic one - does granting robots with capabilities or putting them under that obligation help legal systems overcome any challenges. To answer this question, it is necessary to first identify the basic purposes of the legal systems. For Bryson and others, these are:

(1) to further the material interests of the legal persons [that the legal system] recognizes and
(2) to enforce as legal rights and obligations any sufficiently weighty moral rights and obligations, with the caveat that
(3) should equally weighty moral rights of two types of entity conflict, legal systems should give preference to the moral rights held by human beings. [35]

Turner, on the basis of this formulation, argues that the basic purposes of the legal systems are twofold:

(1) maintaining the integrity of the legal system as a whole and
(2) advancing the interests of human beings[36]

Here, it is emphasized explicitly that the term 'interests' refers to both moral and economic interests of individual human beings.

[34] Bryson, Diamantis, and Grant (n 12) 282.
[35] ibid. 285.
[36] Turner (n 15) ch 5, s 3, sub-s 3.1.

When the two formulations are compared, it appears that that the purpose of 'advancing interests of human beings' in Turner's formulation is broader than the imperative of giving 'preference to moral rights held by human beings' in Bryson and other's formulation.

Indeed, it can be suggested that, in certain circumstances, preferring the interests of legal persons over relevant human beings serves the interests of a larger number of human beings, taking into account the circumstances of contemporary economies. If this were not true, the separate legal personhood of corporations would not be such a fundamental institution for advanced economies. In this respect, it can be submitted that Turner's formulation is more realistic than that of Bryson et others. The present study, therefore, accepts Turner's formulation as the identification of the basic purposes of the legal systems. Below, the arguments as to whether it is desirable to grant legal personhood to robots are examined in relation to these two basic purposes.

As explored earlier in the present study, social robots have the potential to bring about numerous social and economic benefits to society. From a social standpoint, social robots can improve the quality of life for individuals, particularly those who may be isolated or lonely. By providing companionship and social interaction, social robots can help reduce social isolation and improve mental health. Furthermore, social robots can assist in healthcare settings by providing support to patients, reminding them to take their medication, and monitoring their health. In education, social robots can serve as teaching aids and facilitate learning for students, particularly those with learning disabilities. From an economic standpoint, social robots have the potential to increase productivity and efficiency in various industries, such as manufacturing, logistics, and customer service. By automating repetitive and menial tasks, social robots can free up human resources to focus on more complex and creative tasks. Moreover, social robots can improve customer experience and satisfaction by providing personalized and timely assistance. This can result in increased customer loyalty and revenue for businesses.[37] Appropriate regulation can balance the need for innovation in social robotics within a framework that addresses the pertinent safety, privacy and ethical issues and thereby help build trust and confidence among individuals and society.

[37] See Chapter I, Section I.1, 'Background of the Study'.

### VII.2.1 Arguments for New Legal Status

Determination of liability for robot behaviours -or filling in the accountability gap- is an important argument for introduction of the robotic legal personhood. In the last chapter of the present study, it was discussed that the growing autonomy of robots reduces the foreseeability of their behaviours and therefore makes it increasingly difficult to assign liability for the damages arising from these behaviours using existing general and alternative liability instrument. When the question of liability for robots is examined in relation to the basic purposes adopted above, it is found that these purposes pull in opposite directions. On the one hand, advancing the interests of human beings requires that a legal person be assigned liability for the damage caused by robot behaviours. On the other hand, however, in cases where the chain of causality between an existing legal person and an outcome is broken due to the robot autonomy, seeking an existing legal person to be held responsible for the robot behaviour may threaten the integrity of the legal system as a whole, and also damages the interests of legal persons that are being held liable.

In the words of Koops and others, 'for tomorrow's agents, however, applying and extending existing doctrines in these ways may stretch legal interpretation to the point of breaking.'[38] Conferring personhood-like legal status to robots could help overcome this tension between the requirements of the basic purposes of the legal systems. The introduction of robotic legal status will provide entities that would be held accountable for the damages and separate robots' independent activities from the free will of their manufacturers or masters, allowing certain safeguards for their liability to be introduced thereby advancing the interests of human beings.[39] Moreover the interposition of a new legal subject between an existing legal person and an under spouse outcome will allow the determination of liability in accordance with the chain of causation, one of the fundamental legal concepts. Thus, the integrity of the legal system as a whole would be maintained. [40]

---

[38] BJ Koops, M Hildebrandt, and DO Jaquet-Chiffelle, 'Bridging the Accountability Gap: Rights for New Entities in the Information Society?' (2010) 11 Minnesota Journal of Law Science & Technology 497.
See Chapter V, Section V.2, 'Tort Law', for the discussion of the relevant tort concept of *novus actus interveniens.* The doctrine of *novus actus* interveniens is not thought in itself to make the law unable to advance the interests of human beings. The doctrine generally identifies someone else who can indeed be sued as the causa causans, whereas for damage done by a robot, absent personhood, there is no one else to be sued.
[39] Curtis EA Karnow, 'Liability for Distributed Artificial Intelligences' (1996) 11 Berkeley Technology Law Journal 147, 183.
[40] ibid 186.

Another argument for the robotic legal status is concerned with the encouragement of innovation and economic growth. Inspired by the contribution of corporate personhood to the rapid technological and economic growth that was achieved globally in the course of past few centuries, this argument asserts that the structure of legal personhood of corporations is not sufficient to address the economic and social consequences related to liability for robots. Corporate legal personhood stipulates that the rights and obligations of corporations are separate from those of their shareholders and since the creditors only have recourse to corporations' own assets, it limits the risk taken by shareholders to the amount of their investments.[41] In other words, the limited liability provided by corporate legal personhood immunizes shareholders' personal assets from the risk that the corporation fails and cannot pay its debts. Consequently, corporate legal personhood is assumed to encourage enterprise that might prove generally beneficial as it allows for larger investments in more risky fields, such as bio-medical research, banking and finance, or intercontinental trade.[42]

The shareholders of the robot manufacturers, of course, already benefit from the limited liability afforded by their corporate identities. However, even though the personal assets of the shareholders are immunized, robot manufacturers' own assets are exposed to all claims. Considering the distinctive characteristics of robots and how they differ from other products, this means that the robot manufacturers may have to assume liability for unforeseeable damages caused by their products that -by definition- have their own decision-making capacities. [43] In response to the product liability claims, European robot manufacturers can rely on the 'state of the art' defence, set out in the Product Liability Directive (Directive 85/37/EEC), which exempts a producer from liability where it can show that, in light of the state of scientific and technical knowledge at the time it put the product into circulation, it could not have discovered the particular defect. However, this defence requires the disclosure of full details of the facts relating to the products, and even if successful, it is a potentially detrimental process as these products rely heavily on being technological innovations and disclosure of the all the facts relating to them may result in making the details about such products public and therefore making the aspects that are not protected by patents law public.

---

[41] Paul G Mahoney, 'Contract or Concession-An Essay on the History of Corporate Law' (1999) 34 Georgia Law Review 873, 877; Turner (n 15) ch 5, s 3, sub-s 3.3.
[42] Mahoney (n 40) 878.
[43] Gunther Teubner, 'Rights of Non-Human beings? Electronic Agents and Animals as New Actors in Politics and Law' (2006) 33 Journal of Law and Society' 497, 501.

Moreover, it is equally possible that this defence would not apply to robots, as the possibility of unforeseeable behaviour itself is known to designers. Consequently, it is submitted that, in most circumstances, corporate personhood is indeed not sufficient to encourage enterprise in robotics, since robots would not be manufactured, if their manufacturers had to bear their risks, presumably.[44] It is therefore suggested that there is a need for a legal that takes into account the characteristics of the robots and imposes, if necessary new forms of secondary liability for robots on robot manufacturers and masters. Since robots are able to take decisions and act upon these decisions on their own, the most sensible solution seems to be to impose liability on robots for the impairments in their own decision-making processes, or faults. In the proposal detailed in the next chapter, it is suggested that robots should be accorded a dependent legal status, and, like companies, should be entrusted with a certain patrimony by those for whose benefit they act.

### VII.2.2 Arguments Against New Legal Status

One of the most common arguments against robotic legal status is known as the 'Android Fallacy' argument. The 'Android Fallacy' is explained by Richards and Smart as follows:

> [O]ne particularly seductive metaphor for robots should be rejected at all costs, the idea that robots are 'just like people' and that there is a meaningful difference between humanoid and non-humanoid robots. We call this idea 'the Android Fallacy.'[45]

Their basis for rejecting this metaphor is that the tendency to overly anthropomorphise a humanoid robot may lead to holding their manufacturers less liable, compared to those of less anthropomorphic robots who have the same degree of autonomy. It should be noted that Richards and Smarts' assessments are not unrealistic. Indeed, if one is driving a car and it fails to respond when one turns the steering wheel, it is obviously the manufacturer's fault, but if an android is driving the car, and its hands slip on the wheel when attempting to make a turn, the android manufacturer would probably be measured up against different, and probably much more relaxed standards compared to very high standards of the automotive industry. The psychological experiments regarding human beings' evaluation of the wrongness of robots' behaviours supports this statement.

---

[44] Turner (n 15) ch 5, s 3, sub-s 3.3.

[45] Neil M Richards and William D Smart, 'How should the law think about robots?' in Ryan Calo, A Michael Froomkin and Ian Kerr (eds), *Robot Law* (Edward Elgar 2016) 4.

However, it must also be emphasised that objection to robotic legal personhood on the basis of 'Android Fallacy' does not make sense, since it seemingly 'rests on a mistaken conflation of the idea of [legal] personality [with that of] humanity.'[46] In April 2018, a group of artificial intelligence experts from European countries made this mistake, by stating that the legal personality of the robot was ethically and legally inappropriate:

> A legal status for a robot can't derive from the Natural Person model, since the robot would then hold human rights, such as the right to dignity, the right to its integrity, the right to remuneration or the right to citizenship, thus directly confronting the Human rights (…)The legal status for a robot can't derive from the Legal Entity model, since it implies the existence of human persons behind the legal person to represent and direct it. And this is not the case for a robot.[47]

From a technical perspective, this statement is based on an overvaluation of the actual capabilities of even the most advanced robots, a superficial understanding of unpredictability and self-learning capacities.

Still, a significant question in determining whether robots can be accorded legal status is whether being a living human being is a necessary or sufficient condition for being a legal person. As discussed elsewhere in this chapter, neither condition was obtained either historically or in modern-day legal systems.

When considering whether to confer an entity legal personhood, it does not matter whether the entity is able to recognize the consequences of its actions. The legal personhood of human beings who are not aware of their legal personhood (e.g. minor children, permanently mentally impaired persons) is still recognized even though they can only exercise their rights and fulfil their obligations through their representatives. Furthermore, in some legal systems, entities such as rivers, mountains or nature are given legal personhood. [48] Of course, these entities cannot be expected to be conscious of their personhood. As explained earlier in this chapter, the inherent characteristic of an entity does not determine its legal personhood. Therefore, it can be concluded that this argument is not substantial at all. Another argument against granting legal status to robots is that this new status will be exploited by natural persons to pursue their selfish purposes.[49]

---

[46] Turner (n 15) ch 5, s 3, sub-s 3.6.1.

[47] 'Open Letter To The European Commission Artificial Intelligence And Robotics' <http://www.robotics-openletter.eu> accessed 11 September 2019.

[48] For instance, see Bryant Rousseau, 'In New Zealand, Lands and Rivers Can Be People (Legally Speaking)' (*The New York Times*, 13 July2016) Kathleen Calderwood, 'Why New Zealand Is Granting A River The Same Rights As A Citizen' (*Australian Broadcasting* Corporation, 6 September 2016) same-rights-as-citizens/7816456> accessed 11 September 2019; see also Michael Safi, 'Ganges And Yamuna Rivers Granted Same Legal Rights As Human Beings' (*The Guardian*, 21 March 2017).

[49] Bryson, Diamantis, and Grant (n 12) 288.

This argument is explained as follows:

> There is nothing objectionable in itself about actors pursuing selfish ends through law. A well-balanced legal system, however, considers the impact of changes to the rules on the system as a whole, particularly so far as the legal rights of legal persons are concerned. We take the main case of the abuse of legal personality to be this: natural persons using an artificial person to shield themselves from the consequences of their conduct. Recognition of robot legal personhood could present unscrupulous actors with such "liability management" opportunities. [50]

This argument does point to some significant threats, but it nonetheless implies that 'any instance of separate legal personality will be used by human beings on a habitual basis.'[51] That argument is incorrect. On the contrary, throughout history, examples of separate legal personhood, such as corporations, have assumed significant economic roles by allowing human beings to make risky or costly investments without putting all their assets on the line.[52]

The same argument can be asserted against the limited liability of companies, but even the most ardent opponents of the robotic legal status would not argue that all companies should be dissolved. Moreover, in cases where corporate personality is abused, it is possible to hold shareholders personally liable disregarding the company's limited liability by 'piercing the corporate veil'.[53] In principle, the move of 'piercing the corporate veil' can also be adapted to robotic legal personhood, thereby making the argument untenable.

To sum up, it can be determined that the arguments for robotic legal status do not only show that introduction of a new legal personality would further the interests of human beings–but also demonstrate that it may be a viable instrument that would maintain the integrity of the legal systems in face of the challenges posed by increasingly autonomous robots.

[50] ibid 289.
[51] Turner (n 15) ch 5, s 3, sub-s 3.6.2.
[52] Mahoney (n 40) 889.
[53] Robert B Thompson, 'Piercing the Corporate Veil: An Empirical Study' (1999) 76 Cornell Law Review 1036.

## Summary

In this chapter, the nature of legal personhood in the law's ontology and its historical development is examined, and the question of whether the granting of some new legal status to robots would threaten the systematic coherence of the legal order is discussed. In addition, arguments on whether the robotic legal status is pragmatically desirable are considered and the following conclusions have been reached.

The notion of 'legal person' and the word 'person' in everyday language have different meanings. 'Legal person' is a concept specific to the law's ontology and refers to the rights and obligations bearing units. The notion is a creation of law, and it is purely coincidental that any living human individual in the present-day ('person' in everyday language) qualifies as a legal person. Legal orders were able to deny 'legal personhood' to some persons in the past, and they have recognized non-persons -things that are not living human individuals- as 'legal persons' both historically and currently.

'Legal personhood' is a form of legal status that is essentially a bundle of rights and obligations. Not all legal persons have the same rights and obligations since these can come together in different combinations. This implies that the 'legal personhood' is divisible, and the rights and obligations a legal person has can be modified to accommodate the ever-changing social facts.

In addition, being a fiction, the entity accorded 'legal personhood' does not need to have physical existence ('body') or the ability to be conscious of their selves and their decisions ('mind'). Almost all non-human legal persons use their rights and fulfil their obligations through their agents.

Accordingly, as long as the rights and obligations it entails are clearly identified, as well as the boundaries of liability rules about it, there are no obvious conceptual impediments to granting robots legal personhood. Robotic legal status, essentially, is a question of pragmatic desirability.

It is assessed that a well-designed robotic legal status would solve the uncertainty concerning liability for robots, and would encourage the actors in the robotics industry to become generally beneficial enterprises who research and develop robots for more risky (and perhaps less profitable) applications since the robotic legal status would limit the potential liability of the robot manufacturers (further limit, in addition to the limited liability afforded by corporate legal personhood) and robot masters. As such, robotic legal personhood also seems to be pragmatically desirable.

The question about the rights and obligations of the proposed robotic legal status, as well as the liability instruments relating to it, are not answered in this chapter. These questions, along with those about how the proposed personhood would help overcome the legal challenges of sociable robots, are going to be addressed in the next chapter, which details the principles of this new legal status.

# Chapter VIII: Future of Robot Law

> Then you don't remember a world without robots. There was a time when humanity faced the universe alone and without a friend. Now he has creatures to help him; stronger creatures than himself, more faithful, more useful, and absolutely devoted to him. Mankind is no longer alone.
>
> —Isaac Asimov, *I, Robot*[1]

## Introduction

The present study does not only examine the existing law, *de lega lata*. Intended to be both constructive and critical, the research performed in the study also glances to the future and investigates how the law should be modified to keep fulfilling its function as the guardian of social order as humankind enters the age of social robots. In the age of social robots, robots are no longer restricted to behind-the-scenes operations.

With the advent of social robots, human beings are starting to share the social interaction spheres with non-human entities for the first time. Even though social robots are functional artefacts with no desires, wants, or needs that motivate the activities they perform, they are equipped with the capacity to generate sophisticated social interactions. They can make and execute decisions without human intervention, and also display emotional sentiments, engage in relationships, and develop individual styles.

Because of the distinctive characteristics of social robots (relative autonomy and social agency), applying the existing legal instruments to the legal issues surrounding human-social interactions would bring outcomes contrary to the rule of law, such as a breach of legal certainty. As such, it is submitted that the existing legal instruments must be modified for legal systems to produce solutions that align with the underlying legal principles. Therefore, the present study complements the analysis of the existing law through the proposal for a new unified legal framework intended to allow legal systems to resolve legal issues regarding social robots adequately. The necessity and feasibility of the proposed improvements were already explored in the previous chapters of the present study. The following final chapter, therefore, puts forward the unified legal framework proposal and offers a vision of what the law should be, and outlines the *de lege ferenda.* In that regard, the chapter is organised into two sections.

[1] Isaac Asimov, *I, Robot* (first published 1967, HarperVoyager 2018) 10-11.

The first section of the chapter, consisting of three divisions, summarises the present study's findings regarding the *lex lata*. The first division discusses the distinctive characteristics of social robots and approaches these characteristics as the causes of the legal issues surrounding human-robot social interactions. The second division describes the legal issues that the study addresses and highlights the conceptual connections between the legal issues and the distinctive characteristics of social robots. The final division of the first section elaborates upon the need for a new unified legal framework. The reviews several proposals already put forward in the literature, assesses the compatibility of the existing proposals with the distinctive characteristics of social robots, and examines how the existing proposals address the balance between two competing social interests, namely, economic efficiency and legal security.

The second section, incorporating two further sub-sections, elaborates on the present study's proposal for a new unified legal framework to resolve the legal issues regarding social robots adequately. The first sub-section focuses on the challenges in criminal law and advocates that the imposition of criminal liability on robots would bring at least some affirmative benefits of criminal punishment without violating any of the fundamental principles of criminal law. The first sub-section also addresses the legal issues that may arise in connection with robot anthropomorphism and proposes that abusive behaviours toward robots should be recognised as criminal offences, especially where the robot at hand is designed to invoke anthropomorphic responses. The second sub-section focuses on the challenges in private law and suggests a new legal status, 'quasi-personhood' for robots, as part of the new unified legal framework proposed by the present study. The quasi-personhood status grants robots the legal capacity to act (perform legally significant actions) without the juridical capacity (capacity to have rights and obligations). Liability questions stemming from the lack of juridical capacity of the quasi-personhood status are addressed by granting quasi-persons access to their masters' (owners/users) juridical capacities. It is suggested that the extent of that access can be determined according to the nature of robots' actions, the types of injury, as well as by the masters' roles in bringing about the damages related stemming from robots' actions (whether they command, induce or facilitate these actions), and the outcomes of robots' actions over the assets of their masters'.

## VIII.1 Summary of the Findings

The present study maintains that the legal issues that may arise in connection with the independent activities of social robots cannot be answered through the existing legal instruments as these instruments do not recognise the distinctive characteristics of social robots. Indeed, previous chapters of the present study establish that applying any existing legal instrument to the legal issues regarding social robots would either have no effect on these issues or create unfair or inequitable situations for one or more parties involved. As such, it follows that any proposal that offers adequate resolutions to the legal issues surrounding human-robot social interactions must take the distinctive characteristics of social robots into account and focus on the role of these characteristics in bringing about the examined legal issues.

### VIII.1.1 Causes of Legal Issues

#### VIII.1.1.a Robot Autonomy

The abilities and skills that distinguish robots from other functional artefacts can be conceptually summarised under the characteristic of relative autonomy. Though autonomy should be appraised as a spectrum instead of being approached from a binary perspective, it is nonetheless deemed necessary to adopt a definition of autonomy to assess whether any given functional artefact is indeed a robot or not. Inspired by the computer science literature, the present study acknowledges that any functional artefact that can perform at least one behaviour under 'the sense- think-act' paradigm qualifies as autonomous and can be regarded as a robot.[2]

When operating according to 'the sense-think-act' paradigm, first, the artefact must be able to perceive and interpret its surroundings. In other words, the ability to 'sense' refers to the entity's ability to incorporate changes in its external environment into its functioning.[3] As such, the ability to sense is the conceptual prerequisite for any reaction to outside stimuli. The 'sensitivity' of photocells to movements in the surroundings they are placed in and the 'sensitivity' of thermometers to temperature can be evaluated as the simplest examples of the presence of the ability to sense.

---

[2] George A Bekey, *Autonomous Robots: From Biological Inspiration to Implementation and Control* (MIT 2005) 2.
[3] David J Gunkel, *Robot Rights* (Kindle edn, MIT 2018) ch 1, s 1.1, sub-s 1.1.2.

Secondly, to qualify as a robot, any functional artefact should also be able to process information and plan its behaviours - or have the ability to think. The ability to 'think' is perhaps named so because of its similarity to the human ability to think; that is, the cognitive skill that allows human beings to perceive and mentally represent the world and engage in problem-solving and decision-making activities.[4] However, it must be emphasised that in the present day, no robot can think that even approaches the complexity and flexibility of human thought. Thus, concerning robots, the ability to 'think' is better understood as their capability to process any prior knowledge and sensory information through their algorithms. Decisions emerge from these analytical processes as action plans that specify which of the possible behaviours would be performed by robots. The adjective 'autonomous', often used to distinguish robots from other artefacts throughout the study, denotes the presence of 'thinking' abilities. In other words, if any given functional artefact qualifies as autonomous, it indicates that the functional artefact in question can make decisions based on its evaluation processes without receiving any external direction or guidance. Varying degrees of autonomy can be understood by assessing how advanced the given robot's ability to 'think'. For example, robots that can consider their past decisions when evaluating new information that can learn from their experiences can be said to be 'more' autonomous than the others. [5] The presence of artificial intelligence, especially in the form of some ability to develop and improve one's own algorithms, indicates higher levels of autonomy as well. Indeed, the highest possible level of autonomy would indicate that the entity can modify all aspects of its algorithm, including its core code and specified functions, and unlike any of the robots today, it would be able to determine what interests to pursue.

Thirdly, the ability to 'act' refers to the power to carry out decisions produced through thinking processes. In the present study, robots are sometimes also referred as agents. It is submitted that the identification of robots as 'agents' in this manner is justified because of their ability to act, that is, to carry out the decisions they made without any external direction or guidance.[6]

---

[4] Joachim Hertzberg, Raja Chatila, 'AI Reasoning Methods for Robotics' in Bruno Siciliano and Oussama Khatib (eds), *Handbook of Robotics* (Springer 2008) 210; Marc Hanheide and others, 'Robot task planning and explanation in open and uncertain worlds' (2017) 247 Artificial Intelligence 119.
[5] Stuart J Russell and Peter Norvig, *Artificial Intelligence: A Modern Approach* (Prentice Hall 1995) 35.
[6] Samir Chopra and Laurence F White, *A Legal Theory for Robots* (UM 2011) 9.

In the literature, the ability to act is often understood as the initiation of responses to changes in robots' physical environments.[7] Nonetheless, most of the scholarly literature on the topic is based on pre-Internet observations. The bulk of the literature thus overlooks that the Internet has become a platform where most of the business, commerce, education, and entertainment occurs.[8] In all three legal systems chosen for comparison for the purposes of the present study, it is recognised that actions performed over the Internet can impact one's personal and property rights, even though these actions do not always impact some physical environment. Hence, it is no longer feasible to argue that the ability to 'act' must relate to some physical environment and require the entity to have some physical embodiment to qualify as a robot. Instead, the present study holds that ability to 'act' is present if the entity in question can affect changes in its environment, regardless of whether that environment is physical or digital. It must be noted that regardless of the environment they are situated in, robots operate according to the 'sense-think-act' paradigm to carry out the functions determined by their makers. That is to say that, today, no robot can change its pre-set goals or change its designated role. However, even though they, too, cannot change their predetermined functionalities, social robots can enter into unique social interactions with human beings and thus appear to be able to change some of their pre-set objectives or adopt new objectives.

### VIII.1.1.b Robot Anthropomorphism

Anthropomorphism is the attribution of human traits and human internal states -such as feelings and intentions- to non-human entities.[9] It is not, by any means, some new phenomenon that is observed for the first time in connection with the emergence of robots. On the contrary, is one of the innate tendencies of the human mind. The scientific literature regards anthropomorphism as an automatic cognitive response that evolved to enable the human species' survival in environments where they were incorporated in the usual diets of large predators.[10]

---

[7] Roger Clarke, 'Asimov's Laws of Robotics: Implications for Information Technology-Part I' (1993) 26 Computer 53, 54.
[8] Manuel Castells, 'The Impact of the Internet on Society: A Global Perspective' in *19 Key Essays on How Internet is Changing Our Lives* (1st edn, BBVA 2014) 132; Zaryn Dentzel, 'How The Internet Has Changed Everyday Life', in *19 Key Essays on How Internet is Changing Our Lives* (1st edn, BBVA 2014) 241.
[9] Esmeralda G Urquiza-Haas and Kurt Kotrschal, 'The Mind Behind Anthropomorphic Thinking: Attribution of Mental States to Other Species' (2015) 109 Animal Behaviour 167, 168.
[10] ibid 172.

The strength and extent of anthropomorphism correlate to the similarity of non-human entities to human beings; for example, humans are likely to anthropomorphise vertebrates much more than invertebrates. The scientific literature has also established that the parts of the human brain activated when this cognitive response is invoked are the same ones activated during social interactions to determine other parties' internal states. It is established above that the characteristic of relative autonomy resembles the human capacity of free will since it denotes the robot's ability to choose between different courses of action. Until the emergence of robots, no non-human entity had any capacity that had even some resemblance to free will, qualifying it as one of the most exclusive characteristics of human beings.[11] It follows that any non-human entity with the ability to choose between different possible actions would be viewed to have some significant similarity to human beings. In other words, robot autonomy, even when not accompanied by any other human-like features, is sufficient to invoke some anthropomorphic responses.[12]

In addition to robot autonomy, social robots have some other capabilities that allow them to undertake roles as social interaction partners. These 'social' skills, for example the ability to process natural language and anthropomorphic audio-visual features, are intended to strengthen the automatic cognitive response of anthropomorphism. Some robots can already imitate certain responses specific to the organic creatures, such as pain. Robot anthropomorphism is especially relevant in the context of criminal law, where attention must be paid to its potential effects on the society's sensitivity to behaviours that cause physical or emotional pain and discomfort to their victims.

### VIII.1.2 Legal Issues

The function of law is often cited as maintaining the sense of justice in society by ensuring the sustainability of the interactions between the members of society, thereby upholding general peace and social order. [13]

---

[11] The scientific and anecdotal evidence that is examined in the previous chapters of the present study support this statement, such as emotional attachments formed with autonomous vacuum cleaners, see Chapter IV, Section IV.2.1, 'Loss of Genuine Human Connections'.

[12] Matthias Scheutz, 'The Inherent Dangers of Unidirectional Emotional Bonds between Human beings and Social Robots' in Patrick Lin, Keith Abney and George A Bekey (eds), *Robot Ethics: The Ethical and Social Implications of Robotics* (MIT 2014) 213.

[13] Dana Burchardt, 'The Functions of Law and Their Challenges: The Differentiated Functionality of International Law' (2019) 20 German Law Journal 409, 414.

Until the emergence of robots, this function was achieved by regulating human behaviours that affect others or social behaviours. Indeed, private, and criminal law domains prescribe appropriate human behaviours in various social contexts. Legal sanctions are invariably aimed at restricting or enforcing certain human behaviours. However, the proliferation of robots and especially the advent of robots that can act as social interaction partners challenges the law's ability to fulfil its functions while limiting itself to regulating human behaviours. Accordingly, it is found that the domains of law that focus on the regulation of interpersonal relations are not conducive to resolving disputes that may arise due to the presence of robots as interaction partners.

More common in continental legal systems, the distinction between private and public law originates from Roman Law. According to the Roman jurist Ulpian, public law is concerned with the constitution of the State and the functions of the government and administration, and private law is concerned with the interests of individuals.[14] For the present study, private law is understood as the entirety of legal principles and rules that regulate the relations between two equal parties.[15] Public law, on the other hand, incorporates the body of laws that govern the State's organisation, as well as those that delineate the relationship between the State and individuals or organisations.[16] In public law, the State is always in the superior position 'in light of its instruments of power'.[17] Thus, in contrast with private law, the parties concerned in public law are not equal to one another.

### VIII.1.2.a Private Law

The primary distinction between two domains of private law, contract law and tort law, is the difference between the nature of rights they are designed to vindicate. In contract law, the right in question is always relative, meaning that it corresponds to the contractual counterparty's duty and can only be claimed against them, whereas in tort law, the infringed right is always absolute, entailing duties for an undetermined number of individuals.

---

[14] George Long, 'Jus', *A Dictionary of Greek and Roman Antiquities* (John Murray 1875) 658; Alan Watson (ed), *The Digest of Justinian* (University of Pennsylvania Press 2009) vol 1, 72.
[15] Randy E Barnett, 'Four Senses of the Public Law-Private Law Distinction' (1986) 9 Harvard Journal of Law & Public Policy 267, 270.
[16] ibid 271.
[17] Anne Peters, *Beyond Human Rights: The Legal Status of the Individual in International Law* (CUP 2014) 112.

*Contract Law*

Contract law is based on the freedom of contract, which, in turn, follows from the principle of free will. As a general rule, contracting parties are free to determine whether and with whom to enter into a contract and how the contents of that contract should be. The freedom of contract is supplemented with the principle of *pacta sunt servanda* (sanctity of contract), indicating that the parties to contracts must perform their respective obligations under the contract.[18] Since it is impossible to grasp the 'essence' of human will or ascertain what happens inside someone's mind, the selected legal systems instead focus on how the will is expressed and communicated, especially when determining whether there is a valid contract.

There are no legal instruments in any of the selected legal systems that recognise that robots can autonomously decide to enter contracts on behalf of their masters. Nonetheless, the growing autonomy of robots and the increasing number of roles that they can fulfil suggest that, in the near future, they are going to be entering into contracts on their masters' behalf with even less interference from these masters.[19] In other words, it must be acknowledged that robots can analyse external factors with their integrated cognitive instruments, determine their possible courses of action, and then conclude contracts should they assess that doing so is the optimal course of action to help achieve their objectives. Robots' behaviours, in these contexts, can no longer be regarded as communications of their masters' free will. Indeed, though performed on behalf of and to the benefit of their masters, such behaviours are decided by robots themselves.

When one legal person negotiates and enters into some contract on behalf of another legal person, their role in forming the contract is recognised and regulated under the rules that address 'agency'. These rules stipulate specific conditions under which the agent's behaviours will be binding on the principal to protect the interests of both the principal and the third party.

However, the same protection is not afforded in cases where 'robots' are acting as agents, exposing the legal systems to re-emergence the problems solved with the rules that address 'agency'. For example, some robot may be given the order to purchase an antique chair in an auction, offering an amount well above the amount the owner is willing to pay to perform this task.

---

[18] Hans Wehberg, 'Pacta Sunt Servanda' (1959) 53 The American Journal of International Law 775.
[19] Jacob Turner, *Robot Rules* (Kindle edn, Palgrave Macmillan 2018) ch 3, s 2.5, sub-s 2.5.1.

The same can happen during legal transactions in financial markets, where the timing of the transactions and, consequently, the speed and efficiency of decision-making instruments are vital. In such scenarios, as robots' masters cannot benefit from the law of 'agency', all contracts concluded by robots are to be deemed as concluded by their masters.

Consequently, simple errors made by robots during the execution of their contracting tasks may result in the financial destruction of their masters even though they would not have agreed to such contracts. Here, the imposition of contractual liability on the masters without their consent also infringes the principle of party autonomy. For example, in cases where social robots are tasked with caring for people who cannot take care of themselves, such as the elderly or disabled, fulfilling their tasks may require them to make certain transactions on behalf of and to the benefit of their owners. One such robot may dispose of all assets of the person it is caring for because of some unpredictable linguistic error is generated when processing some newspaper article, and none of the selected legal systems provides any recourse to that party.

*Tort Law*

Tort law addresses the issue of how to remedy the damages to one's person or property due to civil wrongs other than breaches of contract and has the central purpose of compensating the victim through the redistribution of damages.[20] The redistribution of damages, in general, is based on the principle of fault. The principle of fault stipulates that one should be liable for any damages for which they are blameworthy.

If one can reasonably contemplate the damages that their behaviours may cause, they are expected to avoid performing such behaviours or take the necessary measures to prevent the damage from occurring.[21] If one fails to avoid performing such behaviours either because of one's intent (desire to bring about harmful consequences) or negligence (failure to exercise necessary care), one will be liable for any resultant injury.

Therefore, it is assessed that tort law is also based on the principle of free will, as both types of 'fault', in essence, are concerned with the unlawful use of one's free will. However, there are some exceptions to the principle of fault.

[20] Percy H Winfield, *The Province of the Law of Tort* (CUP 1931) 32.
[21] Maruerite E Gerstner, 'Liability Issues with Artificial Intelligence Software' (1993) 33 Santa Clara Law Review 239, 241-242.

Under the strict liability law, if one's behaviours are inherently risky, one may be held liable for injuries caused by such behaviours even in the absence of fault. Examples of strict liability include product liability, abnormally dangerous activities, and ownership of wild animals.[22]

In tort law systems, the general liability rules require the defendant to have used their free will to cause undesirable outcomes. The characteristic of relative autonomy makes it impossible to establish *fault* in tort law. [23] In that regard, most of the proposals in this category argue that the autonomy of robots, on its own, sufficiently justifies diverging from the general rules for liability.[24]

Even in strict liability torts, one is held liable, ultimately, for using one's free will to perform some inherently risky behaviour. Under no circumstances can one be held liable for behaviour one did not freely decide to perform. In other words, it can be said that tort law is also shaped around the principle of free will. Even with robots' inclusion in social interaction spheres as interaction partners, they would still be manufactured according to high safety standards, and the behaviours they decide to perform with their decision-making instruments would only rarely cause harm. Since robots are not inherently risky, and the damages caused by them will often be unforeseeable to both their manufacturers and masters, it is uncertain who should be liable for the damages caused by robots who acted after analysing external factors with their independent decision-making processes. For that reason, the present study submits that the existing tort law regimes are not ready for the widespread use of robots.

### VIII.1.2.b Criminal Law

In every legal system that is examined for the purposes of the present study, it is observed that legal entities are subjected criminal punishment when any of their conduct threatens the core moral values of society. That is done independently from any remedial obligations being awarded to the party harmed by such conduct.

[22] William Binchy, 'Tort Law In Ireland: A Half-Century Review' (2016) 54 Irish Jurist 199, 210.
[23] Andrea Bertolini, 'Robots as Products: The Case for a Realistic Analysis of Robotic Applications and Liability Rules' (2013) 5 Law, Innovation and Technology 214, 236
[24] Curtis E A Karnow, 'The Application of Traditional Tort Theory to Embodied Machine Intelligence' in Ryan Calo, A Michael Froomkin, and Ian Kerr (eds), *Robot Law* (Edward Elgar 2016) 155

Indeed, society's right to inflict punishment is grounded on the idea that some harm is done to society as a whole, in addition to the damages suffered by the victim. [25] There are several reasons why punishment has been considered legitimate throughout the centuries: retribution, just deserts, rehabilitation, deterrence, and so on. In criminal law, offences are classically composed of the *mens rea* (the guilty mind) and the *actus reus* (the guilty act). [26] The mental requirements of offences differ between legal systems and even between offences themselves. Generally, the *mens rea* required for guilt goes beyond having foreseen the consequences and requires having intended, desired, or willed such consequences.

When it comes to identifying problems related to criminal law, the present study emphasises that robots can be at both ends of criminal behaviours. As such, the questions arising in criminal law do not merely relate to determining who should be imposed criminal liability for offensive behaviours that robots perform. In light of the socio-legal consequences of robot anthropomorphism, the issue of criminal punishment for offensive behaviours against robots must also be explored.

### *Robots as Offenders*

Any behaviour that is subjected to criminal punishment must embody the objective and subjective elements required for criminal liability. The objective element is determined by how the behaviour is conducted; it assesses whether the series of actions performed as part of the behaviour in question constitute a crime under the relevant criminal law rules. In short, the objective element requires the behaviour to correspond to a crime in terms of its subject.

The concept of free will is central to the imposition of criminal liability in relation to its subjective or moral element. It is only exceptionally possible for someone to be subjected to punishment for behaviours performed without any criminal intent or gross negligence where the crime is identified as a strict liability offence.[27] Considering that criminal punishment, especially in the form of incarceration, is one of the most detrimental infringements that can be on one's freedom, the tendency in the selected legal systems is only to impose such punishment when the mind is 'guilty'.

---

[25] Hans Kelsen, *General Theory of Law and State* (Transaction 2006) 86.
[26] Kenneth W Simons, "The Crime/Tort Distinction: Legal Doctrine and Normative Perspectives "(2008) 17 Widener Law Journal 719, 729.
[27] David Gerber, 'Strict Criminal Liability and Justice' (1974) 60 (4) Archives for Philosophy of Law and Social Philosophy 513, 513.

The characteristic of relative autonomy indicates that robots are expected to be able to analyse their environments using their algorithms and without any external guidance. The decisions made and behaviours performed after such analytical processes will, from time to time, inevitably be unpredictable to their manufacturers and masters alike. In cases where such unpredictable behaviours are ostensibly criminal, the imposition of criminal punishment on robots' manufacturers or masters does not make much sense since their actions would satisfy the subjective element of criminal liability. After all, it would not be possible to say that any of these parties used their free will in such a manner to cause unlawful results. Using the instruments available in criminal laws of selected legal systems, the only way to impose criminal punishment on robots' manufacturers and owners is to expand the number of cases of strict liability offences. However, since strict criminal liability is incompatible with the basic tenets of criminal law, such offences are either being phased out or are already obsolete in all selected legal systems. [28]

Consequently, it remains uncertain who should bear the criminal liability for robot behaviour. If the decision to manufacture or use that robot is deemed to satisfy the subjective element of criminal liability, and some criminal punishment is imposed on either the manufacturer or the master, that punishment would breach the principle of individual criminal liability and therefore be unjustified.

Accordingly, the present study notes that criminal law instruments in selected legal systems fail to acknowledge robots' ability to make decisions based on their independent analytical instruments, even though that ability in question is remarkably similar to the human beings' ability to make decisions based on their on mental processes.

*Robots as Victims*

One of the purposes of criminal law is to express society's condemnation of certain behaviours and help maintain the general sensitivity against these offensive behaviours. To that end, in selected legal systems, criminal liability regimes often focus on offenders and their conduct rather than the internal states of victims. When performed against any robot, including those that can display signs of physical or emotional distress, no offensive behaviour is subjected to criminal punishment in any selected legal system.

[28] Peter Westen, 'Two Rules of Legality in Criminal Law' (2007) 26 Law and Philosophy 229, 232.

It is assessed that the failure to proscribe offensive behaviours against robots is likely to weaken society's sensitivity against such offensive behaviours towards human beings.[29] As such, extending certain protections to robots -regardless of whoever has property rights over them- might be necessary to communicate society's repudiation of such offensive behaviours in general.

### VIII.1.3 Existing Proposals

The present study submits that shortcomings of existing legal instruments cannot be overcome without some recognition of relative autonomy by the legal systems. However, there are some significant caveats to the recognition of the characteristic of autonomy.

One such caveat stipulates that since robots are -however advanced- tools designed to fulfil specific functions, some of their behaviours should still be predictable to their manufacturers or masters. After all, behaviours of robots are determined by action plans determined by their algorithms (created by their manufacturers), in accordance with the commands entered (by their masters). Manufacturers and masters should not be absolved from liability where their robots are programmed or commanded to undertake tasks that may not be unlawful themselves but whose performance may foreseeably involve commission of unlawful behaviours.

The second caveat emphasises that the recognition of robot autonomy might provide disproportionate protection to robots' manufacturers and masters. The recognition of robot autonomy is expected to solve some challenges, but it is feared that it may result in some new challenges that have been prevented from appearing thanks to the modern legal systems' treatment of robots as tools. In order to avert new potential challenges, legal systems should recognise the dual conceptual identity of robots. In other words, in solving the problems arising from the increasing presence of robots in society, in some cases, robots should continue to be treated as extensions of the free will of their manufacturers and masters. In other cases, where robots' behaviours are unforeseeable to their manufacturers and owners, approaching robots as entities with their own free will seems more logical. In fact, from a legal standpoint, it is not unheard of for an entity to have more than one conceptual identity.

---

[29] Kate Darling, 'Who's Johnny? Anthropomorphic Framing in Human–Robot Interaction, Integration, and Policy' in Patrick Lin, Ryan Jenkins and Keith Abney (eds), *Robot Ethics 2.0: From Autonomous Cars to Artificial Intelligence* (OUP 2017) 175.

Consider the concept of *longa manus* used in Italian criminal law doctrine. The concept refers to the role of someone who was forced to commit a crime by material coercion. Here, the person who is made to perform the criminal behaviour in the position of the outstretched hand (longa manus), or the tool of the coercer.

Similarly, when a robot is programmed or controlled with the purpose to bring about an illegal result and, it cannot be said that its behaviour emanates from the robot's free will. In contrast, the robot would be in the position of the extended hand of their manufacturers and owners.

On that ground some of the proposals assert that some of the no-fault liability regimes can be interpreted to apply indiscriminately to all cases where no direct link can be established between the defendant's use of their free will and undesirable outcomes. The reason is that some corresponding social interest individually justifies each of the no-fault liability regimes. The expansion of these alternative liability regimes to allow their indiscriminate application without any corresponding social interest to justify their applications would be contrary to the exceptional nature of these regimes and would shake the foundations of both tort law and criminal law systems. Our rejection of that argument is more pronounced in the context of criminal law, where one of the universally accepted core principles is that 'it is better that ten guilty persons escape than that one innocent suffer'.[30]

Some assert that justifications for expanding the scope of alternative liability regimes can be found in the interests of social fairness.[31] At the EU level, the interests of social fairness are already cited as justifications for the strict product liability regime, which was introduced to provide a fair distribution of damages caused by product defects.[32] According to EU product liability rules, holding producers liable allocates the risks equitably, as defective products are accidental fruits of producers' labours.[33] The present study defends that the same justification cannot be adduced in respect of robots' independent activities, as it cannot be known whose actions have sown the seeds of independent activities that resulted in undesirable outcomes.

---

[30] Jeffrey Reiman and Ernest van den Haag, 'On the Common Saying that it is Better that Ten Guilty Persons Escape than that One Innocent Suffer: Pro and Con' (1990) 7 Social Philosophy and Policy 226, 226-227; Nora Osmani, 'The Complexity of Criminal Liability of AI Systems' (2020) 14 Masaryk University Journal of Law and Technology 53, 55.
[31] Herbert Zech, 'Liability for AI: Public Policy Considerations' (2021) 22 ERA Forum 147, 149
[32] Council Directive 85/374/EEC of 25 July 1985 on the Approximation of the Laws, Regulations and Administrative Provisions of the Member States Concerning Liability for Defective Products [1985] OJ L210/29 (Product Liability Directive).
[33] ibid art 1; Amy L Stein, 'Assuming the Risks of Artificial Intelligence' (2022) 102 Boston University Law Review 979, 992-994.

Admittedly, the actions of manufacturers or masters are the first ones to come to mind. However, since robots operate in uncontrolled everyday environments, it must be acknowledged that the independent activities that result in undesirable outcomes can also be prompted by their random human interactants' or even bystanders' actions.

Some other proposals have suggested that questions about fair reallocation of damages can be bypassed with the introduction of compulsory liability insurance schemes for damages caused by robots' autonomous behaviours which would primarily be funded through the premiums paid by robots' manufacturers.

The present study assesses that such compulsory insurance schemes are likely to be unsustainable. First of all, the types and extent of damages that robots' independent activities can cause are potentially unlimited because of the broad range of robotics applications, meaning that insurers cannot accurately estimate claim frequency rates and mean claim sizes:[34]

> Mandatory insurance for (...) manufacturers will impose liability and deter the wrong tort-feasors. If that is not enough, the augmented harms that characterize AI-based robots, as well as the un-foreseeability problem, are destined to make the task of determining insurance premiums almost impossible.[35]

Under such market conditions, insurers will be more cautious in taking on risks associated with undesirable outcomes of robots' autonomous behaviours.[36] Logically, more caution exercised by insurers should result in higher premiums paid by manufacturers, meaning higher production costs. Apart from manufacturers of specific groups of robots that are high-risk (such as self-driving cars and surgical robots) higher production costs would not return any meaningful protection of revenues for manufacturers. Consequently, they will have to be reflected in higher prices.[37] Higher prices will reduce the quantities sold, and since such prices would not stem from profit motives, lower quantities sold will lead to lower profits for manufacturers, disincentivising the sector.[38] In summary, compulsory insurance schemes would eventually stifle innovation. Instead of protecting interests of social fairness, as claimed by moderate proposals, such schemes are more likely to hamper the pursuit of social interests in general progress.[39]

---

[34] Andrea Bertolini and others, 'On Robots and Insurance' (2016) 8 International Journal of Social Robotics 381, 388; There are some authors who argue that state authorities can also control and subsidise these on behalf of the general public. See Carrie Schroll, 'Splitting the Bill: Creating A National Car Insurance Fund to Pay for Accidents in Autonomous Vehicles' (2015) 109 Northwestern University Law Review 803, 822.

[35] Omri Rachum-Twaig, 'Whose Robot is It Anyway?: Liability for Artificial- Intelligence-Based Robots' (2020) 4 University of Illinois Law Review 1143, 1167.

[36] David Levy, 'Intelligent No-Fault Insurance for Robots' (2020) 1 Journal of Future Robot Life 35, 54.

[37] Ryan Calo, 'Open Robotics' (2011) 70 Maryland Law Review 561, 607.

[38] Amy L Stein, 'Assuming the Risks of Artificial Intelligence' (2022) 102 Boston University Law Review 979, 992-994.

[39] Bugra (n 95) 193.

Even if it is supposed that diverging from the general rules for liability is justified for whatever reason, expanding the application scope of no-fault liability regimes would still produce untenable results. Though the regimes of no-fault liability do not require evidence of any direct causal link between the defendant's use of their free will and undesirable outcomes, that does not necessarily indicate that these regimes are meant to impose liability even when the defendant could not even have contemplated exerting control over the events resulting in these outcomes.

Since robots do not need to act on their own behalf and to their own benefit, they do not have the capacity to have assets or equity that would allow them to cover any damage caused by their autonomous decisions. To meet any liability that may arise from robots' behaviours, it is necessary to utilise the patrimony of another person or persons. Here, in determining whose assets can be utilised to meet the obligations and to what extent, the balance between two fundamental interests should be observed: efficiency (in terms of not hindering innovation and protecting free-market/economic dynamism), and the protection of manufacturers and owners/users of robots and where relevant third parties (in terms of not having their persons and property free from undue interferences). There are already several proposals in the literature to help determine whose assets should be utilised to pay for the damage caused by robots' autonomous behaviours.

One proposal, put forward by Ugo Pagallo, is inspired by the legal status of autonomous robots who could not transact in their own name in Roman Law. Pagallo advocates the revival of the institution of *peculium*, which was used to restrict the master's liability to some specified value.[40] The liability for damages arising from the robot's actions were imposed on the master, but it was limited by the amount determined as the robot's *peculium*. *Peculium* continued to be part of the master's patrimony and thereby allowed the limitation of liability for obligations arising from robot's behaviours without granting the robot the capacity to have rights of their own.[41] Since the *peculium* does not impose any burdens on the manufacturer, it can be deemed to vindicate the interests in economic efficiency and technological innovation. On the other hand, however, the proposal envisages the imposition of all liability on the master, who is not expected to have significant knowledge of the robot's decision-making instruments.

[40] Ugo Pagallo, *The Laws of Robots* (Springer 2013) 103.
[41] Richard Gamauf, 'Slaves Doing Business: The Role of Roman Law in the Economy of a Roman Household' (2009) 16 European Review of History 331.

Another proposal, formulated by Andrea Bertolini, espouses a 'risk-based liability management' approach when determining whose assets will be utilised to cover the damages arising from the behaviours of robots. This risk-based liability management approach stipulates that damages arising from autonomous behaviours of robots should be borne by the party best located to identify potential risks and take precautions. [42] Bertolini's proposal ensures that damages arising from autonomous behaviours of robots are not going to be borne by the injured party. Nonetheless, the proposal does not have any intrinsic safeguards to protect the proposed bearer of liability from potential financial collapse, even though they may have no role, let alone fault, in bringing about the injury for which they are held to account. Bartolini suggests that this lack of integral safeguards can be overcome by employing some additional measures. The additional measures include the introduction of upper limits on the liability that can be imposed, the establishment of automatic compensation funds, and the regulation of compulsory liability insurance for robots.[43] Different additional measures can be employed according to the activity area and functions of the robot in question, making the proposal relatively more flexible compared to those that were examined before.[44] However, especially without additional measures, it is assessed that Bertolini's proposal would result in the over-protection of injured parties' interests to the expense of those of manufacturers. The present study acknowledges that the manufacturer is often best located to identify any risks associated with robots, but that does not change the precept that damages were 'unforeseeable'. If damages were foreseeable to the manufacturer, the source of liability would already be identified as their fault, and no modifications to existing legal instruments would be necessary.

The present study submits that no solution can always protect interests of all concerned parties. If damages are remedied, that is going be at the expense of someone who was not at fault in bringing about these damages and legal security interests of that someone would inevitably be weakened. If that someone is the manufacturer, the general interest in technological innovation and economic dynamism would be damaged as well.

---

[42] Andrea Bertolini, 'Insurance and Risk Management for Robotic Devices: Identifying The Problems' (2016) 16 Global Jurist.

[43] ibid.

[44] Wherever it is impossible to determine who contributed the damages more, it is acknowledged that the fairest practice would be for the manufacturer and the owner to cover the damage caused by the autonomous behaviour of robots jointly. Ultimately, these are the two parties benefit because the robot exists or is functional.

Otherwise, when damages are not fully compensated, legal security interests of the injured party would be disrupted. Any effective framework, therefore, should strive to balance its effects on the individuals' legal security interests with those on the competing general interest in economic efficiency.

## VIII.2 Unified Framework for Robot Law

The inadequacies of existing legal instruments can be overcome by introducing a new legal framework that recognises robot autonomy. Any such recognition must acknowledge that robots can sometimes act 'unforeseeably', beyond the boundaries set and expectations held by the manufacturers and masters of robots.

However autonomous they may be, no robot today can pursue objectives that were not pre-determined. Since the objectives to be pursued by robots are never their own, but rather always of their manufacturers or masters, it appears that no robot can perform behaviours for their purposes or to own their benefit. In other words, even at their most advanced levels, robots are not capable of having interests and values on their own. If they cannot have their interests, they have no use for any legal protection of these interests or rights. That said, especially in criminal law, affording some legal protection to robots may advance the common interests of all members of society. The proposed framework is not meant to offer legal personhood equivalent to human beings and certain corporations. Unlike the status of corporate personhood, the proposed legal status does not equip robots with *capacitas,* or the capacity to have rights and obligations.

Still, robots can make 'original' decisions and perform specific behaviours following those decisions without any supervision. Depending on the parameters of the pre-determined objectives, their autonomy can allow the robot to perform behaviours that cannot be foreseen or sanctioned by their manufacturers and masters. Most science fiction literature on robots, regardless of whether they are presented in a favourable light or not, uses the unforeseeable behaviours of robots as their central dramatic element.[45]

[45] For example, Asimov's robot stories revolve around robots performing sometimes disturbing, and often outright dangerous behaviours, unforeseeable to either their manufacturers or users, but nonetheless in line with their predetermined purposes and values (three laws of robotics). Isaac Asimov, *Robots of Dawn* (Doubleday 1982) 265. Even the most autonomous robot of Asimov's universe, R. Daneel Oliwaw, who comes up with a new purpose-value (the zeroth law) to take precedence over its pre-determined purpose-values, does so on the basis and in fulfilment of its pre-determined purpose-values. Isaac Asimov, *Prelude to Foundation* (Bantam 1989) 170.

Likewise, in the real world, the unforeseeable behaviours of robots planned and carried out by robots to fulfil their pre-determined objectives, already affect others' legally protected interests and private rights. The domains of private and criminal law address different sets of legal questions and have structures that are fundamentally different from one another. Consequently, legal instruments available in each domain stem from different principles and rules. Modifications must reflect these differences between the domains of law. Nonetheless, it must be emphasised that the shortcomings of existing legal instruments -regardless of their domains- either directly or indirectly stem from these instruments' failure to give effect robot autonomy. Therefore, the improvements proposed by the present study are invariably based on the recognition of robot autonomy and its impacts. Since the proposed improvements are united in providing the legal acknowledgement of robot autonomy in their respective domains of law, the proposal espoused by the present study can be regarded as some unified legal framework for robot law.

### VIII.2.1 Modifications in Criminal Law

#### VIII.2.1.a Crimes against Robots

One of the shortcomings of existing criminal law instruments is revealed by the phenomenon of robot anthropomorphism, as the potential desensitisation of society to abusive and violent behaviours is perhaps the most straightforward one to overcome. In selected legal systems, abusive and violent behaviours against some non-human entities are already subject to criminal punishment. The present study asserts that criminalising such behaviours against non-human entities is justified not because these entities feel pain but because they express themselves similarly to human beings, so much so that most human beings 'suppose' they feel pain. These non-human entities, of course, are animals.

When criminal law rules regarding animal rights are examined, it is seen that the behaviours that are prohibited when performed against some animals are allowed against other animals.[46] At least when it comes to prohibiting abusive and violent behaviours against non-human autonomous robots, the main concern is not whether robots in question suffer or are negatively affected by such behaviours.

[46] See Chapter III, Section IIII.2, 'Existing Definitions'.

By contrast, the protections afforded by criminal law appear to correlate with how closely animal behaviour patterns resemble those of human beings. In all three legal systems, a correlation can be observed between the extent to which humanlike qualities can be attributed to an animal and the criminal punishment of abusive and violent behaviours against that animal in criminal law. It is suggested that criminal law rules on animal rights can provide the legal framework model for criminalising abusive and violent behaviours towards robots. The present study proposes that expanding the conceptual infrastructure of the existing legal framework, the criminal law rules concerning animal rights, would sufficiently solve the problem. With the acknowledgement that not all non-human entities expressing their pain in humanlike ways are natural and organic, the abusive and violent behaviours against robots would qualify equally for the imposition of criminal liability.

However, it must be noted that the shortcomings of animal rights protections afforded by criminal law cannot be treated in the same way as punishments for problematic behaviours against social robots. Almost all animals protected by criminal law rules can feel pain and deteriorate their quality of life; however, that is not the case for robots. No matter how humanlike they behave, at least for the present, robots cannot experience positive or negative sensations like any organic entity. Accordingly, it is not in the interest of *robots* that is intended to be protected by expanding the conceptual infrastructure of the legal framework on animal rights abuses to include robot anthropomorphism. Since such behaviours do not inflict any injuries on human beings, no specific human, the punishment of abusive and violent behaviours against robots must be justified in the interests of society. However, no empirical research decisively indicates that impunity for abusive and violent behaviours of robots causes insensitivity towards such behaviours towards human beings and therefore poses a danger to social peace and order that requires criminal law intervention.

The research that can be utilised to clarify the trajectory of the law on this topic is concerned with subsequent behaviours of perpetrators of abusive and violent behaviours towards animals. The statistical data presented earlier in the present study shows that more than half of the perpetrators of crimes such as arson, deliberate injury, and murder have a history of negative behaviours towards animals.[47]

[47] See Chapter IV, 'Challenges of Robot Anthropomorphism'.

However, these statistics do not conclusively prove that the impunity of negative behaviours toward some animals and the resultant insensitivity in society to such behaviours encourage the perpetrators to perform the same behaviours against human beings.

Though the exact nature of this correlation cannot be ascertained, considering that abusive and violent behaviours against animals and witnessing these behaviours are often precursory to criminal behaviours against human beings, and that large part of society attributes feelings of pain to anthropomorphic robots, it is expected that negative behaviours against them will cause the same discomfort to society as such behaviours performed against animals or perhaps even human beings. Consequently, the present study suggests that they may be sufficient public interest to justify criminal punishment of abusive and violent behaviours toward robots.

### *VIII.2.1.b* Crimes By Robots

Robots can perform criminally offensive behaviours thanks to the characteristic of autonomy, unforeseeable to either their manufacturers or masters. Considering that such behaviours emerge from robots' 'guilty minds', in the sense that they are planned through robots' own decision-making processes, it must be evaluated whether imposing criminal punishment on robots themselves can help vindicate the purpose-values of criminal law instruments. The present study considers criminal punishment as the infliction of some hardship or pain on perpetrators in response to their criminal behaviours.[48]

There are several theories regarding the justification of criminal punishment. One of these theories, consequentialism, defends that the imposition of criminal punishment is justified by the positive consequences to society, such as reducing crime through various means, e.g., incapacitation, deterrence, and rehabilitation. It can be said that consequentialism looks to the future to justify the imposition of criminal punishment. According to consequentialists, the imposition of criminal punishment can only be legitimate when the punishment brings about affirmative benefits.

[48] For example, someone sentenced to imprisonment is deprived of their freedom of movement and some of their other fundamental rights, such as the right to privacy.
See Antony Duff, *The Realm of Criminal Law* (OUP 2018) 20.

In other words, no matter how severe the crime is if the imposition of criminal punishment is not anticipated to bring about any social benefits, the imposition of any hardship on the perpetrator by the State cannot be deemed justified.

Another theory for the justification of criminal punishment, retributivism, argues that the moral wrongness attributed to criminal behaviour alone justifies the imposition of punishment. The mixed theories can address other approaches to the justification of criminal punishment. These theories combine elements from both consequentialist and retributivist theories. Mixed theories generally reject the notion that punishment's justification depends on the satisfaction of a single condition. [49] Instead, according to these theories, for any criminal punishment to be justified, it must have desirable consequences and should be imposed on the guilty, proportionately to the gravity of the crime.[50] Inspired by both these mixed theories, the present study submits that the imposition of criminal punishment is justified when its prospective benefits outweigh its associated costs, provided that the punishment is in proportion with the blameworthiness of the perpetrators' conduct.[51] Criminal punishment cannot be regarded as legitimate if its severity exceeds what the perpetrator deserves. Under no circumstances can anyone innocent be justifiably imposed criminal punishment, regardless of the benefits of punishing the innocent.

The present study proposes that the challenges posed by robot autonomy in criminal law can be overcome by imposing criminal liability on robots. Suppose the sense-think-act paradigm resembles the free will of human beings. In that case, it is logical to recognise the 'free will of robots' and hold them liable for the behaviours they decide to perform by processing the available information with their algorithms without breaching the principle of individual criminal liability. The imposition of criminal liability denotes some hardship on perpetrators. However, as robots are incapable of classifying their experiences as positive or negative, it is practically possible to impose hardship. Punishments imposed on robots may partially restrict their functioning, stop them altogether, or reprogram them to disallow any repeat criminal behaviours.

[49] H L A Hart and John Gardner, *Punishment And Responsibility* (OUP 2009) 3.
[50] ibid 9.
[51] Ryan Abbott, *The Reasonable Robot: Artificial Intelligence and The Law* (CUP 2020) 120.

Significantly, the sanctions imposed on those who do not see the hardships they are made to suffer as unfavourable or adverse due to their psychological states are still regarded as criminal punishments since these hardships are 'perceived as negative by the observers' or 'deprives the perpetrator of certain freedoms'.[52] In this context, although they may be happy to be in prison, someone sentenced to imprisonment is subjected to an adverse consequence, and thereby punished according to society's general perception since their freedom of movement is restricted.

The imposition of criminal liability on robots is expected to bring most of the benefits as the criminal liability of human beings. For example, imposing criminal punishments on robots is likely to reduce crime through incapacitation and general deterrence, as the manufacturers would refrain from manufacturing robots that may perform similar behaviours. Consequently, it is submitted that there is no inherent reason that would make imposition of criminal punishment on robots unjustified.

### VIII.2.2 Modifications in Private Law

The present study submits that the unforeseeable behaviours of robots can cause personal injury, property damage, and economic losses to third parties. Moreover, when entering into contracts on behalf of their masters, robots create private rights and obligations for both their masters and counterparts in contracts. In damages that occur due to robots' autonomous behaviours, neither manufacturers nor masters of robots can be attributed fault or unlawful use of their free will, making it unfeasible to use the fault principle to impose liability for the damages caused. Moreover, as robots do not need and cannot have any private rights and obligations, it is impossible to impose any private obligations on robots. Without private rights to fulfil obligations, the imposition of liability directly on robots does not help overcome the challenge. Alternatives to the fault principle in assigning liability, such as various existing forms of secondary or strict liability, do not help overcome the challenges at hand, as explained above since they disrupt the balance that the law is trying to maintain between the interests in efficiency and protection. For instance, vicarious liability would expose the users of these robots to amounts of liability (e.g. obligation to pay compensation for severe personal injury or sizeable economic losses) that would outweigh the advantages of employing robots.

[52] Antony Duff, *The Realm of Criminal Law* (OUP 2018) 20.

For the same reason, strict product liability would outweigh the advantages that are to be gained by manufacturing robots. In both cases, applying any existing no-fault liability rules would disproportionately sacrifice the general interest in technological development and economic dynamism to protect third parties' interests in the safety of their assets and equity.

Moreover, since the damages done to the third parties' private rights would be made good by exacting some diminution of the manufacturers' or masters' private rights without any fault attributable to either of them, the present study holds that employment of any strict liability approach would not have any overall impact on the cumulative interests of all concerned parties. In other words, the public interest in efficiency would be hampered without any benefit to the cumulative interests in the 'safety' of private rights of all parties.[53]

Considering that there is no inherent danger in employing robots -as emphasised using the adjective 'unforeseeable' to describe the damages that their autonomous behaviours may cause- the principle of strict liability for ultra-hazardous activities cannot be utilised as the legal basis for the recovery of damages caused by robots' autonomous behaviours.[54] Given that the existing legal instruments for establishing liability are not of any help, and also in consideration of the submission that robots cannot have the capacity to have private rights, the present study holds that the challenge at hand can be resolved through the legal recognition of the factual reality, that the artificial autonomous robots have the can perform legally meaningful acts and make legally binding decisions, i.e., they have the capacity to act. In other words, in the context of private law, the proposed legal status for artificial autonomous robots entails conferring on these robots the capacity to perform legally relevant actions, but not the capacity to be holders of rights and obligations - some of which may arise because of these actions.

[53] See Chapter VI, Section VI.1.1, 'Core Principles of Contract Law'.
[54] See Chapter VI, Section VI.1.2, 'Shortcomings of Contract Law'.

### *VIII.2.2.a* Quasi-Personhood: The Proposed Legal Status for Social Robots

The proposed legal status for robots, quasi-personhood, is not envisaged to replace the present status of robots as property. Although the quasi-personhood status gives legal effect to robots' capacity to perform legally relevant behaviours, robots are envisaged to retain their status as 'property', with their owners and users maintaining rights of possession, enjoyment, or disposal over them. Since robots are fruits of engineering efforts being exerted over raw materials (over which there are property rights), and also that they cannot determine their own purpose-values and therefore cannot pursue their own individual interests; no matter how autonomous they may be, there can be no justification for equipping them with jurisdictions ducal the capacity to hold rights and obligations. [55]

If the status of personhood signifies that its holder concurrently possesses both capacities, then the envisaged holders of the new legal status can be said to be more capable than other property but not quite as 'capable' as persons. In the literature, proposals for introducing some new legal status also refrain from equipping their prospective holders both capacities associated with personhood are often called 'personhood' but always preceded by some qualifying adjectives.[56] One well-known example can be found in European Parliament's resolution on the recommendations for the civil law rules on robotics which reviewed the possibility of creating the legal status of 'electronic personhood' for robots.[57]

Nevertheless, in all the legal systems reviewed in this study, it is observed that the capacity to have rights and obligations is essential for the status of personhood. Indeed, even when someone's capacity to act is lawfully restricted, for whatever reason; they do not lose their status as persons so long as they have the capacity to hold rights and obligations, and that capacity, in all the selected legal systems, is protected as one of the fundamental human rights that cannot be revoked. Since the proposed legal status does not equip robots with that capacity -the legal capacity, *capacitas*- it can only be described as 'quasi-personhood', indicating that the status is 'almost but not quite that of' personhood.

---

[55] Ryan Abbott, *The Reasonable Robot: Artificial Intelligence And The Law* (University of Cambridge 2020) 4.
[56] Bernd Carsten Stahl, 'Responsible Computers? A Case For Ascribing Quasi-Responsibility to Computers Independent Of Personhood Or Agency' (2006) 8 Ethics and Information Technology 205, 210.
[57] European Parliament Resolution no. (INL) 2015/2103 of 17 February 2017 with Recommendations to the Commission on Civil Law Rules on Robotics [2017] OJ C252/249.

The proposed quasi-personhood allows the legal systems to accept that robots are actors that bring about undesirable consequences through their unlawful behaviours. Since the quasi-personhood does not confer robots the capacity to have rights and obligations, including property rights, holders of this status cannot have any assets to fulfil any liability arising from their behaviours. Any liability for damages arising from robots' improper behaviours will eventually have to be fulfilled by either their manufacturers, owners, or users.

The proposed quasi-personhood ensures that whoever bears liability will do so without being attributed the blameworthiness for wrongful conduct. With the recognition of artificial autonomous robots as quasi-persons, the liability to be borne by their manufacturers, owners, or users becomes indirect or secondary. In other words, these parties will be called to account for the wrongful behaviours of quasi-persons since these new quasi-persons -*actual commissioners of wrongful conducts*- cannot ever have the means to fulfil their obligations.

### VIII.2.2.b Secondary Liability for Quasi-Persons

The notion of secondary liability is not new to any of the national legal systems examined in the present study. In the previous chapters, the present study has examined the most prevalent instances of secondary liability, such as vicarious liability, and it was affirmed that secondary liability is imposed on those who have either the right, duty, or obligation to control entities whose independent activities result in unlawful infringement of other entitities' legally protected interests. For example, one of the common forms of vicarious liability, operating based on *respondeat superior* principle, allows employers to be held liable for actions performed by their employees in the course of their employment. Here, the liability to be shouldered by employers for their employees' wrongful behaviours is limited to those behaviours that are performed in the course of employment. Other forms of secondary liability, such as principal's liability and parental liability, also provide the secondarily liable parties certain ways of absolving themselves of liability. [58] Most, if not all, forms of secondary liability hinge on the notion of control.

[58] Provided that these parties can prove that unlawful consequences would still have occurred even though they exercised sufficient care when controlling actual perpetrators or those undesirable situations in question would still have arisen even if they were diligent enough in exercising control over the original actors.

If damages arising from original actors' behaviours are not foreseeable to secondarily liable parties, these parties can no longer be held accountable, as they could not have been reasonably expected to circumvent these injuries. It was emphasised throughout the present study that the characteristic of robot autonomy, among other things, expresses the unpredictable and, therefore, unforeseeable nature of at least some behaviours of robots. Therefore, the extant secondary liability forms are not expected to have any significant utility when it comes to fulfilling obligations arising that may arise from autonomous behaviours of robots.

It is also observed that, in most forms of secondary liability, the indirectly liable party is held jointly liable with the actual perpetrator of the wrongful behaviour, meaning that the injured party can recover all their damages from either the actual perpetrator or the secondarily liable party. If the indirectly liable party pays for the damages, they can turn to the actual perpetrators for recourse. Since it does not confer robots any capacity to have rights and obligations, the proposed quasi-personhood does not allow for any similar possibility of recourse for the secondary liable parties.

In most forms of secondary liability, indirectly liable party can be determined straightforwardly. After all, that is the party with the right, duty, or obligation to control the behaviours of actual wrongdoers. When it comes to damages arising from robots' independent activities, it is unclear who should bear the secondary liability since three different parties can potentially be imposed secondary liability for robots' autonomous behaviours based on their right to or obligation to control these robots.

The manufacturers, including software programmers who code the decision making and learning algorithms, are the reason why robots are able to perform autonomous behaviours. Indeed, however advanced robots may become, structures of their cognitive processes and the parameters utilised by these processes still must be pre-determined by their manufacturers. Even so, robots' manufacturers cannot determine what decisions are going to emerge from the interactions between the decision-making processes they designed and the inputs to those processes. Indeed, inputs into robots' decision-making processes -regardless of whether these inputs are users' commands, raw data, or sensory information- are beyond any control or supervision of their manufacturers.

In short, manufacturers have some, but not total, control over robots' behaviours. Though they cannot determine all possible consequences of every autonomous action that the given robot may perform; because of their roles in planning and building robots' decision-making processes, manufacturers can prevent robots from performing specific forms of behaviours or perhaps make their performance of specifically risky behaviours less frequent, perhaps subject to the presence of certain conditions pertaining to the inputs. Since manufacturers are performing acts that are bound to impact other members of society by manufacturing artificial autonomous robots, they must adhere to standards of reasonable care per the general duty of care. In other words, manufacturers have the duty to exercise due diligence while performing actions while developing and producing robots. Since the manufacturing process affords some control to the manufacturers over their robots' possible behaviours, the manufacturers can be said to have the duty to exercise reasonable care in their control of robots, making it reasonable to hold them secondarily liable for robots' behaviours.

Owners of robots, on the other hand, have a bundle of rights, referred to as property rights, over robots. This bundle includes the rights to possess the property, control it, exclude others from it, and dispose of it. If so, by virtue of their property rights over robots, their owners clearly have the right to control robots. Having established the owners' right to control robots, it appears logical to hold them secondarily liable for damages arising from their robots' behaviours. It is expected that in most instances, owners and users of robots will be the same parties. In instances where that is not the case, it is assessed that users of robots will also have the right to control their robots. As a result, they too can be justifiably held secondarily liable for the damages caused by autonomous behaviours of robots that they are using.

The existence of three parties that can be imposes secondary liability for robots' behaviours creates uncertainty about how the said secondary liability would be distributed among these parties. Since in most types of secondary liability, there is only one party (employer, parent, principal) to fulfil the obligations arising from the primary actor's behaviours, this uncertainty cannot be addressed by way of analogy to any existing secondary liability norms. The distribution of any secondary liability for damages must balance competing interests in economic dynamism and legal security. In this context, the manufacturers represent free enterprise and technological innovation interests.

If the manufacturers bear all damages arising from autonomous behaviours, that will increase their liabilities to potentially astronomical proportions. No prudent businessperson continues operating their business if it is not sustainable unless their income and assets outweigh their expenses and liabilities. Imposition of secondary liability solely on manufacturers of robots would make ventures in 'robotics and software' industries a lot less profitable and much riskier, with manufacturers becoming increasingly reluctant to invest in developing robots with more autonomy, thereby down the progress rate of technology. Considering that manufacturers would essentially be made to pay for their roles in 'production activities essential for the sustainable growth of any healthy economy, imposing secondary liability on them would eventually decrease production activities, hampering the general interest in economic dynamism.

In addition, the present study adopts the view that imposing secondary liability on the manufacturers would not be fair. In addition to the manufacturers, whose duty to exercise due diligence during manufacturing processes follows from the general duty of care, there are others parties who actually have more comprehensive rights over robots: their owners and users. On the assumption that it would not be feasible to determine whether it is some commands (of their owners/users) or some codes in their algorithms (designed by their manufacturers) that make wrongful actions of robots possible in most cases; imposing secondary liability on the manufacturers means that though the owners/users are entitled to the rewards of their robots' behaviours; the manufacturers would be made to pay for damages caused by these behaviours.

It can be argued that holding manufacturers secondarily liable is akin to holding one's primary school teacher secondarily liable for damages (e.g., bodily injury) arising from the behaviours that one performs in the course of their employment (one assaulting a customer while working as a sales assistant) because the teacher failed to ingrain one with the values of non-physical communications and non-violent interactions. Of course, the above scenario's ending is legally impossible in the real world. As was explored earlier in the present chapter and elsewhere in this study, in all three selected legal systems chosen for comparison in the present study, the employers are held secondarily liable for damages caused by actions performed by their employees in the course of their employment.

The only justifiable basis for imposing any secondary liability on manufacturers of robots can perhaps be found in the product liability norms. In all three legal systems, these rules emerge from the Product Liability Directive, which created a strict liability regime for defective products, applicable in all member states of the European Union.[59] The Directive, in its preamble, affirms that fault-based liability principles no longer adequately protect those who suffer injuries because of defective products and argues that for most defective products, it is practically impossible to prove whether the manufacturer was at fault thanks to 'our age of increasing technicality'.[60] Applying fault-based liability principles would most likely result in no liability being established, ergo, no compensation awarded to the injured party. Instead, the product liability rules impose liability without fault on the manufacturer and establish that manufacturers are to be held liable for damages caused by defects in their products. The Directive also provides several defences for manufacturers, absolving them from liability if they can prove that they made their products as safe as possible and it was not possible for them, in light of the scientific and technical knowledge available at the time of production, to predict and prevent the defects in their products.

The Product Liability Directive does apply to robots since qualify as products according to the definition set by the Directive [61] Nonetheless, the present study agrees that the Directive sets out appropriate boundaries for manufacturers' liability. In damages concerning defective products, where no fault can be attributed to any of the involved parties, it is equitable to impose liability on the party whose lawful behaviours caused defects and subsequent damages (if any), so long as defects were scientifically and technologically predictable.[62] Since the Directive provides equitable boundaries for manufacturers' liability, the present study advocates the designation of robots as products for the purposes Product Liability Directive. In cases where wrongful behaviours of robots are due to the design of their decision-making processes and that it can be proven that their manufacturers could predict and prevent these wrongful behaviours at the time of production, it is equitable to make them to shoulder the burdens caused by defects - wrongful behaviours- of their products.

[59] Council Directive 85/374/EEC of 25 July 1985 on the Approximation of the Laws, Regulations and Administrative Provisions of the Member States Concerning Liability For Defective Products [1985] OJ L210/29.
[60] ibid, Preamble.
[61] ibid, art 2.
[62] ibid, art 7.

However, the defences that draw equitable boundaries for product liability render the Directive useless in recovering most damages caused by autonomous behaviours robots Autonomous behaviours refer to the reactions of robots to changes in their environments or their masters' commands, decided according to the algorithms designed by manufacturers. Unless flaws in robots' algorithms are evident or command given to them are blatantly unlawful, it is impossible to determine whether robots wrongful behaviours are because of some defects in algorithms. In short, the present study holds that manufacturers can be imposed liability on occasions where it is established that robots' wrongful behaviours are proven to be caused by flawed structures in their algorithms, and these behaviours could have been predicted by their manufacturers.

The present study also acknowledges that such occasions will be infrequent and therefore fails to completely solve the question of distribution of liability for autonomous behaviours of robots. After all, manufacturers of robots exert control only during the production of their robots. If they are held liable for damages whose material causes cannot all be traced back to robots' creation, then the liability imposed on them would be for damages that arise beyond the domain of their control, and that would not be conscionable.

Since it is established that manufacturers can only seldomly be equitably held liable for damages arising from behaviours of robots, it should be admitted that secondary liability for wrongful acts of such quasi-persons should not be imposed on them, leaving only two candidates who can shoulder the liability, robots' owners and users. Owners and users have the legal entitlement to control these robots' behaviours because of robots' status as property. Owners have the right to control their robots as part of the whole bundle of property rights (the right of possession, the right of control, the right of exclusion, the right of enjoyment and the right of disposition) they enjoy; and users have the right of control along with any other right in the bundle except for the right of disposition. Since their owners and users can direct or influence most details of robots' behaviours momentarily, their control is bound to be more comprehensive and hands-on than the levels of control that can be exercised during design and production processes by the manufacturers. Moreover, since they will benefit from positive outcomes of robots' behaviours, imposing secondary liability on them is fairer than holding the manufacturers accountable.

The present study proposes the term 'master' to refer to the party actively possessing the rights to control, enjoy, and use the robot in question. If the owner controls the robot or uses the robot, the master would be the owner. If, on the other hand, there is another party -the user- to whom the owner transfers active control over the robot, then that party would be the master. In cases where the owner shares active control with another party, both parties can be referred to as masters of the robot at hand.

### *VIII.2.2.c* Limitation of Secondary Liability for Quasi-Persons

In the present day, no form of secondary liability is unlimited. For each form of secondary liability, there are some options afforded to the entity that is to be imposed secondary liability to limit the extent of their responsibilities or provide some conditions and defences to the one legal entity's liability for the undesirable outcomes of another entity's activities. For instance, employers are held secondarily liable for their employers' behaviours so long as the set behaviours are performed in the course of employment. Moreover, if the employers can establish that they demonstrated the necessary levels of diligence while recruiting and instructing their employees, they may escape secondary liability. Further, the employers are held jointly liable with their employees, meaning that if they assume the liability arising from employees' behaviours, they have recourse to employees' assets. Considering that robots, when recognised as quasi-persons, shall not have any rights or assets, it will not be possible for whoever is held liable for injuries caused by their robots' behaviours to seek recourse against their robots.

Since robots are going to serve not only their masters' interests but also bring some benefits to the members of their broader societies by emancipating human beings from dangerous, dirty, or dull tasks and increasing general safety by reducing the frequency and gravity of injuries caused by the very human tendency to become distracted; imposition of secondary liability on their masters for all injuries caused by their behaviours does not distribute the liability arising from such behaviours equitably. Even though masters of these robots should bear the brunt of liability, making them unconditionally shoulder all injuries arising from activities they did not will the performance of and whose consequences they could not always have reasonably foreseen does not appear to be just.

By their property rights, owners and users are legally entitled to not only control but also enjoy and use their robots, indicating that they automatically possess any wealth produced by their robots' behaviours regardless of whether they will lead to wealth-producing behaviours. Consequently, it is submitted that whoever is bestowed with the rights to control, use, and enjoy will get the bulk of the benefits brought about by their robots. Still, some indirect and non-transferable benefits even to victims in their positions as members of societies.

As such, if secondary liability cannot be imposed according to some fault-based principle but is determined according to whoever was in any position to benefit from robots' autonomous behaviours, there must be limits to masters' secondary liability. Exposing them to unlimited liability for damages caused by their robots' independent activities would make any utilisation of robots an uncertain risk of potentially catastrophic proportions. It is observed that none of the existing strict liability and secondary liability instruments imposes such incalculable responsibilities for potentially infinite damages that might arise because of actions (or events) that the responsible parties could not always have even contemplated, let alone willed to or allowed. The present study cannot identify any justifications for why the observed pattern should be discontinued regarding the proposed assignment of secondary liability for quasi-person behaviours. Therefore, it follows that injuries beyond the boundaries set out for the secondary liability for quasi-persons will have to be borne by whomever they were inflicted upon in the first place to achieve the fairest appropriation of risks based on the distribution of benefits. In other words, the present study holds that it is fair to argue that some of these unpredictable damages would have to be borne by victims as part of the unavoidable risks of everyday life.

The present study proposes three sets of boundaries that can be utilised to determine what limitations on the master's liability would be fair, just, and reasonable. The first set of boundaries focuses on the types of damages that autonomous behaviours of robots can cause. Thanks to their ever-increasing capabilities, the types of damage these robots can cause are becoming more varied with every passing day. Though the owners/users of robots should attract secondary liability in private law for all personal injuries or deaths caused by their robots, they cannot be expected to undertake liability for all purely economic losses, e.g., damages related to opportunities missed because of their robots' autonomous and unpredictable actions.

The second set of boundaries is related to the utility of robots' damaging behaviours to their masters. If robots' independent activities enriched their masters while causing injury to some third parties, regardless of the type of injury caused, they should be liable up to the amount of the enrichment unless there are any other upper boundaries for that specific type of damage. The second set of boundaries does not apply to usual vicarious liability since they are concerned with the mechanism of unjust enrichment. The third set of boundaries focuses on the liability emerging from legal transactions performed by robots.

It was established earlier that robots, under their proposed status as quasi-persons, will have the capability to make legally valid decisions, and these decisions may result in the creation of private rights and obligations for both their owners/users and their counterparties.

The status of quasi-personhood gives the users/owners the ability to decide what legal transactions can be performed on their behalf and what obligations they authorise their robots to assume on their behalf, especially regarding monetary obligations. When making legal transactions on their owners' or users' behalf, robots are not much different from human agents acting on behalf of their principals. Consequently, it is submitted that they should at least authorise the contractual obligations undertaken by robots to be enforceable against their masters.

Robots, as quasi-persons, are envisaged to be entitled to use the power to make legally valid decisions, but the power they are entitled to use neither comes from nor has any impact on their legal status since their legal status does not come with the capacity to have private rights and obligations, unlike various types of legal personhood.

Thus, according to the quasi-personhood theory, when robots use their capacity to make legally relevant decisions, they do that by exercising the powers of their owners/users. If no power is delegated to these robots, their capacity to make legally relevant decisions cannot create rights or obligations on their masters or counterparties.

The most explicit delegation of power in any efficient market remains to allocate some assets to the agents' exclusive use. In such power transfers, the power source may retain their ownership of allocated assets, but they transfer their exclusive capability to exercise a whole bundle of rights over the allotted capital to the delegate. The principal can revoke their delegation at any stage and regain the rights they have transferred over what has become (or remains) of allocated capital.

The most prevalent example of such delegation of powers in the selected legal systems can be found in limited liability corporations. During the formation of such corporations, partners transfer some of their assets to the corporation. As the partners have the ownership (as owners or co-owners) of their corporations, they also maintain some implicit rights over the assets they have allocated to the company. When corporations are dissolved for any reason, their partners are entitled to whatever has become of the assets they each have allocated, determined by the values of their 'shares' in the corporation. If a corporation is dissolved because of its insolvency, its shares would have no value, and any assets allocated by its partners must be regarded as lawfully disposed of.

Since limited liability corporations are legal persons with both legal capacities, they can have rights and obligations, making it easier to accumulate the resources necessary to undertake larger-scale projects over longer periods. Unlike corporate legal personhood, proposed quasi-personhood does not endeavour to facilitate the accumulation of resources from multiple parties to achieve some common objectives. However, it still involves the dedication of part of the owners' assets to meet the obligations arising from their robots' legal decisions. Since robots are 'property' with no capacity to have rights and obligations on their own, the allocation of assets to be used by these robots does not transfer the ownership titles over these assets. Instead, the dedication of assets to robots (over which they have ownership rights) should be regarded as akin to the creation of 'special funds' within the estates of owners/users of these robots. Through the dedication of special funds, owners/users delegate the use of their powers over those funds to their robots. They guarantee that the robot is authorised to undertake obligations on their behalf when limiting their liability for obligations arising from such legal transactions.

The present study does not advocate making the creation of 'special funds' compulsory for owners/users but instead supports introducing the notion of 'special funds' as the means of delegating power to artificial autonomous robots. If their owners/users choose not to dedicate any part of their patrimony as 'special funds', the robots would be delegated no 'power' to use with their capacity to act or exercise power to form contracts. Consequently, the present study envisages that there will be two types of quasi-persons: those that are allotted 'special funds' and those without special funds.

The quasi-persons who are not allotted special funds still retain their capacity to act. Considering that the secondary liability is to be imposed on the owners/users for tortious behaviours of robots is not justified through the powers they allow their robots to exercise, but instead, based on the powers they exercise over their robots, the extent of secondary liability for tortious liability arising from autonomous behaviours of quasi-persons should not be affected by their owners'/users' decision of whether or not to dedicate 'special funds' to their robots. The present study suggests that the dedication of 'special funds' should only signify the limits of binding obligations that may arise from legal decisions made by robots on behalf of their owners/users.

For obligations that may arise concerning the damages caused to third parties by autonomous behaviours of robots, over which they have the right and duty to exercise control, the owners/users should not be able to set boundaries of their liability, just as one is not able to decide the limits of one's liability for the consequences of their behaviours that they have the right and duty to control. Thus, for imposed, non-contractual obligations that may arise from autonomous behaviours of robots, it is proposed that their owners/robots should be liable for the whole of their patrimonies, regardless of whether they have dedicated any special funds to their robots or not.

If no special funds are dedicated to a robot, that robot would not have any power to access its masters' capacity to have rights and obligations. In that case, the robot cannot make any 'binding promises' or assume enforceable contractual obligations since it cannot have any 'things' in its name over which such obligations can be enforced. In short, to use their capacity to assume obligations, the proposed quasi-persons need at least partial access to their masters' *capacity* over which the obligations are going to be assumed - if they are allocated no special funds, though the obligations they assume would be valid, they would not be 'binding'. The obligations assumed by robots without special funds can be regarded as imperfect obligations. That is to say that though the masters of robots may choose to perform those obligations, their failure to do so would not have any legal implications.

Transfers of possessions of certain assets to robots by their masters should be presumed to be dedications of those assets as 'special funds'. The legal enforcement of that presumption would provide the legal basis for artificial autonomous robots performing everyday transactions, such as grocery shopping, with the money they are given.

Similarly, when their masters share their credit card information with robots, that also amounts to granting robots access to their masters' capacity to have rights and obligations. Every transaction that robots complete using the card information shared with them designates the amount of the obligation created by that transaction as 'special funds. For larger transactions that may require further verification of the delegation of said powers, it is envisaged that the dedication of certain assets or monies to robots' use can be made public through official public registries to be established for this purpose, like official registries of corporations.

The proposal to establish a legal status for robots that is in between that of persons and property is largely inspired by the concept of 'digital peculium', coined by Ugo Pagallo. The concept of 'digital peculium', as discussed earlier in this chapter, is proposed as a solution to the question of limitations to liability for obligations created by robots as well. The concept also enables robots to use their contractual capacity through the legal capacity delegated to them by their owners/users. However, the legal status conferred on robots by introducing the digital peculium concept is unclear.

Though Pagallo notes that the concept of 'digital peculium' does not require granting legal personhood to robots, given that the proposed 'digital peculium' is described as the 'de facto property of robots', it remains unclear as to whether it is also going to determine the boundaries to liabilities arising from tortious conducts. With the proposed notion of robotic quasi-personhood, the uncertain scope of the 'digital peculium' is resolved.

To introduce a legal status of quasi-personhood for social robots the EU would need to develop a comprehensive legal framework that would clarify the associated rights and responsibilities, including those applicable to social robots with personhood. This framework would need to address issues such as data protection, intellectual property, and other legal areas, as well as the implications for human-robot interactions, for which, consultation with a wide range of stakeholders, including industry, academia, civil society organizations, and legal experts, would be necessary.

Finally, implementing and monitoring the proposed reform would be essential to ensure that it remains relevant and effective in the face of rapid technological change and evolving social and ethical considerations.

## Concluding Remarks

The present study examined the legal issues surrounding the interactions between human beings and robots. Following the comparative analysis between selected national legal systems, it is concluded that the inadequacies of existing legal instruments in upholding the legal values that they were originally introduced to maintain stems from the structural shortcomings of the legal systems when it comes to the recognition of legally is significant characteristics of social robots.

The present study asserts that social robots are unique in terms of the disturbance balance of legal interests because of two distinctive characteristics they display, the cost of the display of these on the human psyche and the legally relevant for social- economic consequences facilitated either direct or indirectly by these characteristics.

The characteristic of relative autonomy indicates that social robots, as well as the rest of the robots, can perform certain activities to fulfil their specified tasks and that some of these activities may be unforeseeable or beyond the control of their manufacturers and masters. If and when undesirable outcomes arise out of such activities, the existing legal instruments do not adequately assign the liability would be accountable. It is assessed that the characteristic of relative autonomy resembles the human beings' power of free will – even though robots are not performing these activities for their own interests, they are nonetheless willing these activities.

Consequently, it is suggested the resolution of these legal challenges can be achieved by conferral of quasi-personhood on social robots, which provides the capability to act, but not to have rights and obligations.

The characteristic of social agency means that social robots pose threats the legal values that underlie legal systems. These threats arise from the phenomenon of anthropomorphism, prompted at unprecedented levels by the display of humanlike sociability by social robots. The emotional relatability created as part of the design of social robots does have equivalent impacts on human beings' legally significant behaviours as much as other social interactions, and that means that they threaten certain legal(ly protected) values, such as that in solitude, liberty, and freedom in decision making. The same emotional relatability also means that human beings tend to reach to maltreatment of social robots in like manner with they react to maltreatment of human beings.

In conclusion, the present study focused the power of free will as the main concept of the law, as no responsibility mechanism is provided by the existing legal instruments that at least does not somehow link the responsibility for undesirable outcomes for which one is held liable for the use of free will. The study also established that free will is guided by the individual interests of human beings, some of which are protected or recognised by the law as rights or obligations.

The research performed in the present study revealed that the problem behind the inadequacy of existing legal instruments when faced with robots is not the anthropomorphic responses they invoke or not even the lack of recognition afforded to the relative autonomy. The problem lays because these functional artefacts have some capability that is equivalent to free will, which is the key notion for establishing responsibility, but they do not have individual needs, wants, desires, nor even fears to back to use of free will.

To solve that dilemma, the present study suggests the creation of some new legal status between that of property and legal persons and allows the recognition of the relative autonomy of social robots. The proposal is however, intended to be implemented at the EU level, and how it may be integrated into national legal systems that are not within the scope of comparison performed in the present study can provide basis for future research.

# Bibliography

## Books


*The Contemporary Torah: A Gender-Sensitive Adaptation of the Original JPS Translation* (David E Stein ed, Jewish Publication Society 2006)

Abbott R, The Reasonable Robot (CUP 2020)

Al-Jazari I, *The Book of Knowledge of Ingenious Mechanical Devices* (Donald R Hill tr, Springer 1973)

Alexander L,Kessler-Ferzan K and Morse SJ, *Crime and Culpability* (CUP 2009)

Allen CK, *Law in the Making* (OUP 1927)

Alpa G, and Zeno-Zencovich V, *Italian Private Law* (Routledge-Cavendish 2007)

Angelo JA, *Robotics: A Reference Guide to the New Technology* (Greenwood P 2007)

Antoniolli L, and Veneziano A, *Principles of European Contract Law and Italian Law* (Kluwer Law 2005)

Aristotle, *The Politics* (Trevor J Saunders ed and T A Sinclair tr, Penguin 1981),

Asimov I, *How Did We Find Out About Robots?* (Walker 1984)

— — *I, Robot* (first published 1967, HarperVoyager 2013)

— — *Robots of Dawn* (Doubleday 1982)

— — *Robot Visions* (Penguin 1991)

— — *Robots and Empire* (Doubleday 1985)

— — *Prelude to Foundation* (Bantam 1989)

Atiyah PS and Smith SA, *Atiyah's Introduction to the Law of Contract* (6th edn, Clarendon 2005) 36.

Barfield W, and Pagallo U, *Advanced Introduction to Law and Artificial Intelligence* (Edward Elgar 2020)

Baron-Cohen S, *Mindblindness: An Essay on Autism and Theory of Mind* (MIT 1995)

Bekey GA, *Autonomous Robots: From Biological Inspiration to Implementation and Control* (MIT 2005)

Bentham J and others, *An Introduction to The Principles of Morals and Legislation* (Clarendon 1996)

Bohlander M, *Principles of German Criminal Law* (Hart 2009)

Borchardt K, *The ABC of EU Law* (7th edn, Publications Office of the EU 2018)

Breazeal C, *Designing Sociable Robots* (MIT 2002)

Brooks RA, *Flesh And Machines: How Robots Will Change Us* (Pantheon 2002)

Burnham J, *Beyond Modern Sculpture: The Effects of Science and Technology on the Sculpture of This Century* (4th edn, George Braziller 1975)

Charleton P and others, *Criminal Law and Evidence* (2nd edn, Bloomsbury 2020)

Chopra S, and White LF, *A Legal Theory for Autonomous Artificial Agents* (UM 2011)

Clarke I, *British Future Fiction, 1700-1914, Vol 8* (Routledge 2001) .

Cruz PD, *Comparative Law in A Changing World* (Cavendish 1999)

Danaher J and McArthur N, *Robot Sex* (MIT 2017)

*Brainstorms: Philosophical Essays on Mind and Psychology* (MIT 1981)

Dennett DC, *The Intentional Stance* (MIT 1998)

Devlin P, *The Enforcement of Morals* (OUP 1965)

Duff A, *The Realm of Criminal Law* (OUP 2018)

Dumouchel P, and Damiano, *Living with Robots* (Malcolm Debevoise tr, Harvard UP 2017)

Dworkin R, *Taking Rights Seriously* (Harvard UP 1978)
Engelberger JF, *Robotics in Practice* (Kogan Page 1980)
— — *Robots in Service* (MIT Press 1989)
Flaherty DH, *Protecting Privacy in Surveillance Societies* (UNC Press 1989)
Fletcher M, Gilmore W, and Lööf R, *EU Criminal Law and Justice* (Edward Elgar 2010)
Foot P, *Natural Goodness* (Clarendon P 2010)
French PA, *Collective and Corporate Responsibility* (Columbia UP 1984)
Ghafurian M, Hoey J, and Dautenhahn K, 'Social Robots for the Care of Persons With Dementia' (2021) 10 ACM Transactions on Human-Robot Interaction 1
Gibilisco S, *Concise Encyclopaedia of Robotics* (McGraw-Hill 2002)
Gieseke J, *The History of the Stasi* (Berghahn 2014)
Gittinger JL, *Personhood in Science Fiction* (Palgrave Macmillan 2019)
Gray JC, *The Nature and Sources of Law* (Adamant Media Corporation 2006)
Greengard S, *The Internet of Things* (MIT 2015)
Gutteridge HC, *Comparative Law: An Introduction to the Comparative Method of Legal Study and Research* (CUP 1946)
Hallevy G, *When Robots Kill: Artificial Intelligence under Criminal Law* (NEU Press 2013)
— —*Liability for Crimes Involving Artificial Intelligence Systems* (Springer 2016)
Hart HLA, *Essays in Jurisprudence and Philosophy* (Clarendon 1983)
— —*Punishment and Responsibility: Essays in the Philosophy of Law* (OUP 2008)
— —and John Gardner, *Punishment And Responsibility* (OUP 2009)
Homer, *The Iliad* (Bernard Knox ed and Robert Fagles tr, Penguin 1990)
Idel M, *Golem: Jewish Magical and Mystical Traditions on the Artificial Anthropoid* (SUNY 1990)
Jhering R, *Geiſt Des Römiſchen Rechts* (Breitkopf un Härtel 1858)
— —, *Law as a Means to an End* (Isaac Husik tr, Boston Book Company 1913)
— —, *The Struggle foe Law* (John J Lalor tr, Callaghan & Company 1915)
Jordan JM, *Robots* (MIT Press 2016)
Kamm A, *Julius Caesar* (Routledge 2006)
Kant I, *Lectures on Anthropology* (CUP 2012)
— — *Lectures on Ethics* (CUP 1997)
— — *Groundwork of The Metaphysics of Morals* (CUP 1998)
Kelsen H, *General Theory of Law and State* (Transaction 2006)
Kurki V, *A Theory Of Legal Personhood* (OUP 2019)
LaFave WR, *Substantive Criminal Law (Vol 1)* (3rd edn, Thomson Reuters 2018)
Lando O, *Principles of European Contract Law - Part III* (Kluwer Law 2003)
Levy D, *Love + Sex with Robots* (HarperCollins 2007)
List C and Pettit P, *Group Agency: The Possibility, Design, and Status of Corporate Agents* (OUP 2013)
Lochner LP, *The Goebbels Diaries* (Award 1971)
Macias J, *Education In Emerging Global Society* (ED-Tech 2021)
Mak C, *Fundamental Rights in European Contract Law* (Kluwer Law 2008)
Markesinis B, Unberath H, and Johnston A, *The German Law of Contract* (Bloomsbury 2006)
McMahon B and Binchy W, *Law of Torts* (4th edn, Bloomsbury 2013)
Münch B, *Legal Questions with Autonomous Cars* (Akademikerverlag 2014)
Naas CI, and Reeves B, *The Media Equation* (CUP 1996)
Newenham P, *Silicon Docks: The Rise of Dublin as a Global Tech Hub* (Liberties 2015)
Noriega AE, *The EU Principle of Subsidiarity and its Critique* (OUP 2002)

Norvig P, and Russell SJ, *Artificial Intelligence: A Modern Approach* (Prentice Hall 1995)
Nourbakhsh IR, *Robot Futures* (MIT 2013)
Orwell G, *Nineteen Eighty-Four* (Secker & Warburg 1949)
Pagallo U, *The Laws of Robots* (Springer 2013)
Pascal B, *Pascal's Pensees* (Dutton 1958)
Peters A, *Beyond Human Rights: The Legal Status of the Individual in International Law* (CUP 2014)
Pollard J, and Reid H, *The Rise and Fall of Alexandria* (Penguin 2007)
Posner RA, *Law and Literature.* (3rd edn, Harvard UP 2009)
Pound R, *Social Control Through Law* (Routledge 2017)
Rumi, *Selected Poems* (Coleman Barks tr, Penguin 2004)
Robinson PH, *Intuitions of Justice and the Utility of Desert* (OUP 2013) 162
Rosheim ME, *Robot Evolution: The Development of Anthrobotics* (Wiley 1994)
Samuelson P, *Economics* (McGraw-Hill 1973)
Sharpe VA, *Accountability: Patient Safety and Policy Reform* (Georgetown UP 2004) 195.
Shaw I, *Ancient Egyptian Technology and Innovation* (Bloomsbury 2015)
Shelley M, *Frankenstein: The Modern Prometheus* (first published 1818, Routledge 2004)
Smits JM, *Contract Law: A Comparative Introduction* (Edward Elgar 2014)
Stammler R, *The Theory of Justice* (Lawbook Exchange 2000)
Stevens JP, *Five Chiefs* (Little, Brown & Company 2011
Stevens RH, *Torts and Rights* (OUP 2009)
Tadros V, *Criminal Responsibility* (OUP 2010)
Tadros V, *The Ends of Harm* (OUP 2011)
Thielscher M, *Reasoning Robots: The Art and Science of Programming Robotic Agents* (Springer 2005)
Turkle S, *Alone Together: Why We Expect More from Technology and Less from Each Other* (Basic 2012)
Tzafestas SG, *Roboethics: A Navigating Overview* (Springer 2016)
Wagner G, *The Common Frame of Reference: A View from Law & Economics.* (Sellier 2009)
Watson A (ed), *The Digest of Justinian* (University of Pennsylvania P 2009) vol 1
Weizenbaum J, *Computer Power and Human Reason: From Judgment to Calculation* (Freeman 1976)
White EG, *Tort Law in America: An Intellectual History* (OUP 2006)
Winfield PH, *The Province of the Law of Tort* (CUP 1931)
Zweigert K, and Kötz H, *Introduction to Comparative Law* (OUP 2011

## Electronic Books (E-Books)

Bensoussan A and Bensoussan J (eds), *Comparative Handbook: Robotic Technologies Law* (Kindle edn, Larcier 2016)
Clark R, *Contract Law in Ireland* (7th edn, Kindle edn, Round Hall 2013)
Čapek K, *R.U.R.: Rossum's Universal Robots* (David Wyllie tr, originally published 1920, eBooks@Adelaide 2014) <https://ebooks.adelaide.edu.au/c/capek/karel/rur/index.html> accessed 5 October 2018
Castro MA , *Robotic Law: The Legal Personality of the Robot* (Kindle edn, 2019)

Gunkel DJ, *Robot Rights* (Kindle edn, MIT 2018)
Little W, *Introduction to Sociology: 2nd Canadian Edition* (BC Open Textbooks 2016) <https://opentextbc.ca/introductiontosociology2ndedition/> accessed 5 October 2018
Turner J, *Robot Rules* (Kindle edn, Palgrave Macmillan 2018)

## Book Chapters

Afkhami M, 'Gender Apartheid and the Discourse of Relativity of Rights in Muslim Societies' in Howland CW (ed), *Religious Fundamentalisms and the Human Rights of Women* (Palgrave Macmillan 1999)
Asaro PM, 'A Body to Kick, but Still No Soul to Damn: Legal Perspectives on Robotics' in K , Bekey GA, and Lin P (eds), *Robot Ethics: The Ethical and Social Implications of Robotics* (MIT 2014)
Asaro PM, 'Jus Nascendi, Robotic Weapons and the Martens Clause' in Calo R, Froomkin AM, and Kerr I (eds), *Robot Law* (Edward Elgar 2016)
Baum SD, and White TN, 'Liability for Present and Future Robotics Technology' in Abney K, Jenkins R, and Lin P (eds) *Robot Ethics 2.0: From Autonomous Cars to Artificial Intelligence* (OUP 2017)
Bello P, and Guarini M 'Robotic Warfare: Some Challenges in Moving from Noncivilian to Civilian Theaters' in Abney K , Bekey GA, and Lin P (eds), *Robot Ethics: The Ethical and Social Implications of Robotics* (MIT 2014)
Bottolico B, and Santasuosso A, 'Autonomous Systems and Law: Why Intelligence Matters, A European Perspective' in Hilgendorf E and Seidel U (eds), *Robotics, Autonomics, and the Law* (Nomos 2017)
Bratman M, 'What is Intention?', in Cohen PR, Morgan J, and Pollack ME (eds), *Intentions in Communication* (MIT 2012)
Bugra A, 'Room for Compulsory Product Liability Insurance in the European Union for Smart Robots? Reflections on the Compelling Challenges', in Noussia K and Marano P (eds), *InsurTech: A Legal and Regulatory View* (Springer 2019)
Calo R, 'Robots and Privacy' in Abney K , Bekey GA, and Lin P (eds), *Robot Ethics: The Ethical and Social Implications of Robotics* (MIT 2014)
— — 'Robots in American Law' in Hilgendorf E and Seidel U (eds), *Robotics, Autonomics, and the Law* (Nomos 2017)
Castells M, 'The Impact of the Internet on Society: A Global Perspective' in *Change: 19 Key Essays on How Internet is Changing Our Lives* (BBVA 2015)
Chatila R and Hertzberg J, 'AI Reasoning Methods for Robotics' in Khatib O and Siciliano B (eds), *Handbook of Robotics* (Springer 2008)
Chen W, Boase J, and Wellman B, 'The Global Villagers: Comparing Internet Users And Uses Around The World' in Wellman B and Haythornthwaite C (eds), *The Internet in Everyday Life* (Blackwell 2002)
Cooper DM, 'The Application of a "Sufficiently and Selectively Open License" to Limit Liability and Ethical Concerns Associated with Open Robotics' in Calo R, Froomkin AM, and Kerr I (eds), *Robot Law* (Edward Elgar 2016)
Darling K, 'Extending Legal Protection to Social Robots: The Effects of Anthropomorphism, Empathy, and Violent Behavior Towards Robotic Objects' in Calo R, Froomkin AM, and Kerr I (eds), *Robot Law* (Edward Elgar 2016)
— — 'Who's Johnny? Anthropomorphic Framing in Human–Robot Interaction, Integration, and Policy' in Abney K, Jenkins R, and Lin P (eds) *Robot Ethics 2.0: From Autonomous Cars to Artificial Intelligence* (OUP 2017)

Dautenhahn K and others, 'Socially Intelligent Agents: Creating Relationships with Computers and Robots' in Dautenhahn K and others (eds), *Socially Intelligent Agents* (Kluwer 2002)

Davies J, and Nolan D, 'Torts and Equitable Wrongs', in Burrows A (ed), *English Private Law* (3rd edn, OUP 2013)

Dennett D, 'Intentional Systems Theory' in Beckermann A, McLaughlin P, and Walter S (eds), *The Oxford Handbook of Philosophy of Mind* (OUP 2009)

Dentzel Z, 'How the Internet Has Changed Everyday Life' in *Change: 19 Key Essays on How Internet is Changing Our Lives* (BBVA 2014)

Dunne B, 'Power and Sexuality in the Middle East' in Koggel C (ed), *Moral Issues in Global Perspective - Volume 2: Human Diversity and Equality* (Broadview 2006)

Feinberg J, 'The Rights of Animals and Unborn Generations' in Blackstone W (ed), *Philosophy and Environmental Crisis* (University of Georgia P 1974) 51.

Froomkin AM, 'Introduction' in Calo R, Froomkin AM, and Kerr I (eds), *Robot Law* (Edward Elgar 2016)

Gilbert F, 'United States Chapter', in Bensoussan A and Bensoussan J (eds), *Comparative Handbook: Robotic Technologies Law* (Kindle edn, Larcier 2016)

Gitiu SM, 'The Roboticization of Consent' in Calo R, Froomkin AM, and Kerr I (eds), *Robot Law* (Edward Elgar 2016)

Gladden ME, 'The Diffuse Intelligent Other: An Ontology of Nonlocalizable Robots as Moral and Legal Actors' in Nørskov M (ed), *Social Robots: Boundaries, Potential, Challenges* (Taylor & Francis 2018)

Guastaferro B, 'The European Union as A Staatenverbund? The Endorsement of The Principle of Conferral in the Treaty of Lisbon', in Rubini L and Trybus M (eds), *The Treaty of Lisbon and the Future of European Law and Policy* (Edward Elgar 2012)

Heck P, 'Interessenjurisprudenzin' in Schoch M (ed), *The Jurisprudence of Interests: Selected Writings of Max Rümelin, Philipp Heck, Paul Oertmann, Heinrich Stoll, Julius Binder, Hermann Isay.* (Magdalena Schoch tr, Harvard UP 1948)

Henschke A, 'The Internet of Things and Dual Layers of Ethical Concern ' in Abney K, Jenkins R, and Lin P (eds) *Robot Ethics 2.0: From Autonomous Cars to Artificial Intelligence* (OUP 2017)

Hoffman H, 'From Jhering to Radbruch: On the Logic of Traditional Legal Concepts to the Social Theories of Law to the Renewal of Legal Idealism' in Canale D, Grossi P, Hoffman H (eds) *A Treatise of Legal Philosophy and General Jurisprudence* (Springer 2009)

Houttie HV, 'Changed Circumstances and Pacta Sunt Servanda' in Gaillard E (ed) *Transnational Rules in International Commercial Arbitration* (ICC 1993)

Hu Y, 'Robot Criminal Liability Revisited' in Yoon S, Han SH, and Ahn SJ (eds), *Dangerous Ideas in Law* (Bobmunsa 2018)

Hubbard FP, 'Allocating the Risk of Physical Injury from "Sophisticated Robots": Efficiency, Fairness, and Innovation' in Calo R, Froomkin AM, and Kerr I (eds), *Robot Law* (Edward Elgar 2016)

Jacobs F, 'Comparative Law and European Union Law', in Reimann M and Zimmermann R (eds), *The Oxford Handbook of Comparative Law* (2nd edn, OUP 2019)

Karnow CEA, 'The Application of Traditional Tort Theory to Embodied Machine Intelligence' in Calo R, Froomkin AM, and Kerr I (eds), *Robot Law* (Edward Elgar 2016)

Knifka J, 'On the Significance of Understanding Human-Robot Interaction' in M Nørskov (ed), *Social Robots: Boundaries, Potential, Challenges* (Taylor & Francis 2018)

Laaksolahti J, Lönnqvist P, and Persson P, 'Understanding Social Intelligence' in Dautenhahn K and others (eds), *Socially Intelligent Agents* (Kluwer 2002)
Lin P, 'Introduction to Robot Ethics' in G Bekey, K Abney, and P Lin (eds), *Robot Ethics: The Ethical and Social Implications of Robotics* (MIT 2014)
Lober A, Caesar T, and W Ropel, 'Germany Chapter', in Bensoussan A and Bensoussan J (eds), *Comparative Handbook: Robotic Technologies Law* (Kindle edn, Larcier 2016)
Matsuzaki H, 'Robots, Humans, and the Borders of the Social World' in Nørskov M (ed), *Social Robots: Boundaries, Potential, Challenges* (Taylor & Francis 2018)
McKendrick E, 'Contract: In General', in Burrows A (ed), *English Private Law* (3rd edn, OUP 2013)
Michot Y, 'The Image of God in Humanity from a Muslim Perspective' in Harries RD, Solomon N, and Winter T (eds), *Abraham's Children: Jews, Christians and Muslims in Conversation* (T&T Clark 2006)
Monateri PG, 'Introduction' in Monateri PG (ed), *Methods of Comparative Law* (Edward Elgar 2012) 3.
Nørskov M, 'Editor's Preface' in Nørskov M (ed), *Social Robots: Boundaries, Potential, Challenges* (Taylor & Francis 2018)
— —'Technological Dangers and The Potential of Human–Robot Interaction: A Philosophical Investigation of Fundamental Epistemological Mechanisms of Discrimination' in Nørskov M (ed), *Social Robots: Boundaries, Potential, Challenges* (Taylor & Francis 2018)
Panebianco G, 'The Nulla Poena Sine Culpa Principle in European Courts Case Law', Human Rights in European Criminal Law (Springer 2015)
Papadopoulos E, 'Heron of Alexandria (c. 10-85 AD) in Ceccarelli M (ed), *Distinguished Figures in Mechanism and Machine Science: Their Contributions and Legacies* (Springer 2007) 4-6
Richards NM, and Smart WD, 'How Should the Law Think About Robots?' in Calo R, Froomkin AM and Kerr I (eds), *Robot Law* (Edward Elgar 2016)
Ristroph A, 'Criminal Law for Humans' in Dyzenhaus D and Poole T (eds), *Hobbes and the Law* (CUP 2012)
Rummel RJ, 'Social Behaviour and Interaction' in Rummel RJ (au), *Understanding Conflict and War, Volume 2: The Conflict Helix* (first published Sage 1976, HTML edn) <www.hawaii.edu/powerkills/TCH.CHAP9.HTM> accessed 5 October 2018
Scheutz M, 'The Inherent Dangers of Unidirectional Emotional Bonds between Humans and Social Robots' in Abney K, Bekey GA, and Lin P (eds), *Robot Ethics: The Ethical and Social Implications of Robotics* (MIT 2014)
Schütze R, 'EU Competences: Existence and Exercise' in Arnull A and Chalmers D (eds), *The Oxford Handbook of European Union Law* (OUP 2015)
Sharkey N, 'Killing Made Easy: From Joysticks to Politics' in Abney K, Bekey GA, and Lin P (eds), *Robot Ethics: The Ethical and Social Implications of Robotics* (MIT 2014)
Shaw-Garlock G, 'Gendered by Design: Gender Codes in Social Robotics' in Nørskov M (ed), *Social Robots: Boundaries, Potential, Challenges* (Taylor & Francis 2018)
Shay LA and others, 'Confronting Automated Law Enforcement' in Calo R, Froomkin AM, and Kerr I (eds), *Robot Law* (Edward Elgar 2016)
— — 'Do Robots Dream of Electric Laws? An Experiment in the Law as Algorithm' in Calo R, Froomkin AM, and Kerr I (eds), *Robot Law* (Edward Elgar 2016)

Simester A, 'Is Strict Liability Always Wrong?' in Simester AP (ed), *Appraising Strict Liability* (OUP 2005).
Spencer J, and Pedain A, 'Approaches to Strict and Constructive Liability in Continental Criminal Law', in Andrew Simester (ed), *Appraising Strict Liability* (OUP 2005)
Wagner G, 'Children as Tortfeasors Under German Law' in Martín-Casals M (ed), *Children in Tort Law Part I: Children as Tortfeasors* (Springer 2006)
Yaffe G, 'The Voluntary Act Requirement', in Marmor A (ed), *The Routledge Companion to Philosophy of Law* (Routledge 2012)
Zallone R, 'Italy Chapter', in Bensoussan A and Bensoussan J (eds), *Comparative Handbook: Robotic Technologies Law* (Kindle edn, Larcier 2016)

## Dictionaries & Encyclopaedias

'Pandora (Greek mythology)', *Encyclopaedia Britannica* (8 February 2018). <www.britannica.com/topic/Pandora-Greek-mythology> accessed 5 October 2018
'Pygmalion (Greek mythology)', *Encyclopaedia Britannica* (12 February 2018) <www.britannica.com/topic/Pygmalion> accessed 5 October 2018
'robot' (*Merriam-Webster,* 2017)
<www.merriam-webster.com/dictionary/robot> accessed 15 June 2018
'robot' (*OED Online*, OUP 2018)
< www.oed.com/view/Entry/166641> accessed 5 October 2018
'robot' (*CED Online*, CUP 2022)
<https://dictionary.cambridge.org/dictionary/english/robot accessed 30 September 2022.
Alexander L, and Moore M, 'Deontological Ethics', *The Stanford Encyclopaedia of Philosophy* (Winter edn, 2016)
<https://plato.stanford.edu/archives/win2016/entries/ethics-deontological/>
accessed 1 April 2019
Blau L, Jacobs J, and Eisenstein JD, 'Golem', *The Jewish Encyclopaedia* (1904) vol 6
Copeland BJ, 'Artificial Intelligence', *Encyclopaedia Britannica* (17 August 2018) <www.britannica.com/technology/artificial-intelligence> accessed 5 October 2018
Duff A and Hoskins Z, 'Legal Punishment' *The Stanford Encyclopaedia of Philosophy* (Summer edn, 2021)
<https://plato.stanford.edu/archives/sum2021/entries/legal-punishment/>
accessed 15 June 2021
Turkle S, 'The 'Robotic Moment'', *Encyclopaedia Britannica* (25 May 2018) <www.britannica.com/topic/The-Robotic-Moment-2118595> accessed 12 November 2020.
Wedgwood H, 'Agent', *A Dictionary of English Etymology* (2nd edn, 1872)

## Conference Proceedings


Asaro PM, 'Robots and Responsibility from A Legal Perspective', (2007 IEEE International Conference on Robotics and Automation, Workshop on Roboethics, Rome, April 2007) <http://www.roboethics.org/icra2007/contributions/ASARO%20Legal%20Perspective.pdf> accessed 28 February 2019

Denning T and others, 'A spotlight on security and privacy risks with future household robots: Attacks and lessons' (11th International Conference on Ubiquitous Computing, Florida, October 2009) <http://aiweb.cs.washington.edu/research/projects/aiweb/media/papers/p105-denning.pdf> accessed 5 October 2018

Haddadin S and others, 'Injury Evaluation of Human-Robot Impacts' (2008 IEEE International Conference on Robotics and Automation, Pasedena, May 2008) <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=4543534> accessed 28 February 2019

Kidd CD, Taggart W, and Turkle S, 'A Sociable Robot to Encourage Social Interaction Among the Elderly' (2006 IEEE International Conference on Robotics and Automation, Orlando, May 2006) <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=1642311> accessed 1 April 2019

Kneer M, and Stuart M, 'Playing The Blame Game With Robots' [2021] Companion of the 2021 ACM/IEEE International Conference on Human-Robot Interaction 405

Pagallo U, 'Even Angels Need the Rules: AI, Roboethics, and the Law', *European Conference on Artificial Intelligence (ECAI)* (IOS 2016) 209

Rădulescu TV, 'The Principle of Relative Effect of Contracts: A Historical View And Aspects Of Comparative Law', *CKS 2018: Challenges of the Knowledge Society* ("Nicolae Titulescu" University Publishing House 2018)

Siegel M, 'The sense-think-act paradigm revisited' (1st International Workshop on Robotic Sensing, Orebo, June 2003) <https://ieeexplore.ieee.org/document/1218700> accessed 5 October 2018

Wellman B and others, 'The Internet In Everyday Life' (Euricom Conference on e-Democracy, Nijmegen, October 2002) <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.11.8754&rep=rep1&type=pdfZ> accessed 5 October 2018

## Reports and Other Documents


'Definitions, Principles and Model Rules of European Private Law: Draft Common Frame of Reference Outline Edition' (*Study Group on a European Civil Code and the Research Group on EC Private Law*) (1st edn, Sellier, European Law Publishers 2009)

'ISO 8373:2012 - Robots and Robotic Devices — Vocabulary' (Standard, ISO 2012)

'ISO 8373:2021 - Robotics— Vocabulary' (Standard, ISO 2021).

'Principles of The European Contract Law' (*Commission on European Contract Law*) <hwww.trans-lex.org/400200/_/pecl/> accessed 31 May 2017.

'Principles to Underpin Management by Public Entities of Funding' (Office of the Auditor-General 2006) 17-23.

Bossi S and others, 'Robotics in Italy' (Report, ENEA 2014)

DiTecco D, 'New Technology, Same Old Stigma: An Analysis of Feminist Discourses and Sex Work Stigma in Sex Robot Media' (MA, Carleton University 2020)

Docherty B, 'Mind the Gap: The Lack of Accountability for Killer Robots' (Report, Human Rights Watch 2015) <www.hrw.org/sites/default/files/reports/arms0415_ForUpload_0.pdf> accessed 5 October 2018

European Commission, 'A European Approach to Artificial Intelligence' (Shaping Europe's Digital Future: 7 June 2022) <digital-strategy.ec.europa.eu/en/policies/european-approach-artificial-intelligence> accessed 9 July 2022

Fallows D, 'The Internet and Daily Life' (Report, PEW Research Center 2004)

Frey CB and Osborne M, 'The Future of Innovation and Employment' (Report, Citi GPS 2015) <www.oxfordmartin.ox.ac.uk/downloads/reports/Citi_GPS_Technology_Work.pdf> accessed 5 October 2018

Heren C, *Imago Dei: The Doctrine of The Image of God With Regard To Human Origins* (Study, Luther Seminary 2011)

International Federation of Robotics, 'World Robotics 2016 Industrial Robots' (Report, IFR 2016) <ifr.org/img/office/Industrial_Robots_2016_Chapter_1_2.pdf> accessed 5 October 2018

— — 'Robot Density Nearly Doubled Globally' (*International Federation of Robotics*, 2021) <https://ifr.org/ifr-press-releases/news/robot-density-nearly-doubled-globally> accessed 25 July 2022.

— — 'World Robotics 2016 Service Robots' (Report, IFR 2016) <ifr.org/img/office/Service_Robots_2016_Chapter_1_2.pdf> accessed 5 October 2018

Kramer X and others, 'A European Framework for Private International Law: Current Gaps and Future Perspectives' (Study, European Parliament 2012) <www.europarl.europa.eu/RegData/etudes/STUD/2012/462487/IPOL-JURI_ET(2012)462487_EN.pdf> accessed 16 November 2018

Law Reform Commission, *Privity of Contract and Third Party Rights* (LRC 88-2008) 26

Leroux C, and Labruto R, 'Suggestion for A Green Paper on Legal Issues in Robotics' (Report, euRobotics 2012)

Livingston D, *Star Trek: Voyager – Season 7, Episode 20: Author Author* (2001)

Nevejans N, 'European Civil Law Rules in Robotics' (Study, European Parliament 2016) <www.europarl.europa.eu/RegData/etudes/STUD/2016/571379/IPOL_STU(2016)571379_EN.pdf> accessed 5 October 2018

Palmerini E and others, 'Guidelines on Regulating Robotics' (Report, European Commission 2014)<www.robolaw.eu/RoboLaw_files/documents/robolaw_d6.2_guidelinesregulatingrobotics_20140922.pdf> accessed 5 October 2018

Scheerer R, *Star Trek: The Next Generation – Season 2, Episode 9: The Measure of a Man* (1989)

Schur M, *The Good Place - Chapter 7: The Eternal Shriek* (NBC 2016)

Soerjaman K, 'Employer's Liability for Damage Caused By The Employee, A Comparative Study: Germany And The Netherlands' (LLM, University of Tilburg 2013)

Tambornino L, and Lanzerath D, 'Law, AI And Robotics: Germany Country Report' (Report, SIENNA 2019)

UN Human Rights Committee, 'General Comment No 16: Article 17 (Right to Privacy)' in 'Note by the Secretariat, Compilation of General Comments and General Recommendations adopted by Human Rights Treaty Bodies' (1988) UN Doc HRI/GEN/1/Rev.9 (Vol I)

UN Human Rights Council, (48th Session) 13 September–1 October 2021 'The Right to Privacy in the Digital Age: Report of the Office of the United Nations High Commissioner for Human Rights' (15 September 2021) UN Doc A/HRC/48/31

UN Institute for Disarmament Research, 'The Weaponization of Increasingly Autonomous Technologies: Considering how "Meaningful Human Control" might move the discussion forward'(Report, UNIDIR 2014) <www.unidir.org/files/publications/pdfs/considering-how-meaningful-human-control-might-move-the-discussion-forward-en-615.pdf> accessed 5 October 2018

## Journal Articles

Abbott R, and Sarch A, 'Punishing Artificial Intelligence: Legal Fiction or Science Fiction' (2019) 53 UC Davis Law Review 323

Adams B and others, 'Humanoid Robots: A New Kind of Tool' (2000) 15 IEEE Intelligent Systems 25

Airenti G, 'The Development of Anthropomorphism in Interaction: Intersubjectivity, Imagination, and Theory of Mind' (2018) 9 Frontiers in Psychology.

Airenti G, 'The Cognitive Bases of Anthropomorphism: From Relatedness to Empathy' (2015) 7 International Journal of Social Robotics 117

Alheit K, 'The Applicability of the EU Product Liability Directive to Software' (2001) 34 Comparative and International Law Journal of Southern Africa 188

Alkio A, 'Animal Rights in the Third Reich' (2003) 2 (3) Kaltio

Allahverdiyev S, and Othman M, '"Verbandssanktionengesetz"— Corporate Liability For Germany?' (2022) 23 German Law Journal 637

Allan DD, Bartneck C, and Vonasch AJ, 'The Doors of Social Robot Perception: The Influence of Implicit Self-Theories' (2021) 14 International Journal of Social Robotics 127

Allveri RL, 'Protocol No 24: Fact or Fiction for EU Roma?' (2013) 21 Marmara Journal of European Studies 1

Alschuler AW, 'Two Ways to Think About the Punishment of Corporations' (2009) 46 American Criminal Law Review 1359
Arluke A, and Sax B, 'Understanding Nazi Animal Protection and the Holocaust' (1992) 5 Anthrozoös 1
Bakulina LT, 'The Roman Spirit' of the Code of Napoleon' (2014) 5 Mediterranean Journal of Social Sciences 477
Balkin JB, 'The Path of Robotics Law' (2015) 6 California Law Review Circuit 45
Barfield W, "Liability for Autonomous and Artificially Intelligent Robots" (2018) 9 Paladyn, Journal of Behavioral Robotics 193
Barnett R, 'Four Senses of the Public Law-Private Law Distinction' (1986) 9 Harvard Journal of Law & Public Policy 267
Bartneck C, and Hu J, 'Exploring the Abuse of Robots' (2008) 9 Interaction Studies 415
Bartneck C, and Keijsers M, 'The Morality of Abusing a Robot' (2020) 11 Paladyn, Journal of Behavioural `Robotics 271,
Bartoletti M, 'Smart Contracts (2020) 3 Frontiers in Blockchain 1
Bedini SA, 'The Role of Automata in the History of Technology' (1964) 5 Technology and Culture 24.
Bertolini A and others, 'On Robots and Insurance' (2016) 8 International Journal of Social Robotics 381
Bertolini A, 'Robots as Products: The Case for a Realistic Analysis of Robotic Applications and Liability Rules' (2013) 5 Law, Innovation and Technology 214
Bertolini A, and Episcopo F, 'Robots and AI as Legal Subjects? Disentangling the Ontological and Functional Perspective' (2022) 9 Frontiers in Robotics and AI
Binchy W, 'Tort Law in Ireland: A Half-Century Review' (2016) 56 Irish Jurist 199
Binder G, 'Making the Best of Felony Murder' (2000) 91 Boston University Law Review 403
Bordot F, 'Artificial Intelligence, Robots and Unemployment: Evidence from OECD Countries' (2022) 37 Journal of Innovation Economics & Management 117
Botes A, 'Concept Analysis: Some Limitations and Possible Solutions' (2002) 25 Curationis 23
Bouckaert B, 'Corporate Personality: Myth, Fiction or Reality?' (1991) 25 Israel Law Review 156
Bozkurt-Yükse Al, 'Robot Hukuku' (2017) 7 Türkiye Adalet Akademisi Dergisi 85
Bradford A, 'The Brussels Effect' (2012) 1 Northwestern University Law Review 1
Braum S, "Rechtsstaat' And European Criminal Law – From The End Of Sovereignty' (2020) 12 New Journal of European Criminal Law 14
Bryson JJ, Diamantis ME, and Grant TD, 'Of, For, and by the People: The Legal Lacuna of Synthetic Persons' (2017) 25 Artificial Intelligence and Law 273
Buijs L and others, 'Social Needs and Happiness: A Life Course Perspective' (2020) 22 Journal of Happiness Studies 1953
Buijs VL and others, 'Social Needs and Happiness: A Life Course Perspective' (2020) 22 Journal of Happiness Studies 1953,
Burchardt D, 'The Functions Of Law and Their Challenges: The Differentiated Functionality of International Law' (2019) 20 German Law Journal 409
Burchfield K, 'The Nature of Animal Crime: Scope and Severity in Chicago' (2017) 64 Crime & Delinquency 1904
Calo R, 'People Can Be So Fake: A New Dimension to Privacy and Technology Scholarship' (2010) 114 Penn State Law Review 809
— — 'Robotics and the Lessons of Cyberlaw' (2015) 103 California Law Review 513

— — 'Robots as Legal Metaphors' (2016) 30 Harvard Journal of Law & Technology 209
— — 'Open Robotics' (2011) 70 Maryland Law Review 561
Carson H, 'The Trial Of Animals And Insects. A Little-Known Chapter Of Medieval Jurisprudence' (1917) 56 Proceedings of the American Philosophical Society 410
Campa R, 'Social Robots: A Bridge Between the Two Cultures' (2021) 5 KnE Social Sciences
Carlsmith KM and Darley JM, 'Psychological Aspects of Retributive Justice' (2008) 40 Advances in Experimental Social Psychology 193
Carton D, Olszowy W, and Wollherr D, 'Measuring the Effectiveness Of Readability For Mobile Robot Locomotion' (2016) 8 International Journal of Social Robotics 721
Cauffman C, 'Robo-Liability: The European Union In Search Of The Best Way To Deal With Liability For Damage Caused By Artificial Intelligence' (2018) 25 Maastricht Journal of European and Comparative Law 527
Čerka P, Grigienė J, and Sirbikytė G, 'Liability for Damages Caused by Artificial Intelligence' (2015) 31 Computer Law & Security Review 376
Chan W, and Simester A, 'Four Functions ofs Mens Rea' (2011) 70 The Cambridge Law Journal 381
Chesterman S, 'Artificial Intelligence and the Limits of Legal Personality' (2020) 69 International and Comparative Law Quarterly 819
Christensen HG, 'Torts: Strict Liability for Ultra-Hazardous Activities' (1951) 49 Michigan Law Review 919
Ciacchi AC, 'Party Autonomy as a Fundamental Right in The European Union' (2010) 6 European Review of Contract Law 303
Clarke R, 'Asimov's Laws of Robotics: Implications for Information Technology-Part I' (1993) 26 Computer 53
Coeckelbergh M, 'Why Care About Robots? Empathy, Moral Standing, and The Language of Suffering' (2018) 20 Kairos Journal of Philosophy & Science 141
Coghlan S and others, 'Could Social Robots Make Us Kinder or Crueller to Humans and Animals?' (2019) 11 International Journal of Social Robotics 741
Cohen ML, 'The Common Law in the American Legal System: The Challenge of Conceptual Research' (1989) 81 Law Library Journal 13
Coleman J, 'The Morality of Strict Tort Liability' (1976) 18 William & Mary Law Review 259
Collins C and others, 'Artificial Intelligence in Information Systems Research: A Systematic Literature Review and Research Agenda' (2021) 60 International Journal of Information Management
Cottingham J, 'Varieties of Retribution' (1979) 29 (116) The Philosophical Quarterly 238
Cross E and Riddoch K, '"Hit the Robot on the Head With This Mallet" – Making A Case for Including More Open Questions in HRI Research' (2021) 8 Frontiers in Robotics and AI
Csonka P, and Landwehr O, '10 Years After Lisbon – How "Lisbonised" Is The Substantive Criminal Law In The EU?' [2019] (4) eucrim- The European Criminal Law Associations' Forum 261
Čuk T, and Waeyenberge A, 'European Legal Framework for Algorithmic and High Frequency Trading (Mifid 2 And MAR)' (2018) 9 European Journal of Risk Regulation 146

Dacey M, 'Anthropomorphism as Cognitive Bias' (2017) 84 Philosophy of Science 1152,
Dahl O, Eklund B, and Pendrill A, 'Is the Archimedes Principle A Law of Nature? Discussions in an 'Extended Teacher Room'' (2020) 55 (6) Physics Education
Danaher J, 'Robotic Rape and Robotic Child Sexual Abuse: Should They Be Criminalised?' (2014) 11 Criminal Law and Philosophy 71
Danaher J, 'Robots, Law and the Retribution Gap' (2016) 18 Ethics and Information Technology 299
Danaher J, and Petersen S, 'In Defence of the Hivemind Society' (2020) 14 Neuroethics 253.
Dautenhahn K, Fong T, and Nourbakhsh I 'A Survey of Socially Interactive Robots' (2003) 42 Robotics and Autonomous Systems 143
De Graaf M, Hindriks F, and Hindriks K, 'Who Wants To Grant Robots Rights?' (2022) 8 Frontiers in Robotics and AI
Déniz O and others, 'An Engineering Approach to Sociable Robots' (2007) 19 Journal of Experimental & Theoretical Artificial Intelligence 285
Dewey J, 'The Historic Background of Corporate Legal Personality' (1926) 35 The Yale Law Journal 655
DeWolf DK, and Hander DG, 'Assumption of Risk and Abnormally Dangerous Activities' (1990) 51 Montana Law Review 161
Deng E, Mutlu B, and Mataric M, 'Embodiment In Socially Interactive Robots' (2019) 7 Foundations and Trends in Robotics 251
Diamantis ME, 'Corporate Criminal Minds' (2015) 91 Notre Dame Law Review 249
Diamantis ME, and Laufer WS, 'Prosecution and Punishment of Corporate Criminality' (2019) 15 Annual Review of Law and Social Science 453
Dichter A, 'Legal Definitions of Cruelty and Animal Rights' (1978) 7 Boston College Environmental Affairs Law Review 147
Dirik M, 'Al-Jazari: The Ingenious Inventor Of Cybernetics And Robotics' (2020) 1 Journal of Soft Computing and Artificial Intelligence 47
Eisenberg MA, 'Impossibility, Impracticability, and Frustration' (2009) 1 Journal of Legal Analysis 258
Elek B, 'The Connection Between Harmonising Criminal Law And The Occurrence Of An Error In Law — Presented Through Criminal Offenses Against The Natural Environment' [2018] (2) ELTE Law Journal 65
Ely JH, 'Flag Desecration: A Case Study in the Roles of Categorization and Balancing in First Amendment Analysis' (1975) 88 Harvard Law Review 1482
Fletcher G, 'A Fresh Start for the Objective-List Theory of Well-Being' (2013) 25 Utilitas 206
Forsberg L, and Douglas T, 'What is Criminal Rehabilitation?' [2020] Criminal Law and Philosophy
Frank L, and Nyholm S, 'Robot Sex and Consent: Is Consent to Sex Between A Robot and a Human Conceivable, Possible, and Desirable?' (2017) 25 Artificial Intelligence and Law 305
Fuller RC, 'Morals and the Criminal Law' (1942) 32 Journal of Criminal Law and Criminology 624
Funk DA, 'Major Functions of Law in Modern Society' (1972) 23 Case Western Reserve Law Review 257
Gamauf R, 'Slaves Doing Business: The Role of Roman Law in the Economy of a Roman Household' (2009) 16 European Review of History 331

Garcia SM, Chen P, and Gordon MT, 'The Letter versus the Spirit of the Law: A Lay Perspective on Culpability' (2014) 9 (5) Judgment and Decision Making 479
Gardner J, 'Rationality and The Rule of Law in Offences against the Person' (1994) 53 (3) The Cambridge Law Journal 502,
Gardner M, 'The Mens Rea Enigma: Observations on the Role of Motive in the Criminal Law, Past and Present' [1993] Utah Law Review 635
Gavison R, 'Privacy and the Limits of Law' (1980) 89 Yale Law Journal 421
Gerber D, 'Strict Criminal Liability and Justice' (1974) 60 (4) Archives for Philosophy of Law and Social Philosophy 513
Gerstner ME, 'Liability Issues with Artificial Intelligence Software' (1993) 33 Santa Clara Law Review 239
Gestel RV, and Micklitz HW, 'Why Methods Matter in European Legal Scholarship', (2014) 20 European Law Journal 292
Ghodoosi F, 'The Concept of Public Policy in Law: Revisiting the Role of the Public Policy Doctrine in The Enforcement in Private Legal Arrangements' (2016) 94 Nebraska Law Review 685
Gibson CS, 'Legal Impact Analysis: The Ideal and The Practicable' (1982) 24 Journal of the Indian Law Institute 837
Glaeser E, Sacerdote B, and Scheinkman J, 'Crime and Social Interactions' (1996) 111 The Quarterly Journal of Economics 507
Goldberg JCP, and Zipursky BC, 'The Strict Liability in Fault And the Fault In Strict Liability' (2016) 85 Fordham Law Review 743
Gonzalez-Aguirre J and others, 'Service Robots: Trends and Technology' (2021) 11 Applied Sciences 1
Guerra A, Parisi F, and Pi D, 'Liability For Robots I: Legal Challenges' (2021) 18 Journal of Institutional Economics 331
Hanheide M and others, 'Robot Task planning and Explanation in open and uncertain worlds' (2017) 247 Artificial Intelligence 119
Harris J, and Slegers N, 'Performance of a Fire-and-Forget Anti-Tank Missile with a Damaged Wing' (2009) 50 Mathematical and Computer Modelling 292
Hassan S, and Filippi P, 'Decentralized Autonomous Organization' (2021) 10 (2) Internet Policy Review 1
Hay P, 'Frustration and its Solution in German Law' (1961) 10 The American Journal of Comparative Law 345
Henschel A, Laban G, and Cross E, 'What Makes a Robot Social? A Review of Social Robots from Science Fiction to a Home or Hospital Near You' (2021) 2 Current Robotics Reports 9
Hockstein NG and others, 'A History of Robots: From Science Fiction to Surgical Robotics' (2007) 1 Journal of Robotic Surgery 113
Hoecke MV, 'Methodology of Comparative Legal Research' [2015] Law and Method 1
Horstmann AC and others, 'Do a robot's social skills and its objection discourage interactants from switching the robot off?' (2018) 13 PLOS ONE 1,
Hu Y, 'Robot Criminals' (2019) 52 University of Michigan Journal of Law Reform 487
Husak D, 'Retributivism in Extremis' (2012) 32 Law and Philosophy 3
Hutchinson T, 'The Doctrinal Method: Incorporating Interdisciplinary Methods In Reforming The Law' (2015) 3 Erasmus Law Review 130
Ignatescu C, 'Letter and Spirit of The Law: An Interpretation Based on Principles of Law' (2013) 5 (2) Revista Romaneasca pentru Educatie Multidimensionala 75
Ingebretsen M, 'Where's My Personal Robot?' (2009) 24 IEEE Intelligent Systems 90, 91.

Isaacs N, 'The Limits of Judicial Discretion' (1923) 32 (4) The Yale Law Journal 339
Jansen N, and Zimmermann R, 'A European Civil Code in All but Name: Discussing the Nature and Purposes of the Draft Common Frame of Reference' (2010) 69 The Cambridge Law Journal 98
Jensen K, 'Punishment And Spite, The Dark Side Of Cooperation' (2010) 365 Philosophical Transactions of the Royal Society B: Biological Sciences 2635, 2638-2640.
Jeppsson S, 'Retributivism, Justification and Credence: The Epistemic Argument Revisited' (2021) 14 Neuroethics 177
Karnow CEA, 'Liability for Distributed Artificial Intelligences' (1996) 11 Berkeley Technology Law Journal 147, 183.
Keating GC, 'The Priority of Respect over Repair' (2012) 18 Legal Theory 293
Kelly C, 'Privity of Contract- The Benefits of Reform' (2008) 1 Irish Judicial Studies Institute Journal 145
Kerr I, 'Bots, Babes and the Californication of Commerce' (2004) 1 University of Ottawa Law and Technology Journal 285
Khanna V, 'Corporate Criminal Liability: What Purpose Does It Serve?' (1996) 109 Harvard Law Review 1477
King E, 'Clockwork Prayer: A Sixteenth-Century Mechanical Monk' (2002) 1 Blackbird: An Online Journal of Literature of Arts
Kircher A, 'Corporate Criminal Liability versus Corporate Securities Fraud Liability: Analyzing the Divergence in Standards of Culpability' (2009) 46 American Criminal Law Review 157.
Koessler M, 'The Person in Imagination or Persona Ficta of the Corporation' (1949) 9 Louisiana Law Review 435
Koops BJ, Hildebrandt M, and Jaquet-Chiffelle DO, 'Bridging the Accountability Gap: Rights for New Entities in the Information Society?' (2010) 11 Minnesota Journal of Law Science & Technology 497
Kosař D and Lixinski L, 'Domestic Judicial Design by International Human Rights Courts' (2015) 109 American Journal of International Law 713, 743.
Kosterec M, 'Methods of Conceptual Analysis' (2016) 71 Filozofia 220
Lacey N and Pickard H, 'To Blame or To Forgive? Reconciling Punishment and Forgiveness in Criminal Justice' (2015) 35 Oxford Journal of Legal Studies 665
Lacey N, and Pickard H, 'To Blame or To Forgive? Reconciling Punishment and Forgiveness in Criminal Justice' (2015) 35 Oxford Journal of Legal Studies 665
Lachow I, 'The Upside and Downside of Swarming Drones' (2017) 73 Bulletin of the Atomic Scientists 96
Lagioia F, and Sartor G, 'AI Systems Under Criminal Law: A Legal Analysis and A Regulatory Perspective' (2019) 33 Philosophy & Technology 1
Laitinen A and Sahlgren O, 'AI Systems and Respect for Human Autonomy' (2021) 4 Frontiers in Artificial Intelligence
Lederman E, 'Models for Imposing Corporate Criminal Liability: From Adaptation and Imitation Toward Aggregation and the Search for Self-Identity' (2000) 4 Buffalo Criminal Law Review 641
Lee PSN and others, 'Internet Communication Versus Face-To-Face Interaction in Quality of Life' (2011) 100 (3) Social Indicators Research 375
Lefkowitz D, 'Blame and the Criminal Law' (2015) 6 Jurisprudence 451
Lemley M, and Casey B, 'Remedies for Robots' (2019) 86 University of Chicago Law Review 1311

Levy D, 'Intelligent No-Fault Insurance for Robots' (2020) 1 Journal of Future Robot Life 35
Lilkov D, 'Regulating Artificial Intelligence In The EU: A Risky Game' (2021) 20 European View 166
Lima G and others, 'The Conflict Between People's Urge to Punish AI and Legal Systems' (2021) 8 Frontiers in Robotics and AI
List C, 'Group Agency and Artificial Intelligence' (2021) 34 Philosophy & Technology
Lu Y, and Zhao W, 'What Will Robots Be Like In The Future?' (2019) 6 National Science Review 1059, 1059-1061
Lutz C, and Tamó-Larrieux A, 'The Robot Privacy Paradox: Understanding How Privacy Concerns Shape Intentions To Use Social Robots' (2020) 1 Human-Machine Communication 87
MacDougall B, 'Ex Turpi Causa: Should a Defence Arise from a Base Cause' (1991) 55 Saskatchewan Law Review 1
Mahoney PG, 'Contract or Concession-An Essay on the History of Corporate Law' (1999) 34 Georgia Law Review 873
Mamak K, 'Should Violence Against Robots Be Banned?' (2022) 14 International Journal of Social Robotics 1057
— — 'Whether to Save A Robot or A Human: On the Ethical and Legal Limits of Protections For Robots' (2021) 8 Frontiers in Robotics and AI.
Marchisio E, 'In Support of "No-Fault" Civil Liability Rules for Artificial Intelligence' (2021) 1 SN Social Sciences
Mathews M, 'Aspects of the Irish Law of Contract' (1972) 7 Irish Jurist 292
Maza S, 'Politics, Culture, And The Origins Of The French Revolution' (1989) 61 The Journal of Modern History 704
McGuire MR, 'The Laughing Policebot: Automation and the End of Policing' (2020) 31 (1) Policing and Society 20
McPherson T, 'Punishment: Definition and Justification' (1967) 28 Analysis 21
McWhorter L, 'The Morality of Corporate Persons' (2017) 55 The Southern Journal of Philosophy 126
Michigan Law Review, 'Laws that are Made to be Broken: Adjusting for Anticipated Noncompliance' (1977) 75 (4) Michigan Law Review 687
Milgram S, 'Behavioral Study of Obedience.' (1963) 67 The Journal of Abnormal and Social Psychology 371
Mizrahi A, 'The Soul and the Body in the Philosophy of the Rambam' (2011) 2 Rambam Maimonides Medical Journal e0040
Moore MS, 'Justifying Retributivism' (1993) 27 Israel Law Review 15
Moran ME, 'Rossum's Universal Robots: Not the Machines' (2007) 21 Journal of Endourology 1399
Mulligan C, 'Revenge Against Robots' (2018) 69 South Carolina Law Review 579
Naastepad C, and Mulder J, 'Robots and Us: Towards an Economics of the 'Good Life'' (2018) 76 Review of Social Economy 302
Neave N and others, 'The Influence of Anthropomorphic Tendencies on Human Hoarding Behaviours' (2015) 72 Personality and Individual Differences 214
O'Neill A, 'Rescuing the Law of Tort? The Decision of the Supreme Court in O'Neill v. Dunnes Stores' (2010) 45 Irish Jurist 240
Opoku K, 'Delictual Liability in German Law' (1972) 21 International and Comparative Law Quarterly 230
Osmani N, 'The Complexity of Criminal Liability of AI Systems' (2020) 14 Masaryk University Journal of Law and Technology 53

Ossendrijver M, 'Scholars, Priests, And Temples: Babylonian And Egyptian Science In Context' (2021) 8 Journal of Ancient Near Eastern History 1
Owen DG, 'The Five Elements of Negligence' (2007) 35 Hofstra Law Review 1671
Owen RD, 'Tort Liability in German School Law' (1955) 20 (1) Law and Contemporary Problems 72
Pargendler M, 'The Rise and Decline of Legal Families' (2012) 60 American Journal of Comparative Law 1043
Parisi F, 'Alterum Non Laedere: An Intellectual History of Civil Liability' (1994) 39 The American Journal of Jurisprudence 317
Pauw P, 'The Liability of Parents for Loss Caused by Their Children' (1978) 11 The Comparative and International Law Journal of Southern Africa 305
Pettit P, 'Responsibility Incorporated' (2007) 117 Ethics 171
Pieck M, 'A Study of The Significant Aspects of German Contract Law' (1996) 3 Annual Survey of International & Comparative Law 111
Posner RA, 'A Theory of Negligence' (1972) 1 Journal of Legal Studies 29
Pound R, 'A Survey of Social Interests' (1943) 57 Harvard Law Review 1
— — 'Individual Interests of Substance: Promised Advantages' (1945) 59 Harvard Law Review 1
— — 'Liberty of Contract' (1909) 18 Yale Law Journal 454
— — 'Theories of Law' (1912) 22 The Yale Law Journal 114
Preston L, 'Freedom, Markets, And Voluntary Exchange' (1984) 78 American Political Science Review 959
Rachum-Twaig O, 'Whose Robot is It Anyway?: Liability for Artificial- Intelligence-Based Robots' (2020) 4 University of Illinois Law Review 1143
Raja Vora J and others, 'Influence Of A Socially Assistive Robot On Physical Activity, Social Play Behavior, And Toy-Use Behaviors Of Children In A Free Play Environment: A Within-Subjects Study' (2021) 8 Frontiers in Robotics and AI
Rauterberg G, 'The Essential Roles of Agency Law' (2020) 118 Michigan Law Review 609
Reeves N, 'A Rare Mechanical Figure From Ancient Egypt' (2015) 50 Metropolitan Museum Journal 43
Reid E, 'Liability for Dangerous Activities: A Comparative Analysis' (1999) 48 International and Comparative Law Quarterly 731
Reiman J, and Haag E, 'On the Common Saying that it is Better that Ten Guilty Persons Escape than that One Innocent Suffer: Pro and Con' (1990) 7 Social Philosophy and Policy 226
Rheinstein M, 'Common Law and Civil Law: An Elementary Comparison' (1952) 22 Revista Jurídica de la Universidad de Puerto Rico 90
Riordan P, 'Property Rights and Property Duties' (1988) 77 Studies: An Irish Quarterly Review 84
Ronen Y, 'Silent Enim Leges Inter Arma– But Beware the Background Noise: Domestic Courts as Agents of Development of the Law on the Conduct of Hostilities' (2013) 26 Leiden Journal of International Law 599
Rosenfeld M, 'Dworkin and the One Law Principle: A Pluralist Critique' (2005) 3 (233) Revue Internationale de Philosophie 363
Rosenthal-von der Pütten A and others, 'An Experimental Study on Emotional Reactions Towards a Robot' (2012) 5 International Journal of Social Robotics 34
Salvini P and others, 'An Investigation on Legal Regulations for Robot Deployment In Urban Areas: A Focus On Italian Law' (2010) 24 Advanced Robotics 1901

Samuel AL, 'Some Studies in Machine Learning Using the Game of Checkers' (1959) 3 IBM Journal of Research and Development 210
Sarch A, 'Should Criminal Law Mirror Moral Blameworthiness or Criminal Culpability? A Reply To Husak' (2021) 41 Law and Philosophy 305
Sarch A, 'Who Cares What You Think? Criminal Culpability and the Irrelevance of Unmanifested Mental States' (2017) 36 Law and Philosophy 707
Satzger H, 'The Harmonisation Of Criminal Sanctions In The European Union - A New Approach' [2019] (2) eucrim - The European Criminal Law Associations' Forum 115
Schroeder W, 'Limits To European Harmonisation Of Criminal Law' [2020](2) eucrim - The European Criminal Law Associations' Forum 144
Schroll C, 'Splitting the Bill: Creating a National Car Insurance Fund to Pay for Accidents in Autonomous Vehicles' (2015) 109 Northwestern University Law Review 803
Shapo MS, 'Comparing Products Liability: Concepts In European and American Law' (1993) 26 Cornell International Law Journal 279
Shaw-Garlock G, 'Looking Forward to Sociable Robots' (2009) 1 International Journal of Social Robotics 249
Simons KW, 'The Crime/Tort Distinction: Legal Doctrine and Normative Perspectives' (2008) 17 Widener Law Journal 719
Simmons G, 'Free Will and Law: Toward A Pragmatic Approach' (2017) 30 Canadian Journal of Law & Jurisprudence 215
Smithers WW, 'The Code Napoléon' (1901) 49 The American Law Register 127
Smithers WW, 'The German Civil Code (Das Bürgerliche Gesetzbuch) Sources, Preparation, Adoption' (1902) 50 The American Law Register 685
Sneddon L, and others, 'Defining and Assessing Animal Pain' (2014) 97 Animal Behaviour 201
Solaiman SM, 'Legal Personality of Robots, Corporations, Idols and Chimpanzees: A Quest For Legitimacy' (2016) 25 Artificial Intelligence and Law 155
Solove DJ, 'Conceptualizing Privacy' (2002) 90 California Law Review 1082
Solum L, 'Legal Personhood for Artificial Intelligences' (1992) 70 North Carolina Law Review 1231
Stahl BC, 'Responsible Computers? A Case For Ascribing Quasi-Responsibility To Computers Independent Of Personhood Or Agency' (2006) 8 Ethics and Information Technology 20
Stein AL, 'Assuming the Risks of Artificial Intelligence' (2022) 102 Boston University Law Review 979, 992-994.
Stern J, and Walejko M, 'Solitude and Self-Realisation In Education' (2019) 54 Journal of Philosophy of Education 107
Suchman L, 'Subject Objects' (2011) 12 Feminist Theory 119
Sueur C, Forin-Wiart M and Pelé M, 'Are They Really Trying to Save Their Buddy? The Anthropomorphism of Animal Epimeletic Behaviours' (2020) 10 Animals 2323
Taliba AT, 'Rebus Sic Stantibus: A Comparative Survey' (2001) 8 Murdoch University Electronic Journal of Law
Tetley W, 'Mixed Jurisdictions: Common Law vs Civil Law (Codified and Uncodified) ' (2000) 60 Louisiana Law Review 678
Teubner G, 'Rights of Non-Humans? Electronic Agents and Animals as New Actors in Politics and Law' (2006) 33 Journal of Law and Society' 497, 501.

Thomas W, 'Of Systems and Persons: The Ability and Responsibility of Corporate Law To Improve Criminal Punishment' (2016) 78 Ohio State Law Journal 601
Thompson RB, 'Piercing the Corporate Veil: An Empirical Study' (1999) 76 Cornell Law Review 1036.
Turkle S, 'Authenticity in the Age of Digital Companions' (2007) 8 Interaction Studies 501
Ulhøi J, and Nørskov S, 'The Emergence of Social Robots: Adding Physicality and Agency to Technology' (2022) 65 65 Journal of Engineering and Technology Management
Underwood B, and Saiedian H, 'Mass Surveillance: A Study of Past Practices and Technologies to Predict Future Directions' (2021) 4 Security and Privacy e142
Urquiza-Haas EG, and Kotrschal K, 'The Mind Behind Anthropomorphic Thinking: Attribution of Mental States to Other Species' (2015) 109 Animal Behaviour 167
Valcke C, 'Comparative History and the Internal View of French, German, and English Private Law' (2006) 19 Canadian Journal of Law & Jurisprudence 133
Varella M, 'The Biology and Evolution of the Three Psychological Tendencies to Anthropomorphize Biology and Evolution' (2018) 9 Frontiers in Psychology
Veneziano L, and Veneziano C, 'Are Victimless Crimes Actually Harmful?' (1993) 9 Journal of Contemporary Criminal Justice 1,
Vincent G, 'The Laws of Hammurabi' (1904) 9 American Journal of Sociology 737
Vivek Venkatesh G and others, 'Estimation of Volume and Mass of Axi-Symmetric Fruits Using Image Processing Technique' (2014) 18 International Journal of Food Properties 608
Vyver JD, 'The International Criminal Court and the Concept of Mens Rea In International Criminal Law' (2021) 12 (1) University of Miami International and Comparative Law Review 57
Waddams S, 'The Modern Role Of Contract Law' (1983) 8 Canadian Business Law Journal 2
Walby K, 'Research Methods in Law' (2014) 48 Law & Society Review 486
Ware SJ, 'Default Rules from Mandatory Rules: Privatizing Law Through Arbitration' (1999) 83 Minnesota Law Review 703
Warren SD and Brandeis LD, 'Right to Privacy' (1890) 4 Harvard Law Review 193
Waytz A, Cacioppo J, and Epley N, 'Who Sees Human?' (2010) 5 Perspectives on Psychological Science 219
Wehberg H, 'Pacta Sunt Servanda' (1959) 53 The American Journal of International Law 775
Weir K, 'The Dawn of Social Robots: Roboticists Must Dip Heavily into Psychological Science' (2019) 49 Monitor on Psychology 50
Westen P, 'An Attitudinal Theory of Excuse' (2006) 25 Law and Philosophy 289
Westen P, 'Two Rules of Legality in Criminal Law' (2007) 26 Law and Philosophy 229
Wood WR and Suzuki M, 'Four Challenges In The Future Of Restorative Justice' (2016) 11 Victims & Offenders 149
Yntema H, 'Comparative Legal Research: Some Remarks On "Looking Out Of The Cave"' (1956) 54 Michigan Law Review 899
Youngs R, 'Constitutional Limitations on Freedom of Contract: What Can the Germans Teach Us' (2000) 29 Anglo-American Law Review 498
Yuan L and Dennis A, 'Acting Like Humans? Anthropomorphism and Consumer's Willingness to Pay in Electronic Commerce' (2019) 36 Journal of Management Information Systems 450
Zech H, 'Liability for AI: Public Policy Considerations' (2021) 22 ERA Forum 147

Zielinska T, 'Professional and Personal Service Robots' (2016) 4 International Journal of Robotics Applications and Technologies 63

Zuiderveen-Borgesius F, Veal M, 'Demystifying the Draft EU Artificial Intelligence Act — Analysing the Good, the Bad, And the Unclear Elements of the Proposed Approach' (2021) 22 Computer Law Review International 97

## Website and Blogs

'Accident Search Results' (*Occupational Safety and Health Administration,* 1 March 2019) <https://www.osha.gov/pls/imis/AccidentSearch.search?p_logger=1&acc_description=&acc_Abstract=&acc_keyword=robot&Fatal=fatal&sic=&naics=&Office=All&officetype=All&endmonth=02&endday=28&endyear=1984&startmonth=02&startday=28&startyear=2020&InspNr=> accessed 1 March 2019

Ackerman E, 'Lethal Microdrones, Dystopian Futures, and the Autonomous Weapons Debate' (*IEEE Spectrum,* 17 November 2017) <https://spectrum.ieee.org/automaton/robotics/military-robots/lethal-microdrones-dystopian-futures-and-the-autonomous-weapons-debate> accessed 12 April 2018

Armstrong D, 'Emotional Attachment to Robots Could Affect Outcome on Battlefield' (*UW News*, 17 September 2013) <https://www.washington.edu/news/2013/09/17/emotional-attachment-to-robots-could-affect-outcome-on-battlefield/> accessed 15 September 2019

Ackerman E, 'MIT's DragonBot Evolving to Better Teach Kids' (*IEEE Spectrum,* 16 March 2015) <https://spectrum.ieee.org/automaton/robotics/artificial-intelligence/mit-dragonbot-evolving-to-better-teach-kids> accessed 3 December 2018

Bellis M, 'A Brief History of Robots' (*ThoughtCo,* 30 September 2017) <www.thoughtco.com/robot-story-1992373> accessed 15 June 2018

Boran M, 'Saudi Arabia Becomes First Country to Grant a Robot Citizenship' (*The Irish Times*, 2 November 2017) <https://www.irishtimes.com/business/technology/saudi-arabia-becomes-first-country-to-grant-a-robot-citizenship-1.3271620> accessed 11 September 2019.

Bracy J, 'The Future of Privacy: My Journey Down the Rabbit Hole at SXSW' (*Privacy Perspectives*, 20 March 2015) <https://iapp.org/news/a/the-future-of-privacy-my-journey-down-the-rabbit-hole-at-sxsw/> accessed 3 September 2018

Calderwood K, 'Why New Zealand is Granting a River the Same Rights as A Citizen' (*Australian Broadcasting* Corporation, 6 September 2016) <https://www.abc.net.au/radionational/programs/sundayextra/new-zealand-granting-rivers-and-forests-same-rights-as-citizens/7816456> accessed 11 September 2019

Chow D, 'Boston Dynamics' New Atlas Robot Can't Be Pushed Around' (*Live Science,* 24 February 2016) <www.livescience.com/53825-atlas-robot-video-bostondynamics.html> accessed 3 September 2018

Chrisafis A, 'Diamond Heist at Brussels Airport Nets Gang Up To £30M In Gems' The Guardian (19 February 2013) <https://www.theguardian.com/world/2013/feb/19/diamond-robbery-brussels-airport-plane> accessed 16 August 2021

Cuthbertson A, 'An AI Chatbot Just Became a Resident of Japan' (*Newsweek*, 6 November 2017) <https://www.newsweek.com/tokyo-residency-artificial-intelligence-boy-shibuya-mirai-702382> accessed 11 September 2019

Darling K and Vedentam S, 'Can Robots Teach Us What It Means To Be Human?' (*Hidden Brain*, 2017) <https://www.npr.org/templates/transcript/transcript.php?storyId=536424647> accessed 3 April 2019

Dedieu F and Mathieu B, 'Les Robots Vont-Ils Prendre Nos Jobs?' (*L'Expansion.fr*, 3 June 2014) <https://lexpansion.lexpress.fr/high-tech/les-robots-vont-ils-prendre-nos-jobs_1548018.html> accessed 28 February 2019

DeVore V and Nguyen DQ, 'Which countries are testing driverless cars?' (*SWI swissinfo.ch,* 4 March 2016) <www.swissinfo.ch/eng/sci-tech/future-of-transport_which-countries-are-testing-driverless-cars-/41999484> accessed 11 February 2018

Dimov D and Juzenaite R, 'Privacy Concerns About Emotional Chatbots' (*InfoSec Institute,* 16 February 2018) <https://resources.infosecinstitute.com/privacy-concerns-emotional-chatbots/#gref> accessed 26 April 2018

Fisher R, 'Is It OK To Torture or Murder A Robot?' (*BBC Future*, 2013) <http://www.bbc.com/future/story/20131127-would-you-murder-a-robot> accessed 22 November 2020

Griffiths A, 'How Paro the Robot Seal is Being Used to Help UK Dementia Patient' (*The Guardian,* 8 July 2014) <www.theguardian.com/society/2014/jul/08/paro-robot-seal-dementia-patients-nhs-japan > accessed 3 December 2018

Harari YN, 'The Rise Of The Useless Class' (*TED - Ideas*, 24 February 2017) <https://ideas.ted.com/the-rise-of-the-useless-class/> accessed 3 November 2019

Hill K, 'Are Child Sex-Robots Inevitable?' (*Forbes*, 14 July 2014) <https://www.forbes.com/sites/kashmirhill/2014/07/14/are-child-sex-robots-inevitable/#5054c010e46> accessed 3 April 2019.

Kasperkevic J, 'Swiss Police Release Robot That Bought Ecstasy Online' (*The Guardian*, 22 April 2015) <https://www.theguardian.com/world/2015/apr/22/swiss-police-release-robot-random-darknet-shopper-ecstasy-deep-web> accessed 26 July 2022.

Keane D, 'The Six Clues That Have Failed To Solve The Somerton Man Mystery' (*Abc.net.au*, 2021)

Khomami N, 'Egypt Detains Artist Robot Ai-Da Before Historic Pyramid Show' (*The Guardian*, 20 October 2021) <https://www.theguardian.com/world/2021/oct/20/egypt-detains-artist-robot-ai-da-before-historic-pyramid-show> accessed 17 November 2021.

Knight W and Hay J, 'Never Mind Killer Robots—Here Are Six Real AI Dangers To Watch Out For In 2019' (*MIT Technology Review*, 1 July 2019) <https://www.technologyreview.com/2019/01/07/137929/never-mind-killer-robotshere-are-six-real-ai-dangers-to-watch-out-for-in-2019/> accessed 29 July 2020.

Kolesnikov-Jessop S, 'Automatons and Ingenuity' (*The New York Times,* 8 March 2012) <www.nytimes.com/2012/03/08/fashion/08iht-acaw-jaquet08.html> accessed 15 June 2018

MacAskill E and Gabriel D, 'NSA files decoded: Edward Snowden's surveillance revelations explained' (*The Guardian,* 1 November 2013) <www.theguardian.com/world/interactive/2013/nov/01/snowden-nsa-files-surveillance-revelations-decoded#section/2> accessed 16 February 2018

Maruffi F, 'Distributed Ledger Technologies and Smart Contracts In Italy' (*Blockchain - A Blog by Baker McKenzie*, 2019) <https://blockchain.bakermckenzie.com/2019/02/28/distributed-ledger-technologies-and-smart-contracts-in-italy/> accessed 29 July 2020

McGrath D, 'Farewell, Joe Engelberger' (*Electronic Engineering Times,* 12 May 2015) <www.eetimes.com/author.asp?section_id=36&doc_id=1328428> accessed 15 June 2018

Ondrisek B, 'Privacy and Data Security of Chatbots' (*Medium*, 27 October 2016) <https://medium.com/@electrobabe/privacy-and-data-security-of-chatbots-6ab87773aadc> accessed 26 April 2018

'Open Letter to the European Commission Artificial Intelligence and Robotics' <http://www.robotics-openletter.eu> accessed 11 September 2019

Oreck A, 'The Golem' (*Jewish Virtual Library*) <www.jewishvirtuallibrary.org/the-golem> accessed 15 June 2018

Pardes A, 'The Emotional Chatbots Are Here to Probe Our Feelings.' (*Wired*, 31 January 2018) <https://www.wired.com/story/replika-open-source/> accessed 26 April 2018

Parke P, 'Is it cruel to kick a robot?' (*CNN*, 13 February 2015) <https://edition.cnn.com/2015/02/13/tech/spot-robot-dog-google/> accessed 1 April 2019

Read R, 'Rise of the robots will harm the Earth as well as humans' (*The Guardian,* 27 March 2016) <www.theguardian.com/theobserver/2016/mar/27/leters-robots-energy-consuming> accessed 14 February 2018

Rousseau B, 'In New Zealand, Lands and Rivers Can Be People (Legally Speaking)' (*The New York Times*, 13 July 2016) <https://www.nytimes.com/2016/07/14/world/what-in-the-world/in-new-zealand-lands-and-rivers-can-be-people-legally-speaking.html> accessed 11 September 2019

Rowe DG, 'Meet Kismet…' (*New Scientist*, 22 August 1998) <www.newscientist.com/article/mg15921480-800-meet-kismet/> accessed 16 June 2018

Ruiz R, 'This adorable chatbot wants to talk about your mental health′ (*Mashable UK*, 8 June 2017) <https://mashable.com/2017/06/08/mental-health-chatbots/#xhSSFK1TMPqo> accessed 26 April 2018

Safi M, 'Ganges and Yamuna Rivers Granted Same Legal Rights As Human Beings' (*The Guardian*, 21 March 2017) <https://www.theguardian.com/world/2017/mar/21/ganges-and-yamuna-rivers-granted-same-legal-rights-as-human-beings> accessed 11 September 2019

Sanderson D, 'Ai-Da the Art Robot Released from Custody in Egypt' (*The Times*, 21 October 2021) <https://www.thetimes.co.uk/article/ai-da-the-art-robot-released-from-custody-in-egypt-mndhdghxg> accessed 17 November 2021

Shinkle D, 'Self-Driving Vehicles Enacted Legislation' (*NCSL*, 27 August 2018) <https://www.ncsl.org/research/transportation/autonomous-vehicles-self-driving-vehicles-enacted-legislation.aspx> accessed 1 September 2018

Simon M, 'What is a Robot?' (*Wired,* 24 August 2017) <www.wired.com/story/what-is-a-robot/> accessed 14 July 2018

Solon O, 'Google's robot assistant now makes eerily lifelike phone calls for you' (*The Guardian,* 8 May 2018) <www.theguardian.com/technology/2018/may/08/google-duplex-assistant-phone-calls-robot-human> accessed 1 September 2018

Temple E, 'A Brief Literary History of Robots' (*Literary Hub*, 6 April 2017) <https://lithub.com/a-brief-literary-history-of-robots/> accessed 2 August 2022.

Thanaraj T, 'Can Robots Be Prosecuted For A Crime?' (*University of Cumbria*, 14 September 2018) <https://www.cumbria.ac.uk/blog/articles/can-robots-be-prosecuted-for-a-crime.php> accessed 25 December 2021

'The Italian Robot Industry Shows Great Growth Potential' (*#HowToRobot*, 2021) <https://www.howtorobot.com/expert-insight/italian-robot-industry-shows-great-growth-potential> accessed 25 July 2022

Wakefield J, 'Microsoft Chatbot Goes Rogue On Twitter' (*BBC News*, 2016) <https://www.bbc.com/news/technology-35890188> accessed 29 July 2020

Whelan D, 'Guide to Irish Law' (*GlobaLex*, 2019) <https://www.nyulawglobal.org/globalex/Ireland1.html> accessed 28 July 2020

Wilson JH, 'What Is a Robot, Anyway?' (*Harvard Business Review,* 15 April 2015) <https://hbr.org/2015/04/what-is-a-robot-anyway> accessed 14 June 2018

Power M, 'What Happens When A Software Bot Goes On A Darknet Shopping Spree?' (*The Guardian*, 5 December 2014) <https://www.theguardian.com/technology/2014/dec/05/software-bot-darknet-shopping-spree-random-shopper> accessed 17 November 2021

Wale J, and Yuratich D, 'Robot Law: What Happens If Intelligent Machines Commit Crimes?' (*The Conversation*, 1 July 2015) <https://theconversation.com/robot-law-what-happens-if-intelligent-machines-commit-crimes-44058> accessed 20 December 2021

Winder D, 'Is the Future Of Cyber Crime A Nightmare Scenario? - Raconteur' (*Raconteur*, 7 September 2016) accessed 20 December 2021 <https://www.raconteur.net/is-future-cyber-crime-a-nightmare-scenario/> accessed 20 December 2021