Lund University
Department of Political Science

STVM23

Normative Power Europe & AI

How the EU intends to normatively govern artificial intelligence technologies through the Artificial Intelligence Act and its "trustworthy" and "human-centric" approach

LUND UNIVERSITY

Dino Ekdal
FACULTY OF SOCIAL SCIENCE

This thesis is submitted to the Department of Political Science
at Lund University in partial fulfilment
of the requirements for the degree
of Master of Science in
European Affairs

SUPERVISOR:
Ian Manners

“We don’t hate ants, but whenever their presence seriously conflicts with one of our goals we annihilate them without a qualm. The concern is that we will one day build machines that, whether they are conscious or not, could treat us with similar disregard.”

Sam Harris

Abstract

Artificial intelligence is expected to be the next paradigm shift in technology, akin to the steam engine or electricity in its universality and all-encompassing impact. Regulating an innovation of this magnitude is important, however challenging, as evidenced by the European Union's recently proposed Artificial Intelligence Act. This thesis sets out to understand the EU as a normative actor in the field of AI governance, specifically it applies the Normative Power Europe theory to the EU's "human-centric" AI approach and aims to evaluate this self-claimed normativity. It accomplishes this by using a latent qualitative content analysis of 29 documents from official EU channels. The analysis indicates that the EU is indeed intending to act the way a normative power would, although with some significant limitations and incoherencies. However, these can be understood as childhood illnesses owing to the novelty of the EU's approach. A Europe that has the ability to shape what passes as normal in AI governance will be a powerful Europe indeed.

Key words: Artificial Intelligence, AI, Norms, Normative Power, EU, European Union, AI Act, AI Governance

Words: 19 403

Table of contents

Table of abbreviations

AI:	Artificial Intelligence
China:	People's Republic of China
CFSP:	Common Foreign and Security Policy
CFR:	Charter of Fundamental Rights of the European Union
EU:	European Union
EPRS:	European Parliamentary Research Service
GDPR:	General Data Protection Regulation
HLEG-AI:	High-level Expert Group on Artificial Intelligence
LAWS:	Lethal Autonomous Weapons Systems
ML:	Machine Learning
NPE:	Normative Power Europe
OLS:	Ordinary Legislative Procedure
TEU:	Treaty on European Union
US:	United States of America
QCA:	Qualitative Content Analysis

1 Introduction

"In theory, if AI is ever truly able to control itself, it could work 24/7, never sleeping, never lying awake at night wondering why that can't compute that for a certain customer. Imagine such a machine-built software patch that overrides the flaws inherent in human intelligence to make its own decisions and never require a human overseer to tell it no. We haven't even come close to getting there, in theory or in practice. If we ever did, we'd be in a very different place than we are now."

- Inferkit AI-based text-generator 2021

Fed an initial sentence, AI text-completion software, called Inferkit, took creative charge and completed the full ominous paragraph. Notwithstanding the paragraph's lack of coherence or the laziness of yours truly, it illustrates a powerful point. Artificial intelligence is available for you to use, and for it to use you. Social-media companies pin-point your consumer preferences and shoot tailored ads onto whatever flickering device may be before you, it's used in manufacturing to streamline production and in health services to diagnose patients. We're barely scratching the surface of the contemporary applications of AI, but by all accounts, it will become an ever more present factor in human affairs and by extension, European affairs.

Artificial intelligence (AI) is expected to become vital to every sector of the economy, affect the way we consume, produce, communicate, conduct research, provide services, wage war, and to just about anything that involves conventional human intelligence today (Franke & Sartori 2019). AI is more than just another piece of innovation or tech. Some have estimated its impact to be as high as the combustion engine or electricity and more dramatic yet, Russian president Vladimir Putin has said that the nation that leads in AI "will be the ruler of the world" (Vincent 2017). Others, like tech-entrepreneur Ian Hogarth, predict that

"AI policy will become the single most important area of government policy." (Hogarth 2018). Grand words, but certainly not far-fetched or without historical analogies. The United States reaped great economic benefits from the last wave of innovation and became home to the largest tech corporations the world has ever seen, such as Amazon, Facebook, Google, Apple, Intel, and Microsoft. Meanwhile many parts of the world, the European Union included, paid the price by remaining on the side-lines. The next wave of innovation, by many experts' accounts, will come in the form of AI, and missing it would be similarly problematic (Castro, McLaughlin & Chivot 2019).

At the same time, maleficent AI has been used to influence public opinion, such as in the Cambridge Analytica scandal (Shastri, 2019). Governments around the world have begun waking up to the fact that developments and the effects of AI on social life, the economy, or national politics can be erratic (Franke & Sartori 2019). Polities have to adjust to technological developments, to harness the potential of innovation, while also minimizing the risks of said technology. The EU is torn between two ideals. It both wants the status as global leader in AI and while it also wants to regulate AI and ensure "ethical, responsible and sustainable outcomes" (HLEG-AI 2019). Which is a tricky goal considering the fact that the EU has fallen considerably behind the US and China in AI uptake and innovation (Franke & Sartori 2019).

This thesis has identified four important documents that make up the EU's approach to AI regulation: the Communication on Artificial Intelligence (2018), the Ethics Guidelines for Trustworthy AI (2019), the White Paper on Artificial Intelligence (2020) and the AI Act (2021). The EU stakes out its place in the global AI landscape as the epicentre of human-centred AI, in stark contrast to the contemporary hegemons of artificial intelligence innovation like the USA or China. What precisely is this European alternative of human-centred AI and is the EU pursuing it for the sake of "good" values alone?

1.1 Relevance of the thesis

This thesis is an attempt at merging two unnecessarily divorced fields of enquiry, namely political science and the study of emerging technologies, and it does so through the research questions. These are in turn derived from the scientific gap left that previous research has left. None of the identified literature deals precisely with the EU as a normative power in AI governance, except Parviala (2019) who indeed categorized the EU as a normative power in the field of AI. But the choice of material and its comparatively small size warrants a re-evaluation. Especially since the EU has released the, expectedly, influential Artificial Intelligence Act (2021). By analysing a larger set of material and by including the Artificial Intelligence Act, this paper hopes to paint a more detailed picture than existed before. In the process of reviewing literature, a plethora of discourse analyses on the topic emerged (Berger 2018, Humerick 2018, Sharma 2019, Stix 2018 & 2019). By augmenting these approaches with a systematic qualitative content analysis approach hopefully this paper engages in cumulative science.

The introductory section has attempted to explain why we should study AI governance by pointing to the paradigm-shifting consequences of this new technology. To further the relevance of this project the recently proposed Artificial Intelligence Act is scrutinized in its own section of the analysis and compared with the findings of more ideational and visionary documents that are not actionable to the same extent. To summarize: this paper hopes to augment the previous research with a structured qualitative content analysis approach by applying Normative Power Europe theory to the analysed material – and by doing so discovering if we can categorize the EU as a normative power in the field AI governance.

1.2 Research questions & aim

This thesis aims to establish what the EU's concept of AI governance is - an explorative endeavour which lays the baseline for what will subsequently be this thesis' main contribution of exploring to what extent "EU norms" govern the EU's AI approach and whether or not the EU can be considered a normative power in the field of AI. The first research question, to be addressed separately in the analysis section, then becomes:

SRQ 1: What are the prevailing themes in the EU's composite AI strategy, how does it view AI?

This project will identify the central themes and political visions in what is a total of 24 official EU documents and 5 scientific reports commissioned by EU institutions (see 6. Bibliography). These include, but are not limited to proposed regulations, white papers, communications, declarations and scientific reports with the subsequent goal of identifying the European vision of AI. Understanding the EU's AI strategy is a relatively theory-less task, but is nevertheless necessary to reach a fuller understanding of the EU's self-proclaimed normativity in AI governance.

SRQ 2: What norms are conveyed by the EU institutions with the EU's concept of AI?

The theoretical framework used in this paper is Ian Manners' (2002) original theory of Normative Power Europe, but with adaptations derived from Niemann & de Wekker's (2010) work *Normative power Europe? EU relations with Moldova.* The choice of theory is motivated in sections 2.4 – 2.6. The theory becomes extremely useful because of its pre-defined list of EU norms and the operationalizations from Niemann & de Wekker hopefully provide a clear and reproducible method. To understand whether or not the EU is pursuing certain AI governance policies for the sake of "good" values alone, this paper will evaluate the EU's self-proclaimed normativity and determine whether it could be called a normative power in the field of AI governance. This is done through the guiding and overarching main research question:

RQ: Does the EU intend to act the way a normative power would in the field of artificial intelligence governance?

1.3 Disposition

As stated earlier, this thesis is an attempt at merging two unnecessarily divorced fields of enquiry. Namely the political science sub-category of European affairs with studies on emerging technologies. The comparative lack of research on the EU with regards to AI constitutes the main relevance of this thesis and thus its main contribution, The hope is that this endeavour proves both societally and scientifically relevant. Shedding light on a massive issue like AI governance and the immense impact it is expected to have on systems of governance and the ethical status of natural persons, as well as investigating the theoretical adaptability of Manners Normative Power Europe theory to the emerging field of digital policy. An overview of which is given in the following section (see 2.4 The Alternative – Normative Power Europe). The process of locating Manners' ideas in the academic state of the art leads to a discussion on the potential limitations to his way of thinking and subsequently, the thesis presents alternative operationalizations that emerged from this scholarly exchange. Following that is a presentation of the chosen method and research design considered suitable for answering the research questions. The qualitative content analysis approach is structured through the use of a codebook, in which the categories and codes are derived deductively, from the pre-determined categories of the theoretical framework, and inductively through interaction with the material. This was considered necessary in order to adapt NPE to the field of digital governance to not miss out on more nuanced norms and values. In the analysis section, the analysed material is described and key takeaways and findings are presented, in three sub-sections, which were identified during the reading and coding of the material. Finally, a conclusion sums up these findings and is accompanied by a discussion on their meaning and proposes avenues for future scientific enquiry as well as discussing limitations of the theory and method.

2 Theory

2.1 Introduction

Europe likes to see itself as a different species of actor on the international stage. It is a self-perception as a force for good and having the capacity to set the standard for "what is normal" which sets the EU apart (Manners 2002 p.236). In the field of AI, Europe is lagging, however. If Europe is to keep its social values in a world increasingly dictated by software, it must encourage a plethora of actors to internalize the European definitions of freedom, justice, fairness and individual rights, which are embodied in the Charter of Fundamental Rights. Because right now American and Chinese tech-giants seem to be heading in a different direction, and they have a significant head start in terms of empirically measurable capability (Pichot 2019, Franke & Sartori 2019). The field of AI is a new up and coming location where the EU is attempting to exert influence, and this thesis aims to find out whether that influence is normative in nature. The next chapter sets out to identify what AI is, to ensure the paper hits the conceptual mark, and how it has evolved over the years.

Secondly, this section introduces this thesis' theoretical framework which is found in the theory of Normative Power Europe (referred to as NPE) as launched by Ian Manners in the acclaimed paper *Normative Power Europe: A Contradiction in Terms* (2002). It is necessary to complement this presentation with an overview of the historical development of NPE as well as the developments of the theory in the years after its initial publication.

The criticisms of Manners conceptualization are discussed to acknowledge any potential shortcomings of the theory as well as an argument for why the operationalizations of Niemann & de Wekker (2010) are appropriate for the endeavour as captured by the research question. Manners ignited an international

debate on the precise role of the European Union as a foreign-policy actor, with the main-sticking point being the exact qualia of normative power. How do we identify one and what are the criteria of entry into this enlightened club? Through bringing in adaptations and operationalizations from Niemann & de Wekker's work, this thesis makes an attempt at disarming some of the criticisms levied at Manners' original theory (see sections 2.5 & 2.6).

2.2 Past and present understandings of AI

AI is a complicated and technical field of study - but it is unnecessarily divorced from political science scholarship (Valladão 2018). If the prediction of Ian Hogarth (2018) is correct, that AI governance policy will become the single most important policy area for governments, merging the two fields would seem a fruitful endeavour. Starting with how to understand the conceptual evolution of AI, although the political science focus of the overall thesis makes an exhaustive and profound discussion of the historical context of this fascinating technology untenable. Nevertheless, seeing as political scientists are just realizing the impact AI could have. The following section aims to introduce the reader to the central concepts used and analysed in this thesis.

In 1950 the British mathematician Alan Turing, prompted by the rising prominence of machines, and his own experience in cracking the German encryption device Enigma during World War 2, asked "can machines think?". Just five years later American computer scientist Marvin Lee Minsky jotted down the field's fundamental principles, which have remained largely unchanged since (Bussler 2020).

> *A "machine may be 'trained' by a 'trial and error' process to acquire one of a range of input-output functions. Such a machine, when placed in an appropriate environment and given a criterion of 'success' or 'failure' can be trained to exhibit 'goal-seeking' behaviour." (Minsky 1955)*

The geopolitical setting of the Cold War encouraged the United States and the Soviet Union to bolster their surveillance and aerospace capabilities. Silicon

Valley, presently the home of GAFA (Google, Apple, Facebook & Amazon), is an indirect result of this propensity (Pichot 2019).

The later invention of the semiconductor, silicon, in particular, a crucial element in the manufacturing of electronic circuits, gave rise to the personal computing revolution (ibid p.5). This period of tech-history also saw the term Artificial Intelligence first referenced during McCarthy's et al (1959) seminar when discussing machines with goal-seeking behaviour. However, the idea of complementing and amplifying human intelligence through machines was close to two decades old at this point (Bush 1945).

The world has come a long way since the early pioneers, especially with the advent of digital computers and the internet. Lee (2018), in his book *AI superpowers: China, Silicon Valley, and the new world order*, illustrates the development of AI as happening through four waves. The first, so-called internet AI, mainly revolved around algorithmicizing internet-users preferences to suggest content (ibid p.107). The second wave of Business AI makes use of already labelled company-specific data. Banks, insurance companies and medical institutions often possess data sets that sit on decades worth of information pertaining to for example credit histories, claims and fraud cases or archived medical diagnoses and these can be used in developing prediction models that outperform even top-notch experts in their analyses (Rüfenacht 2020 & Lee 2018 p.111). Thirdly comes perception AI, and as the name suggests it is about giving machines senses, a fusion of the digital and physical worlds. Algorithms in this category learn to group pixels in photos, or individual words in snippets of audio, in relevant categories to be able to recognize what is being shown or said. Or advanced facial recognition which amplifies the vulnerability of privacy and is an example of perception AI abuse by despotic governments (Andersen 2020). The fourth wave, still largely theoretical, Lee calls Autonomous AI. Synthesizing the progress of the previous three waves paves the path for machines capable of autonomous decision making (Lee 2018).

Today, Lee argues, we are transitioning from stage three to stage four. Internet, business and perception AI are all at work when it comes to the smartphone in your pocket. It knows what content to suggest to you, based on your digital footprint and behavioural data and you can even have a (somewhat) vibrant conversation with it. It is not fully autonomous, yet.

Three special considerations need to be given the conceptual demarcation of AI, echoing the points made by Heather Roff (2019). We need to be careful not to conflate AI with the more general concept of automation, doing so risks wrongfully aggregating risks and benefits in different ways, when more clarity is reached if we keep them separate. Automation need not involve AI, factory lines, water mills and dams can be automated without algorithmic intelligence that prompt autonomous goal-seeking behaviour. Secondly, AI can be applied to all sectors of the economy as well as in warfare, but that does not mean that it necessarily be best to do so. All problems are not the same and cannot be solved through the use of algorithms. Thirdly. AI is not in itself a moral agent, and so is not equipped to determine right from wrong. Roff argues that AI is a manifestation of 18th-century philosopher David Hume's "is-ought problem" – in that moving from a positive statement about what is to a prescriptive and normative statement about what ought to be is, if not impossible, very difficult. One cannot move coherently from one to the other. The collection of code that informs AI behaviour simply is, while the use of it involves questions of ought and human morality. "If we are clear-eyed about how we build, design, and deploy AI, we will conclude that all of the normative questions surrounding its development and deployment are those that humans have posed for millennia." (Roff 2019). So, this thesis is no evaluation of the normative effects of AI systems, which would require significant technical insight and experience, rather it is an endeavour of investigating whether the EU can be said to live up to the expectations of a normative power in AI governance.

It is up to policy-makers to make ethical judgements of how to allow the application and proliferation of this technology. But AI is an ill-defined term since the meaning ascribed to it has changed over time and as technology has improved. In general, we refer to computers and digital machines that can perform actions that one expects would require human intelligence. Reasoning and decision-making e.g. One way to understand AI is as a shift from humans programming computers into certain patterns, to the computers themselves learning how to act on their own. Systems that do this depend largely on ML techniques like "deep learning" and neural networks. This definition articulated by Ulrike Franke and Paola Sartori is very similar to what the European Commission itself uses

definition-wise (Franke & Sartori 2019). The European Union has adopted its definition of AI in the AI Act:

> *"Artificial intelligence system[s] […] can, for a given set of human-defined objectives, generate outputs such as content, predictions, recommendations, or decisions influencing the environments they interact with;" (EC 2021b p.39)*

The European Commission's definition of AI draws on McCarthy's and Minsky's idea of machines learning and applying goal-seeking behaviour stipulated by humans. However, this definition is highly contested since agreeing on any definition on conventional intelligence is hard "enough (Carrico 2018). But for the purposes of this thesis, it is deemed to capture the essence of AI.

2.3 Traditional notions of European power

If AI is a unique piece of tech, then the EU is a unique 'beast' in the international system (Manners 2002). During the cold war, rather than exert influence through military force it enjoyed a status due to its economic clout and regulatory influence. François Duchêne's notion of a civilian power Europe (CPE) paints the picture of the EC being a peaceful civilian oriented group of states with considerable economic strength who lack combined military power. Geopolitical self-interests are side-lined in favour of promoting norms. To clarify, archetypical civilian power has three main attributes. A) Economic power is used to achieve goals, B) diplomacy is the modus operandi for resolving international problems and C) international progress then, is measured by the willingness to defer to legally binding supranational institutions (Duchêne 1972).

Duchêne congratulatorily pats Europe on the back for its enlightened international status, not dissimilar from Kagan's later idea of Europe as a postmodern paradise (Kagan 2003).

Less than a decade later Bull criticized this notion of civilian power, arguing that it was ineffective and lacked self-sufficiency in military matters. The remedy? He suggested that the EC should become more self-sufficient in defence and

security through seven steps: “the provision of nuclear deterrent forces; the improvement of conventional forces; a greater role played by west Germany; more involvement on the part of France; a change of policy in Britain; careful co-existence with the Soviet Union; and careful co-existence with the United States.” Bull’s solution, unimaginable during the 1980s cold war, was to turn the EC into a military power Europe (Bull 1982).

It might not seem obvious that there would be similarities between Bulls’ and Duchênes’ notions of European power, but upon closer inspection, these two thought systems are closer than what might appear at first glance. Both value empirically measurable forms of power, be they market-share figures or number of fighter jets, over more intangible normative or ideational forms of power. Secondly, their shared understandings stem from the static environment of international relations contemporaneous to these ideas. The status-quo of power relations between the Soviet Union and the United States included assumptions about the fixed nature of the nation-state and the role of national interests. Thirdly, even though Duchêne wanted an end to the “age-old processes of war and indirect violence” and Bull “the regeneration of Europe”, they both intensely value European interests (Manners 2002 p.238). When the cold war came to an end, and the regimes whose ideology was perceived by its citizens to be unsustainable and undesirable, the structures within which these ideas came to be likewise, disappeared. No amount of force seemed to stem the popular uprisings in eastern Europe. This tells us something about the power of ideas over the power of empirically measurable force (ibid).

Autonomous AI remains a theoretical prospect, even though technology inches ever closer every year and many experts believe that Artificial General Intelligence or the singularity, systems capable of human-level thinking, will be achieved before the end of the century (Dilmegani 2021). As the technology continues to emerge and evolve there is considerable opportunity to influence its implementation and direction. Wherein acute ideational awareness of the potential risks and benefits associated with AI becomes important. Rather than simply counting the number of fighter jets or market-shares, or potentially in this case AI start-ups or petabytes of data, conceiving of power as something different, something more abstract will be valuable. If Europe has the potential to shape

perceptions of normalcy in AI governance, then that is a powerful Europe and something worthy of scientific enquiry.

2.4 The alternative - Normative Power Europe

Normative power is the anchor point of analysis for this project and so a presentation of the Normative Power Europe theory as outlined by Ian Manners in his acclaimed work *Normative Power Europe: A Contradiction in terms?* (2002) is in order. The nature and avenues of European power are, as the previous sub-section have shown, a hotly debated issue. The precise example of the Cold war ending and the Soviet regime collapsing is the segue Manners himself uses to assert the power of norms and ideas and transforms the discussion into something other than a dichotomy of military or civilian power. Both these thought systems mainly consider empirically measurable force and the capabilities that emanate therein. By complementing these with a theory that emphasises the power of an ideational and normative nature, the hope is that we get at something deeper. (Manners 2002 p.239). A hidden capability of sorts.

The Treaty on the European Union itself in fact stakes the claim that the organization is built on norms when it says:

> *"The Union's action on the international scene shall be guided by the principles which have inspired its own creation, development and enlargement, and which it seeks to advance in the wider world: democracy, the rule of law, the universality and indivisibility of human rights and fundamental freedoms, respect for human dignity, the principles of equality and solidarity, and respect for the principles of the United Nations Charter and international law." (TEU 2012 Article 21)*

The EU sees itself as built on benevolent norms and points to some fundamental ideas like human rights and dignity, democracy and the rule of law, NPE however, goes further than that, by suggesting that its constitutional norms represent crucial constitutive factors in themselves that determine its international identity. And by extension what leads it to act (Manners 2002 p.241, Groothuisen & Niemann

2012). This in turn rests on the assumption that the most important factor shaping the EU's global role is what it is and represents as opposed to what it does. Not entirely dissimilar to Bretherton and Vogler's concept of presence housed in their wider discussion of EU actorness (Bretherton & Vogler 2013). That is not to say that Manners thinks the role economic power, as some hold to be what drives civilian power (Carr 1962) or military power (Bull 1982) plays in European power projection to be insignificant. Simply that this dichotomy needs to be augmented with a focus on normative power of an ideational nature where common principles and a willingness to disregard Westphalian conventions sets the EU apart (Manners 2002 p.239). Similarly, this paper makes the case that strictly investigating empirically measurable force of the EU in AI misses something interesting.

What is the European Union according to Manners? Historically it emerged from the ruins of a Europe ravaged by the second world war. The nations of this war-torn wasteland opted to pool their resources together to disincentivize future wars of conquest against one's interdependent neighbours (Sjursen 2006). Supranational tendencies have increased over time. Today it is sui generis, combining elements of supranationalism and intergovernmentalism, making it a hybrid-polity (Manners 2002). It is this combination of historical trauma and one of a kind constitutional framework that bolsters a commitment to set universal norms, mainly those defined by the UN, at the very core of the EU's identity in relation to its member states and third parties (ibid).

Normative power can only be conceptually defined in contrast to other entities in the international arena and its member states (Diez 2005). Manners himself asserts that it is the relationship itself is the vessel of influence as it enables the normative actor to shape the actions of a third party to do what it otherwise would not have done. Manners exemplifies this in the 2002 paper which is a case study on the abolishment of the death penalty, and later in 2008 about the adoption of children´s rights (Manners 2008). It is suggested that these policy changes do not necessarily rely on military power or economic clout as the instigator of change. Rather it is norms and their promotion that do this, something which traditionally has been achieved through forceful persuasion or economic pressure. This gives the EU the ability, Manners argues, to shape what is considered normal in international relations. However, Sjursen (2006) and Forsberg (2011) think it

unlikely that the EU has this ability without any empirical force to back up its normative ambitions. Prescribing a certain value system to parts of the world, while at the same time respecting local context seems impossible if you are unable to point to empirically measurable force as the basis of power. Manners needed the answer to two questions: What norms are central to a Normative Power and how can it spread those norms without the direct or indirect use of more traditional forms of power?

Manners points to five 'core norms' of NPE which he finds in the acquis Communautaire, the aggregate legislation, legal acts and court decisions that together form the body of European Union law. Article 21 TEU, as quoted previously, embodies most of these in turn. **Peace** and **liberty** are the first bricks in the mosaic of norms, tracing its origin from the aftermath of World War Two. Historically **democracy, rule of law and human rights** come second as a habit of perception from the dichotomous world order during the Cold War, where the liberal west opposed the totalitarian Soviet east (Manners 2002 p.242-244).

To this list of five founding norms, Manners adds another four minor norms derived from the EU treaties and EU practices, although these are far more contested. As a response to the wave of liberalization in the 80s **social solidarity** became an integral norm. The return of political identities in the 90s gave **anti-discrimination** a place as a core norm. The third norm is **sustainable development** and the fourth **good governance** (ibid p.243).

He then seeks to find an answer to the conundrum of how EU norms are diffused without the support of traditional power. Contagion is the unintentional spread of ideas from the EU to other political actors. Through leading by virtuous example, it can export its experiment in regional integration. Informational diffusion is the result of strategic communications through e.g., new policy initiatives and declaratory communications. The White Paper: On Artificial Intelligence (EC 2020e) and the Declaration of Cooperation on AI are clear examples of this. Procedural diffusion happens when the EU establishes a relationship between itself and a third party, such as inter-regional co-operation agreements, membership of an international organization or enlargement of the EU itself. Transference denotes a kind of diffusion that takes place when the EU exchanges goods, trade, aid or technical assistance with third parties through largely substantive or financial means. The carrot and stick of European foreign

policy consisting of financial rewards contra economic sanctions which allows it to export norms. Conditionality as a way to cement norms happens in both procedural diffusion and transference. The EU can also, by virtue of its sheer physical presence in the form of EU-affiliated organizations, institutions or delegations, export norms which Manners calls overt diffusion. The latter is not exclusive to a normative power. The cultural filter, like an umbrella over it all, affects how well norms are adopted and learned or rejected in other countries (Manners 2002 p.244-245).

Manners launched his NPE theory in 2002 but the world, and the EU with it, has changed markedly since. The above section has introduced NPE in its originally conceived format what follows is a discussion of the limitations and a summary of scholarly objections to NPE, to provide a more contemporary understanding of NPE, which is also the one applied in the code-book of this thesis.

2.5 Criticisms of Manners NPE

Manners' description of how norms are diffused has been criticized for putting too much emphasis on the EU as a passive norm entrepreneur. Haukkala suggests that the EU actively promotes its norms through leading by example (Haukkala 2008). Both Manners and Haukkala maintain that enlargement is one of the main ways the EU diffuses norms. The latter argues that this form of norm-entrepreneurship should not be understood as a side-effect, but rather the tool of norm promotion. Regionally the EU is a normative hegemon that uses a hybrid of economic and normative power to establish sets of, often asymmetrical, bilateral relationships which act as the channels for the active diffusion of values (ibid p.1602). So passively the EU enjoys a status as attractive for future potential members, while once in the accession-process the EU can take a more active role and exercise conditionality through for example the Copenhagen criteria (ibid.).

Hyde-Price (2016) criticizes Manners from the neo-realist camp, arguing that EU member states use the supranational level as a lever for accomplishing their own foreign policy goals. Echoing the traditional realist mantra that self-

actualization and associated interests come before common EU goals in international relations (Hyde-Price 2016). Manners instead argues that the EU, wired to act normatively because of its status as a hybrid-polity, uses this very same influence to shape foreign policy interests of the member states that constitute it (Manners 2002). It does so with a mix of hard and soft power (Ekelund 2019, Haukkala 2008).

Diez reminds us that the desire to export norms is not unique to the EU. Historical- and contemporary powers also diffuse their norms. He fears that normative power Europe theory cements a discourse that establishes a particular identity for the EU through turning third parties into 'others' and represents the EU as a unique force for good (Diez 2005). This mirage of the EU as uniquely benign might stem from a subconscious euro-centrism and a lack of self-reflection – something pointed out by several of the scholars identified in the literature (Diez 2005, Ekelund 2019, Persson 2017, Staeger 2016). Bicchi (2006), taking a distinctly different ontological stance of doubting the universal nature of norms, criticized NPE as being fundamentally misaligned on similar grounds. In summary, the EU may be seeking to spread its norms to third parties to further its influence, rather than idealistically promoting universal norms. If rhetoric becomes inconsistent with policy and the EU unjustifiably is put on this 'high-horse, its credibility as a normative power can be called into question (Nicolaïdis & Howse 2002)

Another recurring criticism is that the framework created by Manners is empirically unviable, or at least unclear. Sjursen recommends more systematic empirical research in discovering whether the EU acts according to norms, or rather from self-interest. But in this endeavour, Sjursen argues that the term NPE lacks precision, especially when it comes to precise criteria and standards that allow for empirical analysis of the concept (Sjursen 2006 p.236). It seems as if a large part of the literature, notably in the early days of NPE, takes the fact that the EU acts normatively for granted and hence rather engaged in conceptual questions. There have been a number of empirically-oriented works on the topic (Whitman 2011, Tocci 2008, Niemann & De Wekker 2010). For this thesis Niemann & De Wekkers contribution *'Normative power Europe? EU relations with Moldova*' is particularly interesting.

2.6 The Framework of Niemann & de Wekker

Niemann & de Wekker tackle the conceptual fuzziness of normative power when they examine it along three dimensions in their investigation into EU relations with Moldova. Breaking up the concept of normative power into clear categories that need to be lived up to if Europe should be able to claim the title of normative power. Firstly, **normative intent** is about how genuine the EU's normative commitment is. Secondly, **normative process** refers to whether the EU is pursuing an inclusive and reflexive normative policy. And thirdly, **normative impact** gets at the heart of Manners idea, namely whether the EU has the ability to actually shape conceptions of normalcy (Niemann & De Wekker 2010 p.7-11).

Normative intent: how serious/genuine is the EU normative commitment?

They argue that this question is relevant in terms of the EU's self-image/identity: as a real normative power, the EU would have to be genuinely promoting its norms. An actor that is normative because of the belief in the goodness of the norms themselves rather than pursuing a self-interest agenda cloaked by high and mighty normative rhetoric. (Niemann & de Wekker 2010 p.7). The authors do not however stipulate that self-interest and norms exclusively exist in a dichotomous relationship. Norms and interests are not subject to an either/or rationale, but that interests and normative concerns tend to go together. This muddies the waters somewhat and the challenge when asking this question is to shed light on how genuine the EU's normative commitment truly is. Niemann & de Wekker identify the genuineness of normative commitment through the use of four normative criteria:

a) First, it is to be asked whether EU/universal norms are at the centre of relations with partner countries, or if norms are peripheral to the EU's engagement. This helps determine how seriously the EU takes norms to be. If norms are not centre stage, genuine normative involvement by the EU is less likely – ergo less likely to be a normative power. (Niemann & de Wekker 2010 p.7)
b) Secondly, we can ask whether the norms in question serve or hurt EU interests. For if the norms conflict with self-interest we have a powerful

indicator for the prioritization of said norm whatever the cost. Which indicates genuine normative commitment. (ibid. p.7)

c) Thirdly, one can examine the extent to which the EU communicates consistently and the degree to which it uses double standards. Double standards, i.e. the EU applies the same standards to third countries as it does internally, would suggest that norms do not constitute the most important basis for decision-making. (ibid. p.8)

d) Finally, to ascertain the normativeness of the EU, we can invoke coherence, the lack of which further exacerbates the problems of an inconsistent policy. "Coherence goes beyond consistency" (Niemann & de Wekker 2010 p.8) and is about the connectedness of claims through shared principles. The authors suggest that policies can be justified by utility-value-, or rights-based arguments and principles and that coherence between different justifications is a significant legitimising factor. Value- and rights-based arguments/principles seem somewhat more in line with a normative power than utility-based principles. (ibid. p.9)

Niemann & de Wekker's notion of normative intent is very similar to Manners'. Indeed, it places the same analytical emphasis on normative principles. Both camps deal with consistency, coherence and legitimacy. However, Niemann & de Wekkers "normative intent" places more emphasis on the aspect of consistency over coherence, distinguishing the two categories from one another and stake out operationalizable categories, which this paper benefits from (ibid p.7-9).

Normative process: does the EU pursue an inclusive and reflexive foreign policy promoting universal norms (or rather an 'our size fits all' approach)?

This category deals with the EU's propensity to include external input, criticisms and self-reflection in the process of decision-making. Openness to learning can thus be seen as an indicator of the EU's virtue and "goodness". The EU furthermore risks acting in a Eurocentric manner unless the norms it promotes are universally recognized (namely those emanating from the UN system). An exclusive and unreflexive policy informed by the conviction that the EU's

experience is a lesson for everyone the authors term an ‘our size fits all’ approach (Niemann & de Wekker 2010 p.9).

Besides reflexivity, the author request inclusiveness as a criterion for normative power. As other academics have suggested, there is a fine line between giving a voice to – and speaking for others (Bicchi 2006 p.289). The question here is whether the EU takes into account the views of those who will be affected (Niemann & de Wekker 2010 p.9).

Promoting universal norms further cements the EU as a potential normative power, that is if that is what it does. The self-perception as a force for good risks being limited to particular contexts unless one promotes universally applicable norms (Niemann & de Wekker 2010). Habermas suggests the principle of universalisation for evaluating the validity of norms. This principle states that “all affected can accept the consequences and the side effects its general observance can be anticipated to have for everyone’s interest” (Habermas 1990 p.65). While this is a sound consequentialist principle, it is extremely difficult to operationalize and seems like a pitfall of political philosophy. So instead, this paper draws on Manners, Niemann & de Wekker and refers to universally applicable norms as those who are generally acknowledged within the system of the UN.

Normative impact: the development of norms in third countries: does the EU have the ability to shape conceptions of what is ‘normal’?

Without any actual external impact, the EU’s normative *power* would be substantially limited. This gets to the heart of Manners’ original definition of NPE, namely “the ability to define what passes as ‘normal’ in world politics” (Manners 2002 p.236). Niemann & de Wekker state that the impact aspect of the framework can be taken to emphasise the *power* side of NPE, evaluating actual impact rather than intention (Niemann & de Wekker p.10).

The novelty of AI governance as a concept and the infancy of the politics surrounding it makes it more difficult to assess the true normative impact of the EU. This is a good avenue for future research to explore. However, what this paper argues, is that normative impact on member states can be assessed to a certain degree and could indicate a regional normative impact and potentially lead us to a tentative conclusion that the EU is exercising normative power. Furthermore, the AI Act can be compared to the impact of the GDPR. Comparing

the impact of the GDPR could provide some insight into the normative impact of European digital legislation. And the GDPR certainly has had a massive international impact, as the new global standard for data privacy (Bradford 2020).

More recently Manners himself also attributed and identified supposed attributes of the normative power ideal type (Manners 2009a & b). He holds that normative power uses normative justifications for its actions that are legitimate, persuasive, socializing and promotes principles of partnership and ownership (Manners 2009a p.1-4). The earlier mentioned Forsberg (2011) also developed a framework along similar lines when he argued that "there are at least four different mechanisms through which normative power is exercised: by persuasion, by invoking norms, by shaping the discourse and by leading through example" (Forsberg 2011 p.236). De Zutter tried "spotting a normative power" by constructing four criteria: material condition, identity, role, relational dimension and impact (De Zutter 2010 p.1117).

Here this thesis agrees with Niemann & de Wekker in that NPE, as conceived by Manners, provides a useful additional concept to the body of work on the EU's role in the world, especially since it highlights previously under-explored ideational and cognitive aspects of European power (Niemann & de Wekker 2010 p.2). Furthermore, it goes beyond the now mundane questions of whether the EU even has a foreign policy. It contends in a similar arena to, and builds on literature concerning EU actorness in world politics (Jupille & Caporaso 1998, Bretherton and Vogler 2006). But it remains empirically hard to pin down and many works lack a systemic formulation of the most important indicators of NPE. How do we recognize, and what qualifies as normative power? The three indicator categories of Niemann & De Wekker, which they derive from Manners original work and have overarching similarities to the other contributions mentioned. Because of their clarity and operationalizable features they are used in this thesis.

3 Methodology

3.1 Introduction

The next section intends to present an overview of the case at hand, this thesis' research design and the methodological approach considered able in answering the research question(s). Since this paper uses qualitative content analysis, a succinct description and arguments for why this method fits the aim of this thesis is provided below. Subsequently, the codebook, which was constructed from both theoretical expectations and through inductive reflection in tandem with the empirical material - to adapt the normative power Europe theory to the field of AI governance – is described. A justification for the selected material and data collection follows suit and the final section provides an overview of the coding scheme used which provides insight into the analysis process.

3.2 Case description

It becomes essential to demarcate exactly what is meant by the EU's AI strategy, which is the subject of this paper, for the reader to understand what is being investigated and what may be left out. This subsection will elaborate on the case and begin to describe what documents were assessed to have value in answering the research question. As so often with political agenda-setting for the EU, the beginning of the AI strategy can be said to have come from the UN and the AI for Good Series held in 2017. It identified sustainable development as an important priority, something which this thesis returns to in the analysis (see 4.3). Thereafter the Communication on Artificial Intelligence (2018) staked out the political

ambitions of the Commission and the Declaration of Cooperation on AI (2018) showed that a clear majority of member states thought cooperation and harmonisation in this field to be important. A critical review of AI policies and reports between 2015 and 2018 by Vesnic-Alujevic et al. (2020) showed that there exists strong normative agreement between supranational actors on the EU level. However, the authors found, besides the broad agreement on ethical ideas, that these documents do not necessarily consider the feasibility of regulation and practical implementation. In 2019, following previous outlines for AI in Europe, two deliberating bodies published the *Ethics Guidelines for Trustworthy AI* (2019). The High-Level Export Group on Artificial Intelligence (HLEG-AI) wrote them in close cooperation with the European AI Alliance, a multi-stakeholder forum assembled by the Commission in 2018 to provide feedback on initiatives on AI. These ethical guidelines as well as the rather practically oriented assessment tool, *The Assessment List for Trustworthy Artificial Intelligence* (2019), framed the discussions for the next phase of actual legislative action. In early 2020 the Commission published the results of these consultations in its white paper *On Artificial Intelligence: A European Approach to Excellence and Trust*.

Preceding the white paper, the EU was under considerable pressure from the uncertainties around Brexit (Berger 2018). This was also a time where other competing actors, in rapid succession to one another, formulated ambitious AI goals with comprehensive investment plans that stoked fears of the EU lagging. During this time China loudly proclaimed that it would be the world leader in AI by 2030 (Franke 2020).

The white paper made clear that we could expect upcoming regulatory action from the Commission and presented the key elements of the future framework, amidst a raging global pandemic. Among these were the mandatory legal requirements, derived from ethical principles as articulated by the HLEG-AI, that should be imposed on high-risk AI systems. The white paper was followed by an open consultation process that involved more than 1200 stakeholders including citizens, academia, member states representatives, civil society and industry to reach where we are today with the EU AI Act (Gaumond 2021).

The AI Act is given its own section under analysis 4.4 for good reason. The European Data Protection Supervisor claims that it is the first proposed regulation in the world that provides a legal framework for AI. One could argue that the

passing of the US National AI Initiative Act as law in January of 2021 precedes it (Floridi 2021 p.216). So even if it's not a world first categorically, it is a world-first in terms of the ethical strictness imposed on AI systems. And even if the process may take several years, (the GDPR took 6 years to be passed in OLS and two more to come into force) it is a massive step towards something which likens the GDPR in scope, namely that it seeks to regulate all applications of this technology. It is an acknowledgement of the importance of digital and a heeding of the experts calls for increased attention towards the rise of AI (Floridi 2021). As a potential normative tool, the AI Act, this paper argues, should be given special consideration.

This form of reasoning warrants a counter-question. If the AI Act is as important as this paper argues – then why not limit the material to this document alone? The synthesis of previous reports, policy papers, guidelines, and press releases paint a clearer picture of the supposed normative ambitions of the EU. Precisely because the AI strategy is so new and since there are varying notions of how to categorize it, see section 2.2, this paper thinks it prudent to begin by pinning down precisely what it is in terms of norms and values. Without this explorative stage, key-takeaways risk getting lost in the technical jargon of the AI Act. This too is an acknowledgement of the author's relative unfamiliarity with the strictly technical aspects of AI.

This thesis will focus on two primary areas. It will attempt to pin down the EU's understanding of- and stance towards AI as represented in the wider strategy documents. We can understand the four aforementioned documents, the coordinated plan, declaration of cooperation, white paper, and the Artificial Intelligence Act, as a pathway for the future of European AI. But do these documents prove that the EU is a normative power? Has Europe struck out on a special path of AI governance and will NPE as conceived by Ian Manners (2002) and operationalized by Niemann & De Wekker (2010) shed light on this?

3.3 Research design

The thesis has a two-pronged approach, starting with an explorative case study of what the EU's AI strategy is and what norms are conveyed, thereafter debouching to a descriptive and comparative design - aiming to highlight the differences or similarities between the ideational documents and the more actionable AI Act. In the explorative step, the researcher can formulate new categories out of the material and theory together inductively, so-called inductive category development (Mayring 2014 p.12).

The initial use of an explorative case study is an acknowledgement of the complexity of the EU system and a choice motivated by the novelty of the field of research. The development of policy goals for value-oriented AI governance has only just begun. As the dates of publication in the appendices indicate, the phenomenon in question is a narrow happening over a relatively short period of time. 2018 – 2021 to be precise. Reading only one or a few official documents about AI governance in the EU, however, runs the risk of missing the bigger picture or context in which the norms this project seeks to make visible are present. And seeing as if this paper can achieve an almost exhaustive list of official EU documents written on the topic of AI, why shouldn't it? Reasonably there are a plethora of actors on the supranational level that input and project onto the AI governance agenda, and it seems likely that an arbitrary selection of just a few documents will misrepresent this ideational stage of the policy process.

Secondly, the field is very new, as is the application of AI-related technologies, which itself perhaps conveniently, limits the amount of EU documentation on the area. AI governance ambitions on the supranational level first appeared explicitly in the 2018 Communication on Artificial Intelligence. Although the EU can trace the start of this policy process to the AI for Good Summit in 2017. The report from the summit can be arguably be said to be the first initiative that emphasises the benefits of AI if used with sustainable development as a top priority (ITU 2017). But crucially it has now morphed into its own EU initiative as part of the larger digital agenda put forward by the von der Leyen Commission.

The final step of analysing the AI Act will use the results gathered from the initial explorative stage of the research to determine whether the normative ambition is still present and how strongly. This second step utilizes the category system of codes derived from both the empirical material and theoretical framework to answer in full the main research question – Does the EU intend to act the way a normative power would in the field of artificial intelligence governance?

3.4 Research method

A qualitative document content analysis approach will allow for a direct investigation of the research question as well as mitigate the problem posed by the fact that the field of research is just coming of age (Mayring 2014). Kuckartz (2019) argues that:

> *"Working with codes and categories is a proven method in qualitative research. QCA is a method that is reliable, easy to learn, transparent, and it is a method that is easily understood by other researchers. In short, it is a method that enjoys a high level of recognition […],"*

One of the inherent strengths of QCA is the default high ecological validity that it provides, especially in a field like "norms in AI governance" which in for example a solely interview-based study could be compromised. Ecological validity, Bryman et al (2016) argues, is ensured when the researcher avoids interfering with the social setting's natural state. But this is also a choice made based on necessity since the access to "inner-circle" settings required to successfully process-trace are hard to gain and the topic at hand is that which only a select few are privy to. The choice of material and method, therefore, reinforce one another.

However, QCA researchers may be placed along a broad spectrum "ranging from an orientation akin to positivist science to one more akin to art and literature" and between these extremes exists a large space of researchers who employ elements of both camps (Ellis & Ellingson 2001 p. 2287).

3.4.1 Qualitative content analysis

Through using qualitative content analysis at both stages of the project, the researcher ensures genuine comparability between the different empirical material. Content analysis involves identifying themes, word categories and possible relationships between categories observed in the empirical material. For this endeavour, a codebook has been key. A codebook may be driven either by the theoretical framework or the empirical material, but as Lynggaard (2019) points out, it is most often constructed with both in mind, which this project does.

Mayring (2014) points to two ways that the researcher can code material for qualitative content analysis: Firstly, the deductive approach infers categories from the theoretical framework or previous research. The categories are created before the process of analysis begins, to extract determined elements from the chosen material. To do this one has to define precisely what would make a unit of analysis fall under a certain category. Secondly, with an inductive approach, the categories emerge from the contents of the raw data rather than derived from the theoretical framework. The text analysis is conducted a priori (Kuckartz 2019 & Mayring 2014).

In this project the codebook will evolve in tandem with the empirical material and the theoretical framework, utilizing a hybrid of deductive and inductive coding. The reason for this is that the proposed theoretical framework of normative power (Manners 2002), with operationalizations derived from Niemann & de Wekker (2010), will lack in-depth with regards to AI policy. More codes will have to be added inductively to make the norms and values in the EU AI strategy visible for examination.

Since this thesis deals with a large quantity of text it uses the Atlas.ti to reduce the complexity and distil concepts, themes and rationales and also allows for visualization in a way that manual coding cannot (Kuckartz 2019 & Mayring 2014). But maybe the frequency of certain codes might also be of importance and show what ideas are more or less seismic to the overarching AI strategy? This would be a so-called manifest analysis where one uses keyword search tools to visualize patterns and the evidence is that content that can be directly seen.

The second approach is latent analysis. This style attempt to dig deeper and refer to the underlying meaning of content, such as interpretations of text. Latent

analysis, which is this thesis' primary approach, has the consequence that single occurrences of concepts, phrases or words can have a large impact on the analysis. These would otherwise be ignored in a manifest analysis, which risks losing out on nuances and findings that could prove interesting (Schreier 2019). Even though the manifest approach risks missing out on the nuances and subliminal nature of values that are implied in the text, this thesis uses the frequency of a code and code co-occurrence as a way to complement in-depth narrative interpretation and to visualize the space given to certain norms over others and their relationship to one another.

The data produced for this paper has been generated through reading all the chosen documents (see 6. Bibliography) through a software called Atlas.ti. Atlas.ti is a qualitative research tool that can be used for coding and analysing transcripts & field notes, building literature reviews, creating network diagrams, and data visualization (Atlas.ti 2021). This program has been instrumental in the analysis, since coding is done, although manually and as you read, much more efficiently than by hand and the coded paragraphs can be easily compared and entire document-sets investigated for code co-occurrences.

3.4.2 Material

The three main documents mentioned in section 3.2, if taken alone, risk leaving out potential conclusions about the EU as a normative power. In order to provide a more detailed description, with more depth and reliability, this thesis includes a total of 29 documents, of which five are EU-commissioned reports and 24 are policy papers of varying sort. By including these the data set hopes to avoid biases and instead image the AI debate in its entirety. The documents selected include, but are not limited to: proposed regulations, brochures, fact sheets, press statements, white papers, communications, declarations etc. Since this is a text-based endeavour, the illustrations found in the brochures will not be given special attention. But with the mix of "simple" and "complex" content and different institutions, the hope is that this thesis will cover a broad spectrum of institutional input and in the end, be able to say something meaningful about how the EU as a whole approaches AI and whether it intends to act normatively or not.

The attentive reader would, after glancing over the bibliography, have noticed a certain bias in favour of documents released by the European Commission over other institutions. The Commission is given this "extra space" since it holds the right of legislative initiative and therefore has the most agenda-setting power in these early stages of legislation. This is an acknowledgement of how the EU system works, but also a conclusion reached after reviewing the relevant literature in the field. The Commission is often at the centre of attention for research on AI governance in the EU (Parviala 2019 & Berge 2021).

This problem of overrepresentation increases since the Artificial Intelligence High-Level Expert Group (HLEG-AI) was launched by the Commission. There is an obvious institutional bias in the proposed dataset. However, including a plethora of institutions was, and remains an ambition. The European Parliament (and its Research Service), the European Economic and Social Committee and the European Council. The actual normative influence, interdependence or tracing of the processes of specific institutions will not be assessed in this thesis but can be a fruitful avenue for future research.

Furthermore, the material chosen limits the window of analysis to the years 2018 – 2021 and so the thesis does not claim to be charting the institutional attitudes towards technology before or after this, although comparison to the GDPR is made. This is mainly because the legislative process that has led us to the Artificial Intelligence Act started in 2018, although one could argue that calls for initiatives like these came at the UN AI for Good Summit in 2017, this thesis restricts itself to analysing the EU and not the UN.

3.4.3 Data analysis & codebook

The thesis delimits itself to a meso-level analysis, that being institutions within the European Union and makes no claims to chart for example the AI landscape in Europe nor any individual member state (Jilke et al 2021). Merely to what extent the selection of European Union institutions indicate that they intend to act in a normative way in the field of AI. This is done by using the deductively and inductively derived codebook found below.

As mentioned previously this thesis approaches the material with a qualitative content analysis. Mayring (2014), however, portrays this as a hybrid between quantitative and qualitative content analysis designs, borrowing approaches from both. He suggests the method be re-named "*category driven qualitative oriented text analysis*", which acknowledges the role categories and coding of the traditionally quantitative content analysis method employs. The frequencies of coding play a role, although minor, in pointing the analysis towards finding patterns. This helped reduce the complexity of the rather large selection of material. With the help of codes, the researcher can label entire paragraphs that in the end make patterns discernible that hopefully can provide indications of normativity. While the frequency of codes plays a minor role, the overall project can be characterized as a latent content analysis, which has a better chance of identifying subliminal norms as well as avoiding the risks of arbitrary coding.

Categories	**Groups**	**Definition**	**Code**
NORMATIVE INTENT	Central value	*Are norms at the centre of the AI concept as in the arguments of the EU institutions? This is indicative of a normative power.*	AI transparency AI safety Antidiscrimination Democracy European values Good Governance Human rights Liberty Peace Rule of law Social solidarity Sustainable development
	Peripheral value	*If the norm has a peripheral value that is a good indication of less genuine normative commitment. It is however a good indicator for the relevance of normativity if norms compete or interfere with interests. Especially if the institution has implemented it, despite economic/political costs.*	AI liability

	Self-interest	*If norms or interests compete with one another we can argue that the norm is considered relevant. The institution seems willing to promote the norm no matter the economic or political cost.*	Competitiveness Threatens core EU values
	Double standards	*Double standards towards different actors suggest that the EU does not consider norms to be the basis of decision-making. Norms have to be followed by actions in a consequent way.*	Intellectual protectionism
	Coherence	*The connectedness of claims or actions through shared principles. Utility-based principles as the basis of an argument are less of an indication of a normative power than a value/rights-based principle.*	Bigger opportunity than barrier Ethical AI Human-centric
NORMATIVE PROCESS	Reflexivity	*How inclined the institution is to change its behaviour and attitude when it is confronted with better arguments. Moreover, the institution is not applying an "our-size-fits-all" approach.*	[Acknowledges: Complexity Future challenges Lack of action Need for data access Opportunities] Includes research Planned adjustment to change Value-related implementation
	Inclusiveness	*Are those who will be affected included in the decision-making process? Those actors whose normality will be affected are consulted and their views are taken into account.*	Consumer protection Civil society EU cooperation Privacy Public-private-AI partnership Public sector Promotes social solidarity
	Universal norms	*Universal norms are those acknowledged by the UN system and not only by the EU.*	AI for the greater Good Tackling Complexity

			Tackling inequality Promotes democracy Promotes sustainable development
NORMATIVE IMPACT	Normative change	*The degree to which the norms projected by the EU are being referred to in the political and media discourse. Norms, becoming part of the discourse, can be seen as the first sign of norm adoption and thus normative change.*	Implemented adjustment to change Initiate value related project
	Internalisation	*When norms are ascribed the same significance and meaning in different contexts and forums, then there is an increased probability that the relevant actors meant what they said.*	Implemented adjustment to change Independent research conducted

(Niemann & de Wekker 2010 p.6-12)

4 Analysis

4.1 Introduction

The EU self-identifies as having a strong scientific and industrial base to build on with world-leading research labs and universities, recognised leadership in robotics as well as innovative start-ups. And at the same time boasting a "'comprehensive' legal framework [that] protects consumers while promoting innovation" (EC 2018e p.19). These are some of the main conflicting ideals, namely innovation versus EU values. Hopefully, the analysis sheds some light on whether the EU will act in a normative manner with regard to AI. To answer this overarching question this segment has been sectioned into three parts. Firstly, a section highlighting distinguishing features of the EU's vision on the future of AI governance, compiled from the documents that preceded the AI Act. Secondly, central norm narratives will be extracted and cross-checked with the pre-defined and inductively gathered criteria for what can be called European values in the hopes of answering the second research question. Thirdly and lastly, the AI Act is given its own section, due to its weight as a normative tool and with it being the world's first proposed regulatory framework to harness the potential of AI, while minimising associated risks.

4.2 The European Union's approach to AI

Europe's approach to AI can be categorized along a few different dimensions, and taken together they underline the EU's normative commitments as expected from the theoretical background of this paper. To establish a globally competitive

European AI landscape the documents analysed call for EU-wide cooperation, both in terms of member-state cooperation and multi-stakeholder alliances on AI. The consensus seems to be that it cannot be up to single member states to regulate and push this innovation forwards. In fact, EU cooperation is the most frequently occurring code across the data-set, from the reports commissioned by the Commission to actionable documents released by the Parliament, Council and the Commission all refer to the importance of EU-wide cooperation in the field of AI. The data set seems to suggest that the Commission in particular advocates for strong ties between the member states, central stakeholders as well as EU forums. This could be explained by the rather lacking competence of the EU in fields of digitalization and technology, by virtue of its character as a hybrid-polity relying on multiple channels of input. The bias towards Commission documents in the bibliography partially explains this outcome as well.

> *"[The Declaration of Cooperation on AI] was signed […] to work together on the most important issues raised by AI; from ensuring Europe's competitiveness in the research and deployment of AI, to dealing with social, economic, ethical and legal questions […]" (Kritikos, 2019)*

The Declaration of Cooperation on AI, signed by 25 countries in 2018, is an instrumental document for the EU as well as in highlighting the value placed on intergovernmental cooperation. Competitiveness in AI cannot be achieved when it is the responsibility of single member states. This is echoed by for example France's national AI strategy when it expresses concerns about Europe becoming a "cyber colony" (Franke 20200). Furthermore, the multipolar world order, in particular the rise of China, is alarming to EU policy-makers, a situation further worsened by the UK, and their resources in AI, leaving the EU (Lee 2018).

The lack of authority is not the only factor making EU cooperation such a central concept in the EU's vision. The analysis discovered several reasons for the centrality of the concept and several alternative ways rather than simply resorting to intergovernmental policy-making. Coming to EU-wide agreement is difficult given the different technological and economic conditions that dictate the interests of the member states. But on the other hand, the material suggests that this diversity of opinion in itself is an asset. It promotes the image of the EU as a normative power because the vision of AI is based on consensus rather than top-

down enforcement. With different cultural backgrounds chiming in, inclusiveness is all the stronger. But the Commission sees the downsides of the different member states being so differently "able" in AI, so much to suggest tools for remedying the disparities.

One such tool seems to be the digital innovation hubs, below mentioned by the Commission's joint research centre:

> *"We suggest putting these hubs at the centre of local ecosystems comprising public administrations, local enterprises, educational and training establishments, and civil society." (EC 2018a).*

These hubs, proposed to be established in every member state, are an attempt at evening out the playing field, promoting the uptake of AI in less digitized countries and countering disruptive tendencies of the technology and "facilitate access of all potential users, especially small and medium-sized enterprises, companies from non-tech sectors and public administrations" (EC 2018f). Once more echoing the inclusive ambitions found throughout the data set. The EU seems to truly want to include a wide variety of stakeholders and less digitized member states in its goal of becoming a world leader in trustworthy AI.

When acknowledging future challenges related to AI, calls for further EU cooperation accompany these sentiments. These two codes co-occur regularly, hinting at a high degree of reflexivity as well as inclusivity present in the EU's AI vision. The Council naturally has a pronounced focus on including member states in the decision-making process.

> *"We call on the EU and its Member States to assess whether existing EU and national legal rules are adequate to take advantage of the opportunities and address potential risks that the use of digital technologies and AI systems in particular may create and to develop them further where necessary. […] where AI applications are used and their cooperation at European level should be enhanced" (Council of the European Union 2020 p.8)*

The first dimension of the EU's AI vision identified in the dataset can hence be ascertained to be EU-wide cooperation, between the member-states as well as

the different stakeholders involved i.e., civil society, academia, the private- and public sector.

> *"We should learn from the examples of successful internet companies and develop European data ecosystems bringing together the public sector, the commercial sector, academia and the third sector, and the general public." (EC 2019a)*

This already supports the notion of the EU as promoting inclusivity in its AI vision, not imposing a one-size-fits-all solution, but rather seeking common ground with the member states and trying to include relevant stakeholders. The EU communicates that AI is a matter of EU cooperation. The interest in harmonisation and cooperation is reciprocal. EU cooperation is necessary to ensure global economic competitiveness for the individual member states and the EU as a whole

The EU outlines a highly inclusive vision of AI. In terms of anti-discrimination, human rights and questions of social solidarity, but also its process-think. The AI High-Level Expert Group is one of the first actors to call for multi-stakeholder alliances to anchor innovation and technological progress in the hands of those affected most.

> *"We believe that sectoral multi-stakeholder alliances that foster trust across policy-makers, industry and academia as well as society, can help securing those investments, as well as steering them towards Trustworthy AI." (HLEG-AI 2019 p.49).*

The EU's High-Level Expert Group on AI, a team of experts from academia, industry and government, the whole purpose is to facilitate discussion and input on the future of AI governance in Europe. They do this by anchoring the policy debate with those stakeholders that might be affected by innovation in AI.

The reason for the EU's inclusiveness comes from an understanding of AI as socio-technical and all-encompassing. Several of the documents analysed in the data set acknowledge the future complexities, challenges and opportunities of this emerging technology. Some even liken it to the steam engine or electricity in order of magnitude and

Trustworthiness in AI represents a frequent and important theme in the analysed material. Its centrality is abundantly clear in documents like the *White Paper on AI* and the *Ethics guidelines for trustworthy AI*. Two central pillars of the EU AI strategy.

> *"A trustworthy approach is key to enabling 'responsible competitiveness', by providing the foundation upon which all those using or affected by AI systems can trust that their design, development and use are lawful, ethical and robust." (HLEG-AI 2020)*

The EU wants to foster an 'Ecosystem of trust' as well as 'ecosystem of excellence' (EC 2020e) and sees these as evolving in tandem. Trustworthiness in AI is described almost as a silver bullet, ready to give Europe the competitive edge by virtue of its AI systems being trustworthy for consumers. This ties back to Niemann & de Wekkers discussion on self-interests in conflict with the value-oriented implementation of policy. Competitiveness need not only be considered a self-interest, in fact, dichotomizing self-interest and normatively guided action is sometimes impossible as the two go hand in hand (Niemann & de Wekker 2010). The EU for example seems to see trustworthy and ethical AI as a competitive edge when compared to China/US. If European AI systems can be labelled and certified as particularly trustworthy with a high degree of consumer protection present, then the EU gains a competitive edge and a panacea for the, comparatively laggard, European start-up community. AI safety, liability and transparency are the main building blocks of this trustworthy AI, and transparency seems to be especially central. Natural persons should be able to understand how decisions are made by the system and that information must be accessible to those who use it (HLEG-AI 2019a).

Furthermore, AI safety, competitiveness, EU cooperation and EU values are often invoked together as guiding for the EU's AI vision, certainly in this quote from the Council on the presidency's conclusions on how to interpret the charter of fundamental rights in the context of AI:

> *"We therefore want to work together on European responses for digital technologies, such as artificial intelligence (AI). We want to ensure that the design, development, deployment and use of new technologies uphold and*

promote our common values and the fundamental rights guaranteed by the Charter, while increasing our competitiveness and prosperity. High levels of IT security must be maintained within a framework that is open to innovation." (Council of the European Union 2020 p.3).

The further development and promotion of AI safety standards and practices of AI transparency and support in EU and international standardisation organisations will, as argued, help enable European businesses to benefit from a competitive advantage and increase consumer trust (EC 2018e). Garnering trust in European AI innovation will be the competitive edge that sets the EU apart and will make consumers want to use European AI systems rather than 'untrustworthy' American or Chinese systems. The commission often criticizes these actors for slacking on their AI transparency.

"Among them, current ML algorithms display some of the inputs and outputs but do not understand fully what happens in-between, and how certain outputs, including decisions and actions, are derived" (EC 2018a p.9)

While the demand for trustworthy and ethical AI is a frequent and central concept in the dataset, it shows some overlap with certain self-interests of the EU. It shall strengthen the global competitiveness of the EU in AI-related technologies, both in practical application and research, and the "made in Europe" label should be on that instils trust, not for the normative value itself, but rather for the economic competitiveness of the Union,

"For the EU, it is not so much a question of winning or losing a race but of finding a way of embracing the opportunities offered by AI in a way that is human-centered, ethical, secure and true to our core values." (EC 2018a p. 12-13)

The Commission refers to the successes of the General Data Protection Regulation in setting normative standards around the world, hoping that the forthcoming AI framework will have a similar impact. The GDPR, opposed by many during its preparation, is perceived as a European asset and inspires many similar actions outside of Europe (EC 2018a p.10). That same regulation is an anchor of trust within the single market for data, it has a strong focus on the rights

of individuals, "[…] reflecting European values, and is an important element of ensuring trust in AI." (EC 2018f p.6). This echoes points made by both Haukkala (2008) and Manners (2002) of the EU as a norm entrepreneur. But the EU seems to know it's not at the forefront of AI. It may not be a race, as the quote above implies, but the EU is certainly afraid of losing its competitive edge in many economic sectors and regularly expresses concerns regarding brain drain. This makes it unclear if norms are a central value or if they are more peripheral to economic growth.

> *"[…] the EU needs to train more specialists in AI, building on its long tradition of academic excellence, create the right environment for them to work in the EU and attract more talent from abroad." (EC 2018e).*

The research sector is particularly incorporated in the EU's vision of AI. The representation of academia is of great importance for the European AI landscape. Many of the documents analysed stress the lack of research conducted so far, especially with regards to consumer experience and behaviour when confronted with AI systems (HLEG-AI 2020 p.6). Then one wonders what would happen in a scenario where the desire for economic growth or innovation directly comes into conflict with EU values. One suspects that the EU would sacrifice technological progress to protect central European values, like consumer protection. The research sector is attributed this central and prioritized role and, in a way, confirms the old stereotype of the EU as a precautionary technocracy (Recuerda 2006). This is very dissimilar from the United States' unregulated and comparatively liberal AI market, where rather than plunge into the unknown for the sake of technological progress, the EU prefers to base AI deployment on empirical findings from its own research initiatives. Several of the reports argue that if the risks of AI are unclear or not assessed, that one should desist from its deployment, as it may "impose unforeseen threats" (HLEG-AI 2019a, Delponte 2018). In her speech to the European AI forum in 2020, vice-president of the Commission Margarethe Vestager reiterated the primacy of European values, beliefs and rules over technological progress.

> *"I am often asked if all the benefits that digital solutions have brought to us during the crisis have changed the Commission's digital policy. Have we*

been too critical of tech companies, too careful and bureaucratic when it comes to digital technologies? My answer to this is that our objectives are more relevant than ever. The more we use and depend on digital technologies, the more important it is that these technologies are in tune with our values, beliefs and rules" (EC 2020b)

AI companies in the EU will have to contend with this common European, attitude towards technological change. The risk is that in seeking to regulate away all potential adverse effects of AI application that innovation and investment are stifled. Throwing the baby out with the bathwater by following the principle "when in doubt – regulate". But going by the theory, it serves to safeguard AI safety in a way that is coherent with a normative power. For if the EU strongly regulates the deployment in favour of fundamental rights and European values despite the costs associated with it, then we can ascertain a large degree of genuineness. These associations hint at another depiction of AI that was often invoked by the EU;

"[…] AI should be understood as a socio-technical system and should be assessed according to the society in which it has been created, while society's role in the development and applications of AI/ML should not be under-estimated […]" (European Parliament Research Service - van Wynsberghe 2020)

Frequent comparison is made to other technological paradigm shifts in history, e.g., the steam engine or electricity and in that sense, AI is a socio-technical system. It is expected that AI will influence life in many ways and result in a dramatic socio-economic paradigm shift. AI is not sector- or industry-specific technology, but rather it is recognized as a potentially (and to some extent already) universal technology that can interfere with all levels of human life (Servoz 2019). It is this universality and socio-technical aspect of AI that compels the EU to seek EU cooperation and to ensure trustworthy AI that is human-centric and promoting "ethics-by-design principles". An all-encompassing challenge warrants a holistic response.

"Crucially, humans should understand how AI makes decisions. Europe can become a global leader in developing and using AI for good and promoting a human-centric approach and ethics-by-design principles." (EC 2018f).

… and the EU makes no secret of its desire to export its way of doing AI.

4.2.1 Interim findings

The EU is fundamentally torn between two varyingly conflicting ideals, namely technological progress for the benefit of economic growth and the preservation and safekeeping of European values. The EU is very concerned with its place in the global competitive landscape and wants to achieve global competitiveness through maintaining a strong research sector that promotes the deployment of trustworthy AI.

EU cooperation is central to the EU's vision of AI. Including a multi-stakeholder approach combined with cooperative inclusion of the Member States portends to a degree of coherence between acting out of self-interest and its self-claimed normativity. These two need not be dichotomized and separated from one another, as norms and self-interests often go hand in hand.

Case in point: Trustworthy AI is a silver bullet of sorts; in that it ties together the self-interests of the EU while also enabling norm entrepreneurship. If consumers can trust the 'made in Europe'-label more than its competitors, then the EU gains a competitive edge and propagates its norms. Comparisons are often made to the GDPR and the impact it had on global personal data standards (Servoz 2019). This is to be achieved through a strong research sector embedded in the socio-technical vision of AI. However, this frequent representation of the research sector and innovative caution leads one to suspect that technocratic stereotypes are being entrenched. This might hamper a potential European start-up culture. On the other hand, the EU is afraid of its young AI talents being brain drained away to its competitors, showing signs of intellectual protectionism and exhibiting double standards in contradiction to what a genuine normative power would.

However, a genuinely normative power would act mainly for the sake of pursuing its normative agenda rather than economic interests. In order to pin

down just how genuinely normative of an actor the EU is this paper will now distil the norms from the EU's human-centric approach. Does it build on European values? The theoretical expectations would have us presuppose that the norms present to be aligned with the EU treaties and also in a central position compared to self-interests.

4.3 What are the norms? Is the EU a normative power?

As discussed in the theory section 2.4, the EU follows a set of norms present in the treaties. These are the basis for its actions, both internally and on the international stage (Manners 2002, Diez 2005). The previous section has shown that the EU follows a human-centric vision on AI, of which inclusiveness, trustworthiness and cooperation are prominent features. Moving on in attempting to answer the second research question: "What norms are promoted by the EU institutions with their concept of AI?", we can formulate more empirically grounded arguments for or against the EU's self-proclaimed normativity when it comes to AI. The analysis has found significant co-occurrences between EU values as defined by Manners (2002) within the examined documents. This is particularly true for the documents that explicitly deal with questions of a normative nature, like the *Ethics Guidelines for Trustworthy AI* while more investment/liability-oriented policy documents, like *Policy and Investment Recommendations for Trustworthy AI*, rarely referenced values other than underlining the EU's commitment towards 'trustworthy and human-centric AI' as laid out by the overall strategy (HLEG-AI 2019 b & c).

> *"The use of AI systems should be given careful consideration, particularly in situations relating to the democratic processes, including not only political decision-making but also electoral contexts (e.g. when AI systems amplify fake news, segregate the electorate, facilitate totalitarian behaviour, etc.)". (EC 2018h)*

Nevertheless, democracy is the value that most often co-occurs with the other pre-defined EU values, suggesting that democratic ideals are deeply embedded within the overall vision and that the EU has the intent of promoting these ideals. Democracy can thus be understood as an "umbrella concept" on which other normative dimensions depend and are connected to. A central narrative is that AI seems to be interpreted as a potential threat to democratic mechanisms. The EU seems to suggest that the public sector can benefit from algorithmicizing many of their processes, but automated decision-making in certain crucial public services might threaten EU values like social solidarity, human rights and anti-discrimination. All of which seem to be included under the umbrella concept of democracy. The EU's AI vision is promoting democracy-related values on many different levels, understandable given the universal and socio-technical aspects of AI.

> *"The EESC believes that technological development can contribute to economic and social progress; however, it feels that it would be a mistake to overlook its overall impact on society. [Job automation is increasing.] This is why the EESC would like to give its input to efforts to lay the groundwork for the social transformations which will go hand in hand with the rise of AI and robotics, by reinforcing and renewing the European social model." (EESC 2018)*

On its own legs, social solidarity was the most frequently occurring code. Even though this paper is not primarily engaged in a manifest analysis (counting word- and phrase-frequencies), it still tells us something. This is perhaps the single strongest norm, based on how frequently it occurs. AI is likely to impact the distribution of employment, and by extension wealth, through the automation of many previously manual jobs while also raising the educational requirements to enter the workforce (HLEG-AI 2019a). Concerns about economic-, gender- and ethnic inequalities leads the HLEG-AI in particular to express the desire to help vulnerable demographics. They encourage and propose a number of policies that are aimed at bracing the European social system for the perceived changes AI might bring. It sponsors coaching initiatives for Women in AI and encourages impact assessments for the goal of AI systems that are socially responsible.

"While digitisation is affecting the structure of the labour market in particular through the automation of middle-skilled jobs, AI could have a more significant impact on lower skilled jobs. If not addressed early and proactively, this may exacerbate inequalities between people, regions and industries in the EU." (EC 2018e)

This ties in closely with the EU's vision of AI as human-centric and truly seems to want to ensure genuine inclusion across all civil society, with the research sector once again taking a central role as norm-agents (HLEG-AI 2019b). Anti-discrimination is ever-present as a sort of appendage to the overall commitment to social solidarity.

"[…] Europe should strive to increase the number of people trained in AI and encourage diversity. More women and people of diverse backgrounds. including people with disabilities, need to be involved in the development of AI, starting from inclusive AI education and training, in order to ensure that AI is non-discriminatory and inclusive" (EC 2018e)

Another value from Manners' list of European norms that occurs with regularity in the data set is sustainable development (Manners 2002). The analysis has found that sustainable development co-occurs with more general norms of democracy and social solidarity.

"The use of AI systems can have a significant role in achieving the Sustainable Development Goals, and in supporting the democratic process and social rights." (EC 2020e p.2)

Together with the recent proposals on the European Green Deal, Europe is normatively well poised to deal with climate and environmental challenges. Here the EU has somewhat of a paradox in how it views AI. While on the one hand AI can be a critical enabler for attaining sustainable development goals, through e.g., smart municipality projects (EC 2018d), and on the other hand the technology has a complicated supply chain and lifecycle, with regards to resource usage for the training of algorithms and the storage of data. This increases the demand for processing power, which come at high energy costs and subsequently increases the demand for natural resources used in battery and chip production (Bird et al. 2020). The EU acknowledges, in a reflexive manner, these risks to the

environment, but seems somewhat undecided whether or not this poses a threat to EU values or if the opportunity is considered bigger than the barrier. Some reports stress the value of AI in reaching climate goals, while others stress its high resource demands. The formulated strategy with regards to this is somewhat unclear.

> *"Public policy should also encourage the wider availability of privately-held data, while ensuring full respect for legislation on the protection of personal data […]." (EC 2018e).*

Data privacy is treated in a similarly ambiguous way by the EU's AI strategy. Simultaneously calling for increased availability of privately-held data while also self-portraying as a defender of personal privacy. However, this need not reduce the EU's genuine normative commitment, in fact, the EU already has regulations like GDPR and the European ePrivacy Directive in force that ensures the right to persons personal data, and these can be seen as normative tools that well complement the EU's AI strategy.

What keeps surprising you while you analyse this large amount of data is how little the founding values of peace, liberty and rule of law are mentioned or elaborated upon. However, all three of these have considerable conceptual overlap to the umbrella-term democracy, which certainly is a central norm. The difficulty in distinguishing these three values from democracy is an inherent limitation to the coding used in this thesis. All three appear in a similarly low frequency. However, AI transparency, safety and right to privacy can be interpreted as constituting elements of liberty. A similar argument could be made for the code rule of law and good governance. While anti-discrimination and sustainable development have clear markers and are easily identifiable in a given document, these other concepts are more overarching and guiding principles that coat the entirety of the AI vision. The strong commitment to AI transparency and the demands for human-centric and ethical AI supports the statement that good governance is indeed present in the EU's overall stance.

> *"AI should adhere to basic principles: that AI should not be used to harm humans in any way;" (EC 2018d)*

Lethal autonomous weapon systems (LAWS) are not something the EU wants to promote and deploy. “These learning machines with cognitive skills that can potentially decide who, when and where to fight without human intervention” (EC 2018d) were proposed to be banned in a resolution passed by the European Parliament in 2018. The argument was that taking someone’s life has to be a choice exercised by a person with accountability and responsibility for their actions (EP 2021). However, the use of AI systems defensively, for purposes of e.g., better cyber-security seems to not be as frowned upon.

“The increasing potential and sensitivity of AI applications in many areas of the digital economy and society, such as autonomous mobility or avoiding power blackouts, means it is highly relevant to establish cybersecurity requirements for AI.” (EC 2018f p.8)

One would think that with this potential uprise in cyber-warfare and digital threats, that peace would be a more central norm, seeing as it is one of the founding norms of the EU. But it is mentioned quite rarely and only in passing when referencing AI-safety in a broader sense. A reason for this could potentially be that defence and security as a policy field is not prioritised in the AI-vision in general – or in fact a subliminal commitment to peace, the EU certainly seems to not be promoting militarized applications of AI. On the contrary, it promotes the use of safe and transparent AI systems.

New challenges in terms of safety also create new challenges in terms of liability for the EU. The fog surrounding accountability in case of accidents or unintentionally malicious AI needs to be cleared, and the challenges addressed to ensure the same level of predictability and protection when compared to victims of conventional tech - “This will help create trust in these new emerging digital technologies and create investment stability.” (EC 2020d p.17). To ensure legal certainty, albeit both for the protection of individuals and to foster investment, for the case of economic gain indicates that individual rights could be side-lined in favour of self-interest in the form of investment stability.

“The concern is not, however, that a self-conscious machine would somehow ‘take over’, but rather that it would effectively be an ‘ethical

object' and therefore subject to the same values as humans" (EC 2018d p.6)

The EU rarely mentions the ethical status of sentient machines, but when it does, the reasoning becomes abstract and very sci-fi. The quote above, taken from the Commission paper *Artificial intelligence, real benefits* (2018) seems to indicate that the EU wants to extend the ethical status of persons to that of AI systems, in a scenario where they become sentient, or at the very least it wants to avoid a situation where a digital system becomes an ethical object. While this is an interesting scenario, it is far beyond the scope of this thesis to interject a discussion on the moral philosophy of conscious machines. Perhaps it should also be beyond the scope of official EU documentation.

4.3.1 Interim findings

Democratic values imbue the EU's AI strategy in so far as we consider democracy an umbrella concept. Values like anti-discrimination, social solidarity and human rights co-occur frequently and depend on one another. However, its stance on sustainable development is more unclear, as both upsides and downsides of AI in the pursuit of climate-change mitigation are brought up. This nevertheless fits Niemann & de Wekkers criteria of reflexivity. Reflexive as it may be, the EU still portrays some double standards, e.g., in the acknowledgement of the need for data access. The acknowledgement that AI systems need to be "trained" on massive amounts of personal data is a utility-based argument in stark contrast to that of the ideals of data privacy outlined in the GDPR (Floridi 2021). The GDPR and ePrivacy Directive are in force however and can be seen as crucial complements to the AI strategy as a normative tool. The pre-defined values of peace, liberty and good governance occurred less frequently than democracy, but can still be said to be present throughout the documentation. Good governance, in particular, can be seen when the European strategy is taken as a whole. Since the repeated and fervent calls for transparent, safe and human-centred AI all indicate good governance by Manners' standards. This can be seen as a limitation of the coding and choice of theoretical framework in relation to AI as a field, perhaps even an inherent flaw with latent analysis on the whole.

AI liability was in the end coded as a *peripheral value*, meaning that norms are side-lined in favour of utility-based interests, since the connotations to legal predictability of AI in case of for example accidents are invoked to ensure investment stability. When push comes to shove there are some conflicting ideals, most notably economic growth versus preservation of European values, that it is hard to tell which is prioritized. But the fact that these norms have such a central value and that the documents show impressive coherence between one another indicates a genuine normative commitment that is indicative of a normative power. But we cannot just take the EU's word for it.

4.4 The AI Act – Early reflections

The EU is the first institution to take a major step in holistically regulating the use of AI systems through the EU AI Act, a proposal for regulation laid out by the Commission on the 21st of April 2021. This document represents an important milestone in not only European AI governance, but global AI governance, a statement supported by the fact that US president Biden upon being elected was greeted by the EU with an EU-U.S. Agenda. The agenda proposed, among other things, that the EU and US "start acting together on AI" to develop regional and global standards based on a shared belief in a human-centric approach (European Commission 2020). The fact that the EU AI Act now is the first proposed AI-specific bill in the world could provide the EU with significant agenda-setting priority and puts it in a pioneer position. Scholars have drawn comparison to the astounding global impact of the GDPR and argue that the AI Act might have similar far-reaching effects (Gaumond 2021). However, it is too early to tell what the impacts of this proposed regulation will be and it will take at least another two to three years before the regulation gets passed with all expected input from lawmakers. Even then it takes another two years before the regulation starts applying. Nevertheless, the AI regulatory proposal is the latest addition to an ambitious digital agenda promised by von der Leyen's Commission and will certainly prove valuable internationally, even as it evolves like the technologies it

seeks to master (EC 2021b). For this thesis, examining to what extent the EU intends to act in a normative way with regards to AI, it is thought reasonable to dedicate this document its own section of the analysis, given its potential weight as a normative tool.

The Proposal for a Regulation Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act), henceforth referred to as the AI Act, is guided by the underlying idea found central in the EU's AI strategy, namely that trustworthy AI is an end in itself and will promote AI uptake in Europe through harmonising rules and avoiding fragmentation of the digital single market (EC 2021b p.1-2). The building blocks for this trust is the safeguarding of people's safety and fundamental rights, while at the same time ensuring that the boundaries are not so burdensome as to hamper the innovation of AI in Europe.

The analysis finds that there is strong value coherence between the more visionary documents explored in sections 4.2 and 4.3, but the focus is shifting more towards practical considerations of how to regulate AI. It is hard balancing fundamental rights without inhibiting innovation in AI, but the AI Act attempts to find a middle ground (Gaumond 2021).

It categorizes different AI applications through a risk-based approach, banning unacceptable uses of AI, regulating those that pose a high risk and encourages adopting codes of conduct for those applications that are of limited- or no risk at all (EC 2021b p.11-14). In this regard, the AI Act follows through on the normative ambitions in the more visionary earlier documents with regards to the norms enshrined in the charter. It now explicitly states that any AI that risks jeopardizing the fundamental rights found in the charter of fundamental rights are to be banned outright and moves down the ladder of severity for the lower risk systems (EC 2021b p. 12-13).

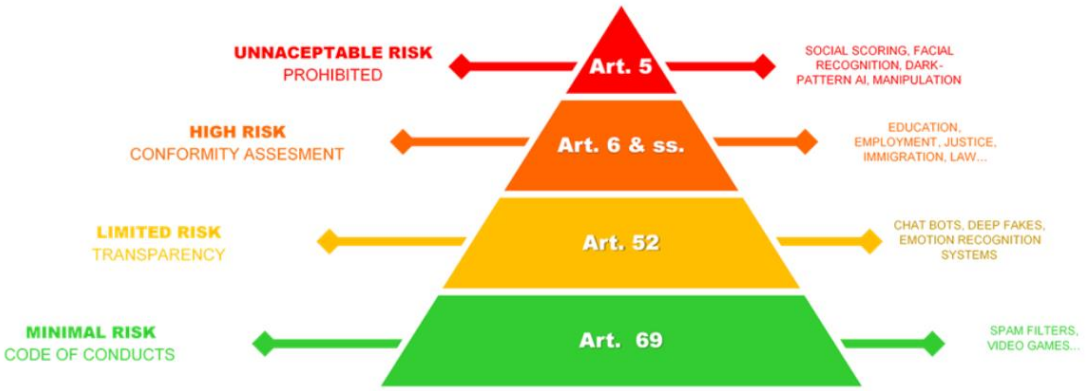

Figure 1. The risk-based approach of the AI Act – categorization of different levels of risk posed by AI systems. Accessed from Lawfareblog.com

"AI systems providing social scoring of natural persons for general purpose by public authorities or on their behalf may lead to discriminatory outcomes and the exclusion of certain groups. They may violate the right to dignity and non-discrimination and the values of equality and justice."
(EC 2021b p.21)

Those practices belonging to the category of unacceptable risk are social credit systems, biometric facial recognition, dark-pattern AI and manipulative AI. Banning social credit systems seems like a sneer directed at Chinese social score-style AI systems. By stating that public authorities are not allowed to assess people's trustworthiness through AI, the EU makes its dedication to fundamental rights abundantly clear. In a similar vein dark-pattern AI systems - which are purported to be "technologies that deploy subliminal techniques beyond a person's consciousness to materially distort the person's behaviour in a manner that is likely to cause either psychological or physical harm" – reaffirms the EU's prioritisation of AI safety as seen laid out in the previous section. Article 5 e.g., would disallow the use of inaudible audio in the cabins of truck drivers to push them to drive for longer than is safe (EC 2021b p.43). The ban on manipulative AI is specifically targeted at situations where the vulnerabilities of a specific group of people, due to their age or any disabilities, risk being exploited in a way that risks them harming themselves or others (ibid). It reiterates commitments to promoting anti-discrimination.

"The use of AI systems for 'real-time' remote biometric identification of natural persons in publicly accessible spaces for the purpose of law enforcement is considered particularly intrusive in the rights and freedoms of the concerned persons, to the extent that it may affect the private life of a large part of the population, evoke a feeling of constant surveillance and indirectly dissuade the exercise of the freedom of assembly and other fundamental rights." (EC 2021b p.21)

It also moves to ban biometric facial recognition used in mass surveillance – and this I think is a key takeaway. Democracy is once again present as a central overarching norm, it occurs much more frequently than peace or liberty, although as discussed previously there is considerable overlap between these categories and so this can be considered a limitation of the coding and not an intrinsic lack in the AI Act. However, the Commission has chosen to add a limited number of cases where facial recognition with biometrics is to be allowed. Here it distinguishes between two instances of biometric facial recognition. The first one applies to all biometric identification systems, used by either public or private actors. These have to comply with all the requirements for a "high-risk", but with even stricter criteria added on – such as a third-party conformity assessment. The second category concerns biometrics used by law enforcement agencies, i.e., the police and their real-time surveillance of public spaces. Generally speaking, these technologies can take a picture of a person's face, analyze it, and cross-compare with a database and almost instantly have their identity figured out. This way of using AI-powered biometric recognition is generally considered an "unacceptable risk" and is banned. But there is a list of specific instances where this still can be considered legal. For example, in the case of a national security threat or for finding a missing child – all in all, it is a list that feasibly allows a multitude of uses for biometric facial recognition (EC 2021b p.22). The EU is willing to sacrifice some ideals of privacy for the sake of security, which in itself is no cause to suspect that the EU is not a normative power, but indicates a pragmatism towards the potential uses of AI in line with the "maximize benefit while minimizing risk"- position identified earlier.

With regards to other values, like sustainable development or social solidarity, the analysis indicates that there has been a move away from guaranteeing positive

rights and instead doubling down on negative rights, where the freedom from harm, social scoring, surveillance and anti-discrimination are given more room than for example sustainable development or social solidarity. The concerns expressed by the EESC in its document *Artificial intelligence: anticipating its impact on work to ensure a fair transition* seem to have gotten lost in the consultation process. However, why social solidarity appeared less frequently can perhaps be explained by the fact that the EU has limited competence in the field of social policy.

> *"AI systems used in employment […] should also be classified as high-risk, since those systems may appreciably impact future career prospects and livelihoods of these persons." (EC 2021b p.26)*

It does however mandate that recruitment tools powered by AI be classified as "high-risk" since they risk perpetuating historical patterns of discrimination against women, certain age groups, disabled people, or people belonging to a certain race (EC 2021b p.36). This is the only real place where issues of employment are brought up, the AI Act itself does nothing to address the concerns of the EESC about job loss due to automation or digitalization.

> *"By improving prediction, optimising operations and resource allocation, and personalising digital solutions available for individuals and organisations, the use of artificial intelligence can provide key competitive advantages to companies and support socially and environmentally beneficial outcomes." (EC 2021b p.18).*

The EU is less reflexive now when it comes to sustainable development. The earlier synthesis of documents could point to arguments made in both directions, that AI can be both harmful and beneficial in the pursuit of climate-change mitigation. However, in the AI Act, the consensus seems to be that AI will promote economic growth and ensure progress in issues of sustainable development. No distinguishable action is taken to ensure this will happen in the AI Act itself though.

The issue of militarized AI is similarly ambiguous. While the AI Act bans the use of AI systems that risk harming people physically or psychologically, it gives a blanked exemption to AI systems developed for use in fields governed by the

CFSP. This area is instead regulated by title V of the TEU (EC 2021b p.20). This, one could argue, leads us to assume a less peacefully oriented EU or at least one that allows the use of AI-aided militaries.

Beyond these caveats, however, fundamentally the analysis has found a continuation of the ethical stance found in the previous section. The ethics guidelines proposed by the HLEG-AI are explicitly adopted. It mitigates risks associated with AI, and sees public trust in these technologies as paramount and wants to further the development of AI in Europe with humans at the centre of consideration (HLEG-AI 2019).

The AI Act is a regulation and not a directive, meaning that the proposed law will enter into force on the same date across all member states (while a directive simply requires the member states to reach a certain goal, not prescribing exactly how to transpose the aims into national law, that is left up to the member states) (Europa.eu 2021). The regulation will also enjoy direct effect, given that the rights conferred upon individuals are sufficiently clear and precise, and so any natural or legal entity will be able to invoke these rules before a national court of law (Floridi 2021 p.216). This speaks volumes about the potential normative impact that we can expect the AI Act to achieve. And even though some norms have been side-lined, the overall commitment to ensuring trustworthy and human-centric AI is still present and now embodied through a potentially huge regulatory proposal. Even though NPE is primarily concerned with the external impact of EU policies, it is still a little soon to predict what the external impact of the AI Act will be precisely, but the GDPR can give us some indication. The “Brussels-effect” is a well-known mechanism whereby the EU exercises extraterritoriality by imposing rules for the single market that affect all those that wish to do business in said market. With regards to the digital, the GDPR is one of the most consequential regulatory developments in information policy of this generation. It has brought personal data into a detailed regulatory regime that has influenced personal data usage worldwide and effectively become the international standard (Hoofnagle et al 2019).

Arguing that the AI Act could achieve a similar status would not be entirely unreasonable. Either the EU becomes the global standard-setter for AI, much like it has for data management and exercises the extraterritoriality of its single market to a great degree. The extraterritorial reach of the AI Act indicates the EU’s aims

of promoting universally acceptable norms globally: It covers actors in AI located within but also outside the Union where “the output produced by those systems is used in the Union” (EC 2021b p.21). Or this piece of legislation effectively dissuades AI uptake and investment in the EU. What we can discern is that the values promoted in the AI Act come at a potentially great cost for certain economically minded actors in the EU, but still the EU is pushing a strict sector-wide regulation that favours fundamental rights. Therefore, confirming once again the central value of these norms.

5 Concluding Discussion

This thesis has attempted to discover whether the European Union indicates that it intends to act the way a normative power would in the field of AI governance. It has done so by analysing the state-of-the-art legislative documentation surrounding the Artificial Intelligence Act and applying the criteria of Manners (2002) theory of Normative Power Europe with the operationalizations, in the form of three categories of normativity, from Niemann & de Wekker (2010). The EU seems to be driven by alternative motivations for regulating AI, an excellent case to analyse, which Manners would argue stem from its complex hybrid-polity nature. The qualitative content analysis of the 29 documents, the AI Act included, indicate that the EU can be classified as a normative power. But with certain limitations or cases where the genuine normative commitment of the EU seems more ambiguous. In this way, the thesis partially confirms the findings of Parviala (2019) and fulfil the theoretical expectations, garnered from this paper's framework, of the European Union as a normative power, but with limitations.

With regards to the first research question – the vision on AI governance can be said to be one of trustworthiness, cooperation and reflexivity. The EU acknowledges the risks of market fragmentation in promoting its norms and calls for EU-wide cooperation among the member states and through public-private partnerships to remedy this. In ways echoing the findings of Franke & Sartori (2019), who called for increased efforts at EU-wide cooperation as the main approach for ensuring a competitive edge for Europe. Furthermore, within its often-portrayed concept of "human-centred AI" the EU wishes to promote AI systems that are transparent, safe, trustworthy and that always place "humans-in-the-loop". While the overall commitment to trustworthy AI shines through, it is evident that the EU is torn between two conflicting ideals here, mainly between economic growth and ethical AI. This can be likened to Niemann & de Wekkers idea that utility-based principles should generally be sidelined in favour of value-based principles for justifying policy when acting as a genuine normative power.

The second research question asked what norms are identifiable within the EU's AI vision and subsequently what these then say about the EU's self-proclaimed normative power role. Democracy can be considered the umbrella concept that binds together other central values like peace, liberty, rule of law and good governance. Good governance, in particular, can be seen when the European strategy is taken as a whole. Since the repeated and fervent calls for transparent, safe and human-centred AI all indicate good governance by Manners' (2002) standards. Some documents and reports stress the value of AI in reaching climate goals, while others stress its high resource demands. In the pursuit of climate-change mitigation, the EU is more ambiguous in how it views AI systems. The EU acknowledges, in a reflexive manner, risks to the environment, but seems somewhat undecided whether or not this poses a threat to EU values or if the opportunity is considered bigger than the barrier. We can interpret the EU's stance towards data privacy to be similarly ambiguous and even portrays some double standards e.g., in the acknowledgement of the need for data access. The acknowledgement that AI systems need to be "trained" on massive amounts of personal data is a utility-based argument in stark contrast to that of the ideals of data privacy outlined in the GDPR (Floridi 2021). The GDPR and ePrivacy Directive are however in force and can in this regard be seen as crucial complements to the AI strategy. The analysis indicates that ensuring liability has a dual purpose, both to ensure legal predictability for the safety of persons in cases where AI systems are involved in accidents or cause unintentional harm, but also to ensure investment stability. This utility-based interest indicates that EU norms have a peripheral value in this instance. But the fact that the list of pre-defined norms holds such a central position and that the documents show impressive coherence between one another indicates a genuine normative commitment that is indicative of normative power. This largely fulfils the criteria for normative intent as described by Niemann & de Wekker.

The process behind the AI Act seems to have been inclusive and reflexive enough to warrant the judgement that the criteria of normative process are fulfilled. What we then get from the AI Act, however, is slightly less ambitious. Certainly, the Commission acknowledges the need for the research sector's involvement, sometimes so much so that one wonders if technocratic stereotypes are not entrenched, and had ambitions for inclusion of civil society actors and

social solidarity. These are seen to a lesser extent in the AI act, which instead focuses on the safety and democracy aspects of dysfunctional or maleficent AI. It does posit that the different risk groups and associated requirements on AI systems have to be adapted in the future. The EU wants to evolve with the technologies it seeks to master. But one wonders who will provide the input that reappraises this list of AI-safety requirements.

Assessing normative impact was more difficult, however. The NPE-framework was originally developed as a way to analyse the EU in global affairs and its foreign policy, while this thesis has taken on a more EU-internal view. Niemann & de Wekker (2010) look for definitive signs in Moldova that the EU acts normatively. Without this third-party comparison, the analysis falls, in retrospect, a bit short in genuinely assessing the normative impact of EU AI governance. However, this paper has argued that the very institution of the HLEG-AI, the EU AI Alliance and AI Act points us in a direction of what we can expect. It is not entirely unlikely that the AI Act will enjoy similar status as the GDPR as the global standard (Floridi 2021), although claiming this definitively based on the findings of this paper would be conjecture. The EU-internal view could very well have been complemented with a third country for the sake of comparison and to better determine the true normative impact. But this is also a question of timing. The AI Act just came out this April and will not apply to the member states until a couple of years in the future. We will simply have to wait and see; only then can we reveal normative interrelations and impacts of the EU's AI strategy.

A reflective note on the choice of theory is in order. While Manners' normative power Europe theory does not in its original form apply to digital technologies, with some adaptation it seems a suitable lens through which to view AI governance. Other researchers have similarly had success with applying it to different areas, like Niemann & de Wekker (2010) do to geopolitics, and Manners himself indeed does to the rights of children and the abolition of the death penalty (Manners 2002 & 2008). Adapting it specifically to the digital was a challenge, however, and a different researcher might not agree that AI transparency, safety or liability for example constitute European norms. However, this paper argues that these can be derived from the EU's foundational treaties and have a basis in universally acknowledges norms within the UN system. This echoes the

insinuation made by Manners, that the EU will have to adapt its value portfolio as the world changes, as indeed happened when sustainable development became a guiding principle. The theory needs adaptation to be successfully applied to different policy fields and modern matters. The all-encompassing nature of AI and its influence on international politics warrants this adaptation.

The EU and its AI act represents one of the world's - if not *the* – first attempts at harnessing the potential of AI systems, while also minimising their risks in line with a normative approach. The ambiguity in certain selected areas can perhaps be regarded as childhood flaws and the chance that other countries follow suit as they have with the GDPR, is not entirely unlikely. The AI Act protects the citizens of Europe and her passed down values. If this has a global impact and the EU gets to define what passes as normal in AI governance, then that must be considered a very powerful Europe indeed.

6 Bibliography

6.1 Primary documents

Council of the European Union 2020 *Presidency conclusions - The Charter of Fundamental Rights in the context of Artificial Intelligence and Digital Change*. Brussels: European Council of the European Union. Available at: <https://www.consilium.europa.eu/en/press/press-releases/2020/10/21/artificial-intelligence-presidency-issues-conclusions-on-ensuring-respect-for-fundamental-rights/>

EESC 2018 *Artificial Intelligence: anticipating its impact on work to ensure fair transition (own-initiative opinion)*. (537). Brussels: European Economic and Social Committee. Available at: <https://www.eesc.europa.eu/en/our-work/opinions-information-reports/opinions/artificial-intelligence-anticipating-its-impact-jobs-ensure-fair-transition-own-initiative-opinion>

European Commission 2018a. *Artificial Intelligence - A European Perspective*. Available at: <https://publications.jrc.ec.europa.eu/repository/bitstream/JRC113826/aiflagship-report-online.pdf>

European Commission 2018b. *Artificial Intelligence: Commission kicks off work on marrying cutting-edge technology and ethical standards* [Press release]. Available at: <https://ec.europa.eu/growth/content/artificial-intelligence-commission-kicks-work-marrying-cutting-edge-technology-and-ethical_en>

European Commission 2018c. *Artificial Intelligence*: *Commission outlines a European approach to boost investment and set ethical guidelines* [Press release] Available at: <https://ec.europa.eu/commission/presscorner/detail/en/IP_18_3362>

European Commission 2018d. *Artificial Intelligence, real benefits*. Content and Technology, European Commission. Available at: <https://ec.europa.eu/digital-single-market/en/news/artificial-intelligence-real-benefits>

European Commission 2018e. *Communication from the Commission to the European Parliament, the European Council, the Council, the European Economic and Social*

Committee and the Committee of the Regions - Artificial Intelligence for Europe. Brussels: European Commission. Available at: <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=COM%3A2018%3A237%3AFIN>

European Commission 2018f. *Communication from the Commission to the European Parliament, the European Council, the European Economic and Social Committee and the Committee of the Regions - Coordinated Plan on Artificial Intelligence*. Brussels: European Commission. Available at: <https://ec.europa.eu/digital-single-market/en/news/coordinated-plan-artificial-intelligence>

European Commission 2018g. A European approach on Artificial Intelligence [Press release] Available at: <https://digital-strategy.ec.europa.eu/en/policies/european-approach-artificial-intelligence>

European Commission 2018h. *Liability for emerging digital technologies*. Brussels: European Commission. Available at:<https://ec.europa.eu/transparency/regexpert/index.cfm?do=groupDetail.groupMeetingDoc&docid=36608>

European Commission 2018i. Member States and Commission to work together to boost artificial intelligence „made in Europe“ [Press release] Available at: <https://ec.europa.eu/commission/presscorner/detail/en/IP_18_6689>

European Commission 2019a. *The European AI Landscape*. Available at: <https://ec.europa.eu/jrc/communities/sites/jrccties/files/reportontheeuropeanailandscapeworkshop.pdf>

European Commission 2019b. *Trustworthy AI - Joining efforts for strategic leadership and societal prosperity* [Brochure] Available at: <https://www.ltu.se/cms_fs/1.192206!/file/TrustworthyAI.pdf>

European Commission 2020a. *European enterprise survey on the use of technologies based on artificial intelligence*. Retrieved from Leuven, Belgium: Executive summary. Available at: <https://ec.europa.eu/digital-single-market/en/news/european-enterprise-survey-use-technologies-based-artificial-intelligence>

European Commission 2020b. *Margrethe Vestager Opening message for the European AI Forum*. In: European Commission. Available at: <https://ec.europa.eu/commission/commissioners/2019-2024/vestager/announcements/opening-message-european-ai-forum_en#:~:text=We%20want%20to%20develop%20European,an%20inspiring%20and%20productive%20conference!>

European Commission 2020c. *The pivotal role of research and innovation in artificial*

intelligence policy [Fact Sheet] Available at: <https://era.gv.at/public/documents/4196/FINAL_AI_factsheet.pdf>

European Commission 2020d. *Report from the Commission to the European Parliament, the Council and the European Economic and Social Committee - Report on the safety and liability implications of Artificial Intelligence, the Internet of Things and robotics*. Available at: <https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52020DC0064&from=en>

European Commission 2020e. *White Paper: On Artificial Intelligence - A European approach to excellence and trust*. Brussels: European Commission. Available at: <https://ec.europa.eu/info/sites/info/files/commission-white-paper-artificial-intelligence-feb2020_en.pdf>

European Commission 2021a. *Digital Europe - For a more competitive autonomous and sustainable Europe* [Press release] Available at: <https://ec.europa.eu/digital-single-market/en/news/digital-europe-more-competitive-autonomous-and-sustainable-europe-brochure>

European Commission 2021b. *Regulation of the European Parliament and of the Council Laying down harmonized rules on Artificial Intelligence (Artificial Intelligence Act) and amending certain Union legislative acts*. Brussels: European Commission. Available at: <https://eur-lex.europa.eu/legal-content/EN/TXT/?qid=1623335154975&uri=CELEX%3A52021PC0206>

European Parliament 2020. *Artificial Intelligence and Law Enforcement - Impact on Fundamental Rights*. Available at: <https://www.europarl.europa.eu/RegData/etudes/STUD/2020/656295/IPOL_STU(2020)656295(SUM01)_EN.pdf>

HLEG-AI 2019a. *Ethics Guidelines for Trustworthy AI*. High-Level Expert Group on Artificial Intelligence Available at: <https://ec.europa.eu/futurium/en/ai-alliance-consultation#:~:text=The%20Ethics%20Guidelines%20for%20Trustworthy,strategy%20announced%20earlier%20that%20year.>

HLEG-AI 2019b. *Policy and Investment Recommendations for Trustworthy AI*. High-Level Expert Group on Artificial Intelligence Available at: <https://ec.europa.eu/digital-single-market/en/news/policy-and-investment-recommendations-trustworthy-artificial-intelligence>

HLEG-AI 2020. *The Assessment List for Trustworthy Artificial Intelligence (ALTAI) for self assessment*. High-Level Expert Group on Artificial Intelligence Available at: <https://ec.europa.eu/digital-single-market/en/news/assessment-list-trustworthy-artificial-intelligence-altai-self-assessment>

6.2 EU-commissioned reports

Bird, Eleanor. Jasmine Fox-Skelly, Nicola Jenner, Ruth Larbey, Emma Weitkamp & Alan Winfield 2020. *The ethics of artificial intelligence: Issues and initiatives*. This study has been drafted at the request of the Panel for the Future of Science and Technology within the Directorate General for Parliamentary Research Services by the University of West of England. Available at: <https://www.europarl.europa.eu/thinktank/en/document.html?reference=EPRS_STU(2020)634452>

Delponte, Laura 2018. *European Artificial Intelligence (AI) leadership, the path for an integrated vision*. Available at: <https://www.europarl.europa.eu/thinktank/en/document.html?reference=IPOL_STU(2018)626074>

Kritikos, Mihalis 2019. *Artificial Intelligence ante portas: Legal & ethical reflections*. EU guidelines on ethics in artificial intelligence. In E. P. R. Service (Ed.), (Vol. PE 640.163). Brussels: European Parliament. Available at: <https://www.europarl.europa.eu/RegData/etudes/BRIE/2019/634427/EPRS_BRI(2019)634427_EN.pdf>

Servoz, Michel 2019. *The Future of Work? Work of the Future!* European Commission Available at: <https://ec.europa.eu/digital-single-market/en/news/future-work-work-future>

van Wynsberghe, Aimee. 2020. *Artificial intelligence: From ethics to policy*. European Commission. Panel for the Future of Science and Technology. Available at: <https://www.europarl.europa.eu/RegData/etudes/STUD/2020/641507/EPRS_STU(2020)641507_EN.pdf >

7 Literature

Andrus Ansip 2017a. *Keynote speech by Vice-President Ansip in charge of the Digital Single Market*. Available at: <http://europa.eu/rapid/press-release_SPEECH-17-1652_en.htm>

Andrus Ansip 2017b. *Speech by Vice-President Ansip in Athens on „A Digital Strategy for Greece: path to growth*", European Commission. Available at: <https://ec.europa.eu/commission/commissioners/2014-2019/ansip/announcements/speech-vice-president-ansip-athens-digital-strategygreece-path-growth_en>

Atlas.ti 2021. *Qualitative Data Analysis: ATLAS.ti*. NYU Libraries – web access. Available at: <https://guides.nyu.edu/QDA/atlasti#:~:text=ti-,ATLAS.,building>

Bicchi. 2006. *'Our size fits all': normative power Europe and the Mediterranean*. Journal of European Public Policy, 13(2), 286-303. Available at: <https://www.tandfonline.com/doi/abs/10.1080/13501760500451733 >

Bradford, Anu 2020. *The Brussels Effect: How the European Union Rules the World*. Oxford Scholarship Online. Available at: <https://oxford.universitypressscholarship.com/view/10.1093/oso/9780190088583.001.0001/oso-9780190088583>

Bretherton, Charlotte & Vogler, John 2013. *A global actor past its peak?* Sage journals. Available at: <https://journals.sagepub.com/doi/abs/10.1177/0047117813497299>

Bryman, Alan 2012, '*Research Designs', in Social Research Methods*, 4th edition (Oxford: Oxford University Press), pp. 44-78. Available at: <https://www.academia.edu/38228560/Alan_Bryman_Social_Research_Methods_4th_Edition_Oxford_University_Press_2012_pdf >

Bull, H. 1982 '*Civilian Power Europe: A Contradiction in Terms?*'. Journal of Common Market Studies, Vol. 21, No. 2, pp. 149–64. Available at: <https://onlinelibrary.wiley.com/doi/10.1111/j.1468-5965.1982.tb00866.x>

Bush Vannevar 1945. *As We May Think*. July 1945 Issue The Atlantic Journal. Available at: <https://www.theatlantic.com/magazine/archive/1945/07/as-we-may-think/303881/>

Bussler, Frederik 2020, *"A History of Artificial Intelligence – From the Beginning".* Towards Data Science. Available at: <https://towardsdatascience.com/a-history-of-artificial-intelligence-from-the-beginning-10be5b99c5f4>

Carr, E.H. 1962 *The Twenty Years' Crisis 1919–1939: An Introduction to the Study of International Relations.* 2nd edn (London: Macmillan).

Carrico, Goncalo 2018. *The EU and artificial intelligence: A human-centred perspective*. European View, 17(1), 29-36. Available at: <https://journals.sagepub.com/doi/full/10.1177/1781685818764821>

Castro, Daniel, McLaughlin, Michael & Chivot, Eline 2019. *"Who Is Winning the AI Race: China, the EU or the United States?"*. Center for Data Innovation. Available at: <https://www2.datainnovation.org/2019-china-eu-us-ai.pdf>

Cooper, Robert 2004*. "Breaking of Nations: Order and Chaos in the Twenty-First Century".* Grove Press

Diez, Thomaz & Antje, Wiener 2018. *"Introducing the Mosaic of Integration Theory".* Berlin: Freie Universität Berlin, FB Politik- und Sozialwissenschaften, Otto-Suhr-Institut für Politikwissenschaft Kolleg- Forschergruppe "The Transformative Power of Europe". Available at: <https://nbn-resolving.org/urn:nbn:de:0168-ssoar-57752-1>

Diez, Thomaz 2005. *Constructing the Self and Changing Others: Reconsidering `Normative Power Europe'*. Millenium, 33(3), 613-636. Available at: <https://journals.sagepub.com/doi/10.1177/03058298050330031701>

Dilmenagi, Cem 2021 *Will AI reach singularity by 2060? 995 experts' opinions on AGI. AI Multiple*. Available at: <https://research.aimultiple.com/artificial-general-intelligence-singularity-timing/>

Duchêne. 1972. *'Europe's Role in World Peace'*. London: Fontana.

EDPS European Data Protection Supervisor 2021. *Artificial Intelligence Act: a welcomed initiative, but ban on remote biometric identification in public space is necessary*. Available at: <https://edps.europa.eu/press-publications/press-news/press-releases/2021/artificial-intelligence-act-welcomed-initiative_en>

Ekelund, Helena 2019. *Normative Power Frontex? Assessing agency cooperation with third countries. In The External Dimension of EU Agencies and Bodies.*: Edward Elgar Publishing. Available at: <https://www.elgaronline.com/view/edcoll/9781788973748/9781788973748.00012.xml>

European Commission 2020. *A new EU-US agenda for global change.* Available at: <https://ec.europa.eu/info/sites/default/files/joint-communication-eu-us-agenda_en.pdf>

European Parliament 2021. *Guidelines for military and non-military use of Artificial Intelligence*. Available at: <https://www.europarl.europa.eu/news/en/press-room/20210114IPR95627/guidelines-for-military-and-non-military-use-of-artificial-intelligence>

Forsberg, T. 2011. *Normative Power Europe, Once Again: A Conceptual Analysis of an Ideal Type**. JCMS: Journal of common market studies, 49, 1183-1204. Available at: <https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1468-5965.2011.02194.x>

Franke, Ulrike & Sartori, Paola 2019. *"Machine politics: Europe and the AI revolution"*. *European Council on Foreign Relations*. Available at: <https://ecfr.eu/publication/machine_politics_europe_and_the_ai_revolution/>

Franke, Ulrike 2020. *France and Germany: Where do they agree on AI?*. *European Council on Foreign Relations*. Carla Hobbs (ed.) Essay collection. Available at: <https://ecfr.eu/publication/europe_digital_sovereignty_rulemaker_superpower_age_us_china_rivalry/#Artificial_intelligence:_Towards_a_pan-European_strategy>

Gabriel. 2018. *Keynote speech by Commissioner Gabriel on the ethical and economic impact of artificial intelligence at GSMA Mobile 360 Serries* - Europe event. In (Vol. SPEECH/18/5827). Brussels: European Commission. Available at: https://ec.europa.eu/commission/presscorner/detail/en/SPEECH_18_5827>

Gaumond, Eve 2021 *Artificial Intelligence Act: What Is the European Approach for AI?* Lawfare Journal. Available at: <https://www.lawfareblog.com/artificial-intelligence-act-what-european-approach-ai>

Groothuis, Marga & Niemann, Arne 2012. *Normative Power Europe? The power of the EU in its Relation to the USA in the Policy Field of Counter-terrorism*. Mainz Papers on International and European Politics, 3. Available at: <https://international.politics.uni-mainz.de/files/2012/10/mpiep03.pdf>

Haas, Ernst B. 1958: *The Uniting of Europe: Political, Social, and Economic Forces 1950–57*, Stanford, CA: Stanford University Press.

Habermas, J. 1990, *Moral Consciousness and Communicative Action*, Cambridge: Polity Press.

Haukkala. 2008. *The Russian challenge to EU normative power: The case of European European Neighbourhood Policy*. Italian Journal of International Affairs. Available at: <https://www.tandfonline.com/doi/abs/10.1080/03932720802057117>

Hogarth, Ian 2018. *"AI Nationalism".* Ian Hogarth blog. Available at: <https://www.ianhogarth.com/blog/2018/6/13/ai-nationalism>

Hoofnagle, Chris Jay; van der Sloot, Bart & Borgesius, Frederik Zuiderveen 2019. *The European Union general data protection regulation: what it is and what it means,* Information & Communications Technology Law, 28:1, 65-98. Available at: https://www.tandfonline.com/doi/pdf/10.1080/13600834.2019.1573501?needAccess=true>

Hyde-Price, Adrian 2006. '*Normative'power Europe: a realist critique*. *Journal of European Public Policy*, *13*(2), 217-234. Available at: <https://www.tandfonline.com/doi/abs/10.1080/13501760500451634>

Inferkit AI-based text generator 2021. Online text-generation software. Available at: <https://inferkit.com/>

ITU 2017 (International Telecommunications Union). *AI for Good Global Summit 2017 Report*. Available at: <https://www.itu.int/en/ITU-T/AI/Documents/Report/AI_for_Good_Global_Summit_Report_2017.pdf>

Jilke, Sebastian. Favero, N. Wolfson, J. & M. Young 2021. *Messenger Effects in Covid-19 Communication: Does the Level of Government Matter?* Health Policy OPEN. Available at: <https://sebjilke.github.io/publications/>

Jupille, Joseph & Caporaso, James 1998. *States, agency, and rules: the European Union in global environmental politics*. The European union in the world community vol 17 p.157-182. Lynne Rienner.

Kagan, Robert 2003. *Of Paradise and Power: America and Europe in the New World Order*. New York, Alfred A. Knopf.

Kuckartz U. 2019 *Qualitative Text Analysis: A Systematic Approach*. In: Kaiser G., Presmeg N. (eds) Compendium for Early Career Researchers in Mathematics Education. ICME-13 Monographs. Springer, Cham. Available at: <https://link.springer.com/chapter/10.1007/978-3-030-15636-7_8>

Lee, K. F. 2018. *AI-Superpowers, China, Silicon Valley and the New World Order*. Boston: Houghton Mifflin Harcourt.

Manners. 2002. *Normative power Europe: a contradiction in terms?* JCMS: Journal of common market studies, 40(2), 235-258. Available at: <https://onlinelibrary.wiley.com/doi/abs/10.1111/1468-5965.00353>

Manners. 2006. *Normative power Europe reconsidered: beyond the crossroads*. Journal of European Public Policy, 13(2), 182-199. Available at: <https://www.tandfonline.com/doi/abs/10.1080/13501760500451600>

Manners. 2008. *The EU's international promotion of the rights of the child*. In O. Tortell (Ed.), The European Union and the Social Dimension of Globalization: How the EU influences the World (pp. 229–241). New York: Routledge. Available at: <https://forskning.ruc.dk/en/publications/the-eus-international-promotion-of-the-rights-of-the-child>

Manners. 2009a. *The Concept of Normative Power in World Politics*. Institut for Internationale Studier / Dansk Center for Internationale Studier og Menneskerettigheder., 1-5. Available at: <https://pure.diis.dk/ws/files/68745/B09_maj_Concept_Normative_Power_World_Politics.pdf >

Manners. 2009b. *The EU's Normative Power in Changing World Politics*. In Gerrits (Ed.), Normative Power Europe in a Changing World: A Discussion (pp. 9-24). The Hague: Netherlands Institute for International Relations. Available at: <https://forskning.ruc.dk/en/publications/the-eus-normative-power-in-changing-world-politics>

Marcut, Mirela 2017. *Digital Policies as instruments in the European integration process.* Analele Universităţii din Oradea, Seria Relaţii Internaţionale şi Studii Europene, TOM IX, pag. 169-176. Available at: <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3648913>

Mattli, W. 1999. *The Logic of Regional Integration*: Europe and Beyond. Cambridge University Press.

Mayring, Philpp 2014. *Qualitative content analysis – theoretical foundation, basic procedures and software solution.* Available at: <https://www.researchgate.net/publication/266859800_Qualitative_content_analysis_-_theoretical_foundation_basic_procedures_and_software_solution>

McCarthy, J. J., Minsky, M. L., & Rochester, N.,. 1959. *Artificial intelligence. Research Laboratory of Electronics (RLE) at the Massachusetts Institute of Technology (MIT).*

Nicolaidis & Howse. 2002. '*This is my EUtopia...': narrative as power*. JCMS: Journal of common market studies, 40(4), 767-792.

Niemann & De Wekker. 2010. *Normative power Europe? EU relations with Moldova*. European Integration Online Papers, 14(1). Available at: <https://www.researchgate.net/publication/227430803_Normative_Power_Europe_EU_Relations_with_Moldova>

Parviala, Tuulia 2019. *EU Entering the Era of AI – Ai Qualitative Text Analysis of the European Union's Policy on Artificial Intelligence.* Uppsala Universitet <http://uu.diva-portal.org/smash/get/diva2:1280201/FULLTEXT01.pdf>

Persson, A. 2017. *Shaping Discourse and Setting Examples: Normative Power Europe can Work in the Israeli–Palestinian Conflict*. JCMS: Journal of common market studies, 55, 1415– 1431. Available at: <https://onlinelibrary.wiley.com/doi/full/10.1111/jcms.12578>

Pichot, J. 2019. Tech Tighrope - *Europe's Alternative to Silicon Valley*. Are We Europe, 5, 4-7. <https://www.archive.areweeurope.com/stories/code-of-conscience/tech-innovation-europe>

Recuerda, Miguel A. 2006. "*Risk and Reason in the European Union Law".* European Food and Feed Law Review. Available at: <https://www.jstor.org/stable/90008300>

Roff, H. M. 2019. *Artificial intelligence: Power to the people*. Ethics & International Affairs, 33(2), 127-140. Available at: <https://ethicsandinternationalaffairs.org/2019/artificial-intelligence-power-to-the-people/>

Rüfenacht, Mattia 2020. *The 4 waves of AI: Does the winner take it all?* Parashift archives. Available at: <https://parashift.io/en/the-4-waves-of-ai-does-the-winner-take-it-all/>

Shastri, S., Wasserman, M., & Chidambaram, V. 2019. *The Seven Sins of Personal-Data Processing Systems under {GDPR}.* Workshop on Hot Topics in Cloud Computing (HotCloud 19). <https://arxiv.org/abs/1903.09305#:~:text=We%20illustrate%20these%20conflicts%20via,security%20as%20a%20secondary%20goal.>

Sjursen, Helene 2006. *The EU as a 'normative'power: how can this be?* Journal of European Public Policy, 13(2), 235-251. Available at: <https://www.tandfonline.com/doi/abs/10.1080/13501760500451667>

Smith, Leigh 2018. *"EU Member States sign AI declaration of cooperation Ahead of publication of upcoming AI communication"*. Available at: <https://talkingtech.cliffordchance.com/en/emerging-technologies/artificial-intelligence/eu-member-states-sign-ai-declaration-of-cooperation-.html>

Staeger. 2016. *Africa–EU Relations and Normative Power Europe: A Decolonial Pan African Critique*. Journal of Common Market Studies, 54(4), 981-998. Available at: <https://onlinelibrary.wiley.com/doi/abs/10.1111/jcms.12350>

Stix, Charlotte 2019. "*A survey of the European Union's artificial intelligence ecosystem".* Leverhulme Centre for the Future of Intelligence. Available at:

<https://ec.europa.eu/jrc/communities/en/community/digitranscope/document/survey-european-union%E2%80%99s-artificial-intelligence-ecosystem>

Treaty on the European Union 2012. Consolidated version TEU. Available at: <https://eur-lex.europa.eu/resource.html?uri=cellar:2bf140bf-a3f8-4ab2-b506-fd71826e6da6.0023.02/DOC_1&format=PDF>

Vachudova. 2005. *Europe undivided: democracy, leverage, and integration after communism.* New York.

Valladão, Alfredo 2018. *Artificial Intelligence and Political Science*. OCP Policy Center. Policy Paper. Available at: <https://www.policycenter.ma/sites/default/files/OCPPC-PP1807_0.pdf>

Vesnic-Alujevic, Lucia. Susana Nascimento & Alexandre Pólvora 2020. *Societal and ethical impacts of artificial intelligence: Critical notes on European policy frameworks.* Telecommunications Policy vol.44 is.6, Juli 2020 101961. Available at: <https://www.sciencedirect.com/science/article/abs/pii/S0308596120300537?via%3Dihub>

Vincent, James 2017. *"Putin says the nation that leads in AI 'will be the ruler of the world'". The Verge.* Available at: <https://www.theverge.com/2017/9/4/16251226/russia-ai-putin-rule-the-world>

Vor dem Berge, Moritz Frederik 2021. *The EU as a Normative Power in the field of artificial intelligence? Challenges and concepts in the governance and regulation of digital technologies using the example of the EU and its human-centred approach to AI.* Westfälische Wilhelms-Universität Münster.

von der Leyen, Ursula. 2020. *Press remarks by President von der Leyen on the Commission's new strategy: Shaping Europe's Digital Future*. In E. Commission (Ed.), (Vol.SPEECH/20/294). Brussels. <https://ec.europa.eu/commission/presscorner/detail/en/speech_20_294>