# Trade Secrets, Big data and Artificial Intelligence Innovation: a Legal Oxymoron?

Ana Nordberg

Associate Senior Lecturer, Faculty of Law, Lund University, Sweden.

Key words: Big data; AI; Machine learning; Trade secrets; Digital Single Market

## 1. Information is (market) power

'Information is power'[1]. Established economic research tells us that asymmetries in information are a fundamental transaction cost and in this sense informational quality and quantity translates into economic competitive advantage. Different types of informational resources have for long been an important asset of businesses. Data, especially big data can be used in a large variety of decision-making processes concerning research, development and marketing of all types of products and services in all sector of innovation. In an ideal future, it can be used for personalised offerings and for personal decision making (from deciding on what movie to watch, to major financial decisions, wellness and health care choices). It also can be used for policy making and regulatory purposes.

'Big data is the oil of internet'.[2] Not only information is valuable, data has a market of its own. An increasingly valuable one. The Data Market is the marketplace where digital data is exchanged as "products" or "services" as a result of the elaboration of raw data.[3] The calculated value of the European data economy was €300 billion in 2016. This value could grow to up to €739 billion (4% of the EU's GDP) by 2020.[4] If information translates into market power, whoever best masters the development and use of advanced information technologies will have an unprecedented competitive advantage against those market agents that do not access to such technology.[5]

Artificial intelligence and big data analytics are capable of gathering and processing large amount of information and transform these into innovation. Different types of informational resources have for long been an important asset of businesses. Informational technologies, automated data retrieval and cross reference will produce large quantities of valuable data that can be used for research, development and marketing of all types of products and services. Big data, has explained bellow, is not merely static data sets, among other things it is characterized by being real time heterogenic data, in constant update and also able to continuously originate new data. Such data will constitute an important immaterial asset that companies will want to monetarize. Intellectual property (IP) protection, in this regard will serve as an essential tool for monetarization of investments in obtaining,

---

[1] Common aphorism whose origins can be traced to ancient texts in different cultures.

[2] "Personal data is the new oil of the internet and the new currency of the digital world." in Meglena Kuneva, European Consumer Commissioner, 2009, Keynote Speech, Roundtable on Online Data Collection, Targeting and Profiling, Brussels, 31 March 2009. Transcription available at http://europa.eu/rapid/press-release_SPEECH-09-156_en.htm.

[3] European Data Market SMART 2013/0063 Final Report (01.02.2017), p. 25.

[4] European Data Market SMART 2013/0063 Final Report (01.02.2017) study prepared for the European Commission (Directorate-General for Communications Networks, Content and Technology) by IDC and Open Evidence, pp 131 seq. Available: http://datalandscape.eu/ (viewed March, 2019).

[5] Concerning the use of AI and its economic implications on markets see: Marwala, Tshilidzi; Hurwitz, Evan (2017). *Artificial Intelligence and Economic Theory: Skynet in the Market*. (London: Springer, 2017).

maintaining and making such data useful. Trade secrets (TS) will also form an important alternative or complementary form of legal protection.

However, and as a reverse side of the coin, artificial intelligence and data mining tools pose a huge challenge to TS as legal concept. Their functioning and objective is in essence to discover or establish previously undetected or not sufficiently known correlations between large and heterogeneous sets of data. Such correlations produce new data, allowing to discover or confirm patterns, for example concerning consumer behaviour, price of commodities, labour market, on-going scientific research and technological innovation, and regulatory initiatives. The use of predictive models also allows for extrapolation of conclusions as to the future and are a useful planning tool for companies and regulators. Conventional wisdom would say that despite legal protection, in the future few secrets are expected to survive, but is it really so or can the legal framework for trade secrets (TS) offer legal shelter against technology?

This paper will analyse the Trade Secrets Directive[6] (TS dir.) from a technological informed legal perspective, looking at the possibilities and scope of protection that it offers for knowledge-based activities and business models. It will start by looking into what is big data and its relevance (section 2); proceeds with an analysis of the relevance of TS (section 3); TS object and nature of the afforded protection (section 4); requirements for protection of TS (section 5); Scope of TS protection in light of a predicted future use of AI and big data analytics as a business tool (section 6); and finalizing with concluding remarks (section 7).

## 2. What on Earth is AI and Big data and why is it a 'Big deal'?

> Data, *'noun* [U, + sing/pl verb] information, specially facts or numbers, collected to be examined and considered and used to help decision-making, or information in an electronic form that can be stored and used by a computer'[7]

This is the literal sense, but what does *data* means in an EU law sense? Article 4 (1) of the General Data Protection Regulation (GDPR)[8] states that "Personal data" means any information relating to an identified or identifiable natural person'. Thus, generalizing and by analogy, data can be *any information* relating to any given topic or subject.The proposal for and Open Data and Public Sector Information Directive[9], a recast of the Public Sector Information Directive (PSI Dir),[10] relies on a

---

[6] Directive (EU) 2016/943 of the European Parliament and of the Council of 8 June 2016 on the protection of undisclosed know-how and business information (trade secrets) against their unlawful acquisition, use and disclosure (Text with EEA relevance), *OJ L 157, 15.6.2016, p. 1–18, ELI: http://data.europa.eu/eli/dir/2016/943/oj.*

[7] Cambridge Advanced Learner's Dictionary & Thesaurus (Cambridge University Press). Available: https://dictionary.cambridge.org/dictionary/english/data.

[8] Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation) (Text with EEA relevance), OJ L 119, 4.5.2016, p. 1–88

[9] Proposal for a Directive of the European Parliament and of the Council on the re-use of public sector information (recast) COM/2018/234 final - 2018/0111 (COD). Document 52018PC0234. Available: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=COM:2018:0234:FIN (viewed April 2019).

[10] Directive 2003/98/EC of the European Parliament and of the Council of 17 November 2003 on the re-use of public sector information, OJ L 345, 31.12.2003, p. 90–96, already amended by Directive 2013/37/EU of the European Parliament and of the Council of 26 June 2013 amending Directive 2003/98/EC on the re-use of public sector information (Text with EEA relevance), OJ L 175, 27.6.2013, p. 1–8. ,

generic definition of the term 'document'. The concept of document 'covers any representation of acts, facts or information — and any compilation of such acts, facts or information — whatever its medium (written on paper, or stored in electronic form or as a sound, visual or audiovisual recording)' and excludes computer programmes.[11] The proposal also uses the terms 'dynamic data',[12] 'research data'[13] and 'high value datasets'[14] all of which are linked to the notion of documents.

In scientific usage, it is difficult to find a uniform definition or understanding of *data*, as the term is used differently in different disciplines. In similarity, in a business setting the identification and delimitation of the concept of *data* is also not clear-cut. *Data* can be conceptualized and categorized based on the substantive characteristics of its content (e.g. personal, non-personal, health, consumer, traffic, etc.) or based on legal categories (e.g. general personal data, special personal data, non-personal data, public sector data, open data, proprietary data, etc.). The scope of the concept of *data* can also vary, and for example it may 'refer to individual pieces of data (e.g. single fields in a relational database), the structured files in which they are combined, the metadata describing the data or the files, the information contained in the data, the software processing it, the algorithms on which that software is based, and any resulting knowledge derived from the data.'[15]

The notion of big data is usually associated and characterized by the presence of the '4Vs' – Volume, velocity, variety and veracity.[16] Volume, relates to an exploding volume of data produced by different sources: internet of things, social media, apps, sensors, internal networks (e.g. systems for invoicing, tracking goods in transit and costumer deliveries); repositories of information (e.g. data bases, libraries, scientific repositories and biobanks); public sector information, etc. Velocity, relates to the dynamic nature of big data. Information is processed in real time and accessed while new data is constantly produced. Variety, relates to the fact that big data is data from multiple sources, in various kinds and formats (e.g. pictures, texts, audio, video, nonverbal communication such as emoji's, memes, hashtags, tags, likes, swipes, geolocation, time, etc). Veracity, corresponds to a need for meaning (semantic information) to be accurate (at least in light of the state of the art). Not all big data will have this last element present. Veracity is an aspect that computer scientists are devoting considerable attention, especially in applications in assisted or automated decision-making (e.g. in the judicial or health sector), but also in the automation and automotive industry (e.g. robotics and driverless vehicles). Asserting the veracity of source and training data requires strict control over its origin and exclude many data sources. Not only it is difficult to assert the veracity of source data, as

[11] Proposal Open Data & Public Sector Information Directive, Recital 26.

[12] Article 2 (6) Proposal Open Data & Public Sector Information Directive 'dynamic data' means documents in an electronic form, subject to frequent or real-time updates.' – check if remains unchanged

[13] Article 2 (7) Proposal Open Data & Public Sector Information Directive 'research data' means documents in a digital form, other than scientific publications, which are collected or produced in the course of scientific research activities and are used as evidence in the research process, or are commonly accepted in the research community as necessary to validate research findings and results'

[14] Article 2 (8) Proposal Open Data & Public Sector Information Directive 'high value datasets' means documents the re-use of which is associated with important socio-economic benefits, notably because of their suitability for the creation of value-added services and applications, and the number of potential beneficiaries of the value-added services and applications based on these datasets.'

[15] Study on emerging issues of data ownership, interoperability, (re-)usability and access to data, and liability (EU Commission, April 2018), p. 75 Available: https://ec.europa.eu/digital-single-market/en/news/study-emerging-issues-data-ownership-interoperability-re-usability-and-access-data-and

[16] Some sources include only volume, velocity and variety: For the purposes of this discussion veracity should be considered.

veracity plays a role in data mining and machine learning data inferences. This type of data product can be extremely helpful in objectifying a variety of decision-making processes, but its veracity is extremely difficult or lengthy to confirm by human hands. This factor is important for regulatory issues, but also plays an important role when we consider the implications of such technology in the legal framework for IP and other immaterial rights (see below).

In summary, big data corresponds to the aggregation of large datasets, processed by computerized means. For the purposes of this analysis on trade secrets, it was adopted a broad concept of data that includes more than just the tangible data sets but also underlying information and derived knowledge, and the software and algorithms used to process data. The expression big data analytics and AI, as used in this paper, refers generally to the process of examining large and varied data sets, at a unprecedented velocity (big data), in order to uncover relevant know-how and business information -- such as for example hidden patterns, unknown correlations, market trends and customer preferences.

## 3. Trade Secrets for Protecting Big Data Analytics: Highway or Scenery Route?

Sometime ago at a workshop while discussing what were the most relevant legal issues concerning the commercial use of AI and big data, someone stated: 'Whoever has better data, will come up with a better product. We try to protect our data in anyway we can'. The corporate researchers and industry representatives at the table nodded their heads and agreed.[17]

The anecdotic statement above seems to be at odds with openness policies prioritized by the EU Commission, that announced for example that 'the aim of the Juncker Commission is to create a digital single market where the free movement of goods, persons, services, capital and data is guaranteed.'[18] In fact, over the last years it is visible a number of EU legislative initiatives to regulate the digital space and in particular data. On 25 April 2018, the EU Commission adopted the 2018 Data Package,[19] intended to address for the first-time different types of non-personal data (public, private, scientific) making use of different policy instruments within a coherent framework. This includes the review of the Public sector Information Directive (PSI)[20] and improving the framework for re-use of data generated by public sector bodies for commercial and non-commercial.

In practice, however, data markets are complex. Access and sharing of source data seem to be important for many sectors of activity, as it avoids duplication of research efforts. Simultaneously, exploitation of information asymmetries is a traditional ingredient to successful commercial strategies. Depending on the business model, companies perceive the ability to extract value from data as linked to the ability to maintain exclusivity over such data. Companies will seek to establish

[17] ehealth@LU 6th workshop, 24 August 2018, Lund University, Sweden.
[18] European Commission, Priorities, Digital Single Market. Available: https://ec.europa.eu/commission/priorities/digital-single-market_en#background (viewed april 2019).
[19] European Commission - Press release 'Data in the EU: Commission steps up efforts to increase availability and boost healthcare data sharing', 25 April 2018. Available: http://europa.eu/rapid/press-release_IP-18-3364_en.htm (viewed April 2019).
[20] Directive 2013/37/EU of the European Parliament and of the Council of 26 June 2013 amending Directive 2003/98/EC on the re-use of public sector information (Text with EEA relevance), *OJ L 175, 27.6.2013, p. 1–8. ELI: http://data.europa.eu/eli/dir/2013/37/oj.*

legal entitlements and commercial exclusivity over both AI and big data technology, and the output of big data analytics and creations by Artificial intelligence entities.

There are several type of immaterial creations linked to AI and big data for witch intellectual property rights (IPRs) and TS will be sought: a) AI and big data technology (algorithms and computer programs); b) source data sets (created or selected); c) methods to create and select data sets; d) output information (including immaterial creations of technical, scientific or artistic nature).

TS offers a legal entitlement that complements IPRs or replaces them when these are unavailable or inadequate. Furthermore, and paradoxically, because TS is not a *whole or nothing* form of protection, it is also in some cases perceived as a useful tool to enable the introduction of open innovation policies.[21] Unlike IPRs such as for example patents where the invention has to be completely described and disclosed,[22] or copyright that generally only protects an expression and not ideas, TS protection can be selective, for example a computer implemented invention,[23] a database[24] or computer program,[25] can be partially disclosed allowing IPR over the knowledge and the expressions released in the public domain, while certain specific features are kept undisclosed and protected against misappropriation.

Besides AI and big data technology, the output or product of the usage of these new ICT technologies are also often particularly valuable for companies. Big data analytics and generally AI, can be used to produce derivative data. This data can be statistical inferences about a multitude of subjects; a given arrangement of a list of information; technical information related to a product or process; aesthetical appearance of a product; sound or visual artistic expression, a literary text, etc.

In many cases big data analytics and AI output will comprise subject matter that typically can be the object of one or more IPRs. Because IPRs have at their core a logic of either a personal link with a person creator of intellectual goods or an economic incentive to innovation, AI produced art and innovation raises legal challenges to the concepts of authorship and inventorship. Requirements for protection and their assessment is often dependent on standards of notional persons (example the person skilled in the art in patent law or the informed user in design law), which have been constructed as a notional human person. In copyright the originality requirement is also linked to the work being an artistic expression of a human person, even if mediated by the use of tools and various mediums, and even if in some jurisdictions the author can be a legal person.[26] Furthermore, term of protection is usually somehow linked to the life of the author, except in jurisdictions that allow legal persons to be authors.[27]

---

[21] Recital 3, TS Dir.

[22] Article 83, Convention on the Grant of European Patents (European Patent Convention) of 5 October 1973 as revised by the Act revising Article 63 EPC of 17 December 1991 and the Act revising the EPC of 29 November 2000 (EPC).

[23] Article 52, EPC. See guidelines for examination at the EPO F-IV 3.9 'Claims directed to computer-implemented inventions'.

[24] Article 1 and 3, Directive 96/9/EC of the European Parliament and of the Council of 11 March 1996 on the legal protection of databases, *OJ L 77, 27.3.1996, p. 20–28,* ELI: http://data.europa.eu/eli/dir/1996/9/oj.

[25] Article 1 and 4, Directive 2009/24/EC of the European Parliament and of the Council of 23 April 2009 on the legal protection of computer programs (Codified version) (Text with EEA relevance) OJ L 111, 5.5.2009, p. 16–22, ELI: http://data.europa.eu/eli/dir/2009/24/oj.

[26] See for example Article 2, computer programs Dir.

[27] Article 1 (3) and (4), Directive 2006/116/EC of the European Parliament and of the Council of 12 December 2006 on the term of protection of copyright and certain related rights (codified version), *OJ L 372, 27.12.2006, p. 12–18,* ELI: http://data.europa.eu/eli/dir/2006/116/oj.

While, these type of challenges and legal difficulties raise questions that undoubtedly will occupy researchers and upper courts in the coming years, contributing to the importance of TS as a means of establishing legal entitlements where protection by IPR is uncertain. In short, TS will certainly be an important part and complement to IPR's strategy and IPR's portfolios concerning emerging ICT technologies. In some cases, where IPR are not possible or undesirable TS may even the major venue for protecting valuable commercial information.

## 4. Object and nature of protection

Trade secret is defined as meaning information which meets the following three requirements:

> (*a) it is secret in the sense that it is not, as a body or in the precise configuration and assembly of its components, generally known among or readily accessible to persons within the circles that normally deal with the kind of information in question;*
> *(b) has commercial value because it is secret;*
> *(c) has been subject to reasonable steps under the circumstances, by the person lawfully in control of the information, to keep it secret.*[28]

Thus, accordingly in principle any type of undisclosed know-how and business information can be object of a TS. There are however grey areas and it can be questioned whether the secret information has to be belong to the categories of technical know-how or business information, or may include also other commercial information. In other words, whether it is or not necessary that the information is linked to the specific business of the secret holder in a specific economic sector of activity.

The recitals clarify that trade secrets can relate to a 'diverse range of information that extends beyond technological knowledge to commercial data such as information on customers and suppliers, business plans, and market research and strategies.'[29] It is also further emphasised the importance of establishing a homogenous EU definition of TS without restricting the subject matter protected against misappropriation. It includes in the concept both business information, technological information and know-how and excludes trivial information, experience and skill obtained by employee in the normal course of their employment, and generally known information.[30] Considering the directive in its whole it does appear that the legislator meant for the concept of TS to be interpreted broadly as long as it does not conflict with public interest (e.g. administration of justice, regulatory oversight, freedom of movement and competition) and fundamental rights (privacy, data protection, freedom of expression and information, right to work, good administration, right to justice – fair trial, effective remedies and right of defence).[31]

Another related matter is whether the information has to be positive, correct and legal. In other words, does protection against misappropriation subsists if the secret is negative information, incorrect, vague or undetermined information, information illegally obtained, or information consisting or related to an illegal conduct?

Research and development activities often result in a large body of valuable negative information. For example, a company may already know that a given material cannot be mass produced in a cost-

[28] Article 2, Trade Secret Directive.
[29] Recital 2, TS Dir.
[30] Recital 14, TS Dir.
[31] Article 1 (2) and (3) and Recitals 11, 12, 13, 34, 35 and 38, TS Dir.

efficient manner, or that a given technical solution cannot be implemented due to safety concerns. unsuccessful research projects, failed experiments and abandoned prototype are often the most valuable secrets, because knowing what doesn't work provides an enormous competitive advantage. The directive does not restrict the type of information protected to positive information, nor would such make sense in light of its objectives. The directive is set up to protect undisclosed know-how and business information that has commercial value and such includes information that grants any type of commercial advantage, even if it is of a negative nature.

The same can be said for information that is incorrect, or incomplete. In certain situations, it is possible that secret business information turns out to be incorrect, false, incomplete or outdated (e.g. client contacts information, preferences or suppliers' prices). Such does not per se exclude the undisclosed information from the object of protection. However, this type of information may not meet the requirements for protection, because in specific cases the commercial value derived from the secrecy of the information may be conditional or linked to its actual commercial usefulness in providing a competitive advantage.[32]

Finally, a very different issue is whether secret information concerning an illegal conduct or for ex. contradicts the official released information about a company, can be object of a TS and protected as such. Article 5 (b) TS Dir. creates exceptions from enforcement when the alleged acquisition, use of disclosure of a TS was carried out for the purpose of protecting the general public interest and reveals 'misconduct, wrongdoing or illegal activity'[33]. According to recital 20 TS Dir., these whistleblowing activities are restricted to those that reveal directly relevant misconduct, wrongdoing or illegal activity. It thus covers information pertaining to a given conduct that is either prohibited by law or breaches a general legal duty (e.g. general duty of good faith) or contractual obligation.[34] Whether legal conduct that may be considered unethical is also covered by the exception for whistleblowing activities under the directive results unclear, and national implementation solutions variation may be expected.

More generally, it is also possible to question whether all data can be considered information for the purposes of defining the object of protection. Namely, whether valuable AI training data sets and large amounts of raw data can be the object of TS protection. Although the directive uses the expression 'undisclosed know-how and business information', the recitals indicate that the wording information is sometimes used as a synonym for knowledge – e.g. recital 1 uses the expression 'knowledge that is valuable to the entity and not widely known'. Another reason not to necessarily exclude all types of data is that data scientists do not agree on a sufficiently clear distinction of what might constitute raw data and processed data. Imposing such boundaries would also create issues in determining whether real-time data can be considered secret information and thus be object of protection.

Data can be any character, text, word, number, and, if not put into context, means little or nothing to a human. Information is data formatted or contextualized in a manner that allows it to be used by human beings in a significant way. Distinction based on how humans perceive information between semantic (meaning), syntactic (signs) and structural (physical medium) information has in this sense

[32] See Recital 1, TS Dir.
[33] Article 5 (b), TS Dir.
[34] See also: Article 1, Proposal for a Directive of the European Parliament and of the Council on the protection of persons reporting on breaches of Union law COM/2018/218 final - 2018/0106 (COD). Once this directive is in force it will create a protection for whistleblowing activities concerning certain but not all types of illegal activity.

legal implications.[35] However, in business life of companies these different layers of information are cumulative: structural data contains a collection of signs and these form (or not) the meaning. AI can extract meaning of large data sets in real-time (or near real-time) in ways that a human mind cannot, and as such the difference between data and information is bridged and difficult to establish.

Raw data or no-order data, unclassified, unorganised and without selection, and/or which is deprived of any meaning would not fulfil the directive requirements as object of protection for a TS. This data is deprived of meaning and cannot be seen as information, and difficultly will have commercial value.

However, processed raw data, data that has been collected, categorised and selected, already contains a minimum level of meaning and context and thus informational character. For example, training data for visual recognition technology selected and categorized to show different representations of an object e.g. - 'cars' or 'persons' will be valuable information; as opposed to a random collection of images.[36] This type of big data sets are the result of well-thought, well-developed processes, that builds up large data sets with proper classification, labelling and data quality control in place. The data selection process, in itself, can be highly valuable undisclosed information.

As seen above, the actual scope of the object for protection resides in the nature of the TS protection, where the object of protection is more focused on the secrecy than on the actual informational content.

The protection against misappropriation harmonised by the directive, does not create an intellectual property right over information, it creates a system of liability for a specific tortious conduct – because it requires unlawful conduct in acquiring the information. Both recital 16 and articles 3 and 4 TS Dir. point to a strong concern in ensuring that the scope of protection will not extend to the information in itself but rather will focus on its secrecy. It is the secret character that is protected against 'unfair methods of disclosure', the information in itself remains in the public domain. Likewise, the ratio legis of the specific rights and corresponding obligations are the protection of the commercial value of the information and the efforts and investment made to maintain secrecy. In this sense a broad interpretation of the object of protection would include data and big data, even when it may not be strictly considered as information but from which information can be retrieved.[37]

## 5. Requirements for protection

The TS dir. contains three requirements for protection: the Information needs to be secret; existence of a causal link between secrecy and commercial value; a duty to perform reasonable steps to kept secrecy. These further shape the object of protection, as not all information is necessarily TS protected. The next sections will take a closer look at the different elements in article 2 (1) TS Dir. which establishes these requirements for protection.

[35] Herbert Zech (2015) 'Information as Property', 6 JIPITEC 192, para 10-13, mentioning the framework proposed by Benkler and Lessig. See: Benkler, 52 Federal Communications Law Journal, 2000, 561,562; Lessig, The Future of Ideas, The Fate of the Commons in a Connected World, 2002, 23.

[36] Aditya Khosla et al. 'Undoing the Damage of Dataset Bias', in European Conference on Computer Vision (ECCV) 2012. Available: http://undoingbias.csail.mit.edu/.

[37] However, also here the requirements for protection can become an obstacle, namely the requirement that the information is not known as a body or in its precise configuration and assembly of its components.

### 5.1. Secrecy

A basic requirement for protection of TS is that its object has to consist necessarily of an actual secret, not known in the concerned circles. More, the information has to be undisclosed, as a body or in the precise configuration and assembly of its components. Generally, the TS directive does not impose absolute secrecy as a standard for protection, as the use of the wording 'generally known' and readily accessible' denotes.[38] The information may be communicated or licenced under a non-disclosure clause and remain a TS. Machine learning, Artificial intelligence and data mining tools can pose a challenge to TS as legal concept in this regard.

As with other recent legal instruments dealing with data,[39] the EU legislator took a static view on the notion of know-how and business information. A static view can be easily applied to a chemical formula or a prototype but what about dynamic real-time information? Real-time data can be extremely valuable commercially, but depending on the information, and because we are talking about data mining tools in an online environment, in real time, it might not be possible to completely describe the precise configuration and assembly of its components. Specially in cases where information is combined from multiple, undetermined number of sources by the use of machine learning algorithms. Dueto the known back-box issues the human brain cannot reverse engineer and describe the components of the information. Often its commercial value comes exactly from being real time information constantly updated (for example information feeding dynamic pricing systems). An argument could be to interpret this provision in the sense of referring not necessarily to the complete precise configuration of the information in a static sense, but to information defined by categorization under defined criteria (e.g. customer mentions in social media, sales, etc) independently of its dynamic nature and the fact that the derived informational content (semantic meaning) produced is also not static.

Another element to consider, is that the information most not be 'generally known among or readily accessible to persons within the circles that normally deal with the kind of information in question'[40] Here it will be relevant whether we are trying to protect a selected data set from which to derive statistical correlations and extrapolations; processed information (final informational product) or methods for retrieving, processing and producing information. When the data sources are purely internal and kept in house it might be possible to argue successfully that the data is not being readily accessible. However, this might be difficult in many situations, when the data concerns for example market analysis: (a) consumer preferences, location, behaviour patterns, values, likes, tags and check-ins etc; b) active suppliers and c) competitors pricing, sale volumes, strategies, etc. A similar reasoning applies if the raw data is technical/scientific information, or when we are talking about Biobanks, medical records, clinical trials and data subject to regulatory oversight and transparency provisions. However, here due to a collision of different public interests the situation might be different. Raw data producers might have a policy or legal obligation to licence or grant access to the data. In some circumstances and depending on the actual industry sector and business model, it might

[38] See with further references Nuno Sousa e Silva (2014) 'What exactly is a trade secret under the proposed directive?' 9(11) IPJLP: 923-931, 929.

[39] See for example the approach followed in Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation) (Text with EEA relevance), *OJ L 119, 4.5.2016, p. 1–88, ELI: http://data.europa.eu/eli/reg/2016/679/oj.*

[40] Article 2 (a), TS dir.

be easier to protect the actual type of technical method used to process and obtain the information – the algorithm, or the specific combination of data sources.

Considering the above, I would argue for a broad interpretation of 'readably accessible' because the intention of the directive is to avoid creating an exclusive right over information.[41] Furthermore any information that cannot be considered as 'readily accessible to the concerned circles', but can be accessed by lawful means under article 3 TS Dir. would have to be considered as excluded from the scope of protection. Even when the data is not public information, the raw material - the data points used to train an AI or perform specific big data analysis that result in a given type of information - might be available to everyone in the concerned circles using a similar data analysis technology (as long as they obtain a license to use it). It may however be possible to argue the existence of secrecy, even when part of the data is known on the concerned circles, based on the fact that the standard is set to the body of information or its precise configuration and not the sum of its elements. If the specific data set configuration is undisclosed, then secrecy can be established.

A related but different issue that might be important concerning AI and Big data TS is to determine the standard of this notional person 'persons within the circles that normally deal with the kind of information in question'. Should we presume that this notional person is a data technology company? Or perhaps it should be constructed instead as a company in the concerned industry or sector of activity using the same AI technologies, e.g. data mining and machine learning tools? I would suggest that again here (and depending on what are we trying to protect) it will play an important role, whether the data mining and machine learning tools used (the specific algorithmic) are known to these 'concerned circles' or instead are in themselves a trade secret or protected by an IP right. It will also play a role what type of training data and input data is used, as caution is advisable when mixing in house data with external data.

The directive creates a single EU definition of TS, but however implementation differences are possible and likely since MS are allowed' to provide for more far-reaching protection' provided that limitations are established by the directive are complied with.[42] Does this means that the notional 'persons within the concerned circles' do not have to be evaluated by reference to the entire EU, but rather by reference to national commercial circles in the relevant sector?

In theory, the fact that subject-matter can be added in national implementation and that subject-matter may be excluded also from protection in the public interest by national authorities[43] favours a national approach. On the other hand, not only such appears to contradict the objectives of the directive harmonization and construction of a single market, as the nature of the subject-matter – information, does not allow for national compartmentalization. Furthermore, notional persons in EU law are usually constructed by reference to the entire EU internal market.

The directive intention is to create a uniform TS right concerning its object and basic scope of protection. However, it is a separate issue whether a specific and concrete piece of undisclosed know-how or business information is or not considered as 'not generally known or readably accessible to persons within the circles that normally deal with such information'[44]. Indirectly, constructing this notional person by reference to the EU as a whole, instead of a national standard, strengthens the

[41] Recital 16, TS Dir.
[42] Article 1, TS Dir.
[43] Article 1 (2) (b) and (c), TS Dir.
[44] Article 1 (1) (a), TS Dir.

scope of protection. Even in the days of the information society, due to language differences, access to ICT tools and other factors, information may be known locally but not generally known in the relevant industry across the EU. The construction of this notional person has also procedural implications, due to rules governing the burden of proof. An EU wide construction of the relevant circles will tend to favour right holders when enforcing their TS rights. Parties challenging the status of information as a TS will have to present evidence that the information in known in the relevant circles throughout the whole EU and such will be more difficult and more unlikely to obtain than if the relevant circles are to be understood nationally.

However, given that know-how and commercial information is often intrinsically linked with a given activity or business model, the notion of concerned circles should be restricted to those undertakings engaged in producing and/or offering similar goods or services. Thus, for example, if such activity is related to local products, traditions and culture, eventually legally protected as such,[45] the relevant circles will necessarily be geographically restricted to a region in one or a few MS.

This question is also related to the issue of determining whether a once TS ceases to fulfil the secrecy requirement once it has been acquired by one or a few undertakings in the concerned circles (e.g. through reverse engineering) or when it has been revealed under confidentiality obligation to a considerable number of entities (e.g. through licencing agreements), but has not been further disclosed. In the first case the information cannot be considered generally known nor readily accessible.[46] As for the second case, it would seem a paradox that non-exclusive licensing of a TS would result in loss of legal protection, as the objective of the directive is to contribute to reduce barriers to the free flow of information in the internal market.[47]

### 5.2. Causal link between secrecy and commercial value

Article 2 TS dir. states that the information retrieved has to have commercial value because it is secret. Establishing a causal link between secrecy and commercial value might be more or less easy depending on the nature of the undisclosed information. For example, the secret may concern concerning a physical entity (e.g. a prototype in the company vault), static technical information (e.g. a list of ingredients or chemical formula) or at least more or less stable information (e.g. a list of clients or suppliers) or real-time big data analytics inferences (e.g. predictions on future market behaviour). Concerning big data establishing this causal link between secrecy and commercial value in litigation, might be a procedural shoot in the dark. Being that real-time data is particularly problematic. Expert opinions may be used but these can be easily rebated by a divergent expert assessment. Market value could be an argument, if the TS has been object of licensing or similar data has been transitioned. When the commercial value of secret information is derived from facilitating

[45] Examples are protected designation of origin (PDO), protected geographical indication (PGI), and traditional specialities guaranteed (TSG). See: Regulation (EU) No 1151/2012 of the European Parliament and of the Council of 21 November 2012 on quality schemes for agricultural products and foodstuffs, *OJ L 343, 14.12.2012, p. 1–29.*

[46] Pires de Carvalho argues that secrecy under article 39 TRIPS is maintained as long as the last competitor withing the circle that normally deals with the information is unware of the information. See: Nuno Pires de Carvalho, *The TRIPS regime of Antritrust and undiscloused information* (Kluwer Law International, 2008), p. 233; Cf. With I Meitinger, 'Art. 39 TRIPS' in T Cottier and P Véron, Concise International and European IP Law (Wolters Kluwer Law International, 2011), p. 115 arguing that there is a limit above which the information is in the public domain.

[47] See Recitals 2 and 3. Cf. Sousa e Silva above n. ? , p.929, defending that 'there is a limited number of licensees a trade secret can have before it becomes generally known'.

a decision-making process. Another argument is competitive advantage granted by, or the investment made in acquiring or creating the object of the TS. Recital 14 mentions that the TS has to have a value and that this value may be either actual or potential. Therefore, it might be enough to demonstrate potential value. Examples of commercial value are mentioned in recital 14, and relate to situations when unlawful acquisition, use or disclosure of know-how or information 'is likely to harm the interests of the person lawfully controlling it, in that it undermines that person's scientific and technical potential, business or financial interests, strategic positions or ability to compete.'[48] Thus, it seems that the directive has chosen a low threshold for the requirement of commercial value.[49] An additional debate is on whether the information needs to be accurate to have commercial value or not, and whether information that is vague or non-understandable to humans has commercial value (e.g. AI training data). As mentioned before a secret can be negative or even void of material content and still have commercial value, e.g. the TS could precisely be that, unlike common belief, there is no secret (ingredient).[50] Veracity is not a static concept, what is valid knowledge today will perhaps be scientifically disproved next year. While it is difficult that incorrect or false information may have commercial value, that is not necessarily true in all situations (e.g. a list of client profiles containing only alias will still be commercially valuable if it allows communication and advertising). Knowledge that a certain information is incorrect contrary to common believe in the concerned circles, is valuable information providing a competitive advantage, and as such negative secrets should also be within the scope of the right. Furthermore, recital 14 mentions that commercial value can be potential and does not need to be actual, thus it should enough that the TS owner believes the information to be correct or that the incorrect information is potentially valuable.

A trade secret does not need to have a technical character, nor do they need to perform a function and thus veracity, cannot serve as legal criteria by itself. To base an assessment of commercial value on veracity or accuracy of the data would narrow down the scope of protection in a sense that does not appear compatible with its objectives

### 5.3 Duty of diligence

Finally, TS protection is conditional on a duty of care in preserving confidentiality: the information must have 'been subject to reasonable steps under the circumstances, by the person lawfully in control of the information, to keep it secret'. The standard for such duty is set to what is reasonable under the circumstances. From a practical standpoint reasonable steps will depend on the sector of activity, proportionality between the commercial value of the secret, the annual income of the company and the expenses incurred with protective measures, etc. A small restaurant cannot be expected to protect their 'secret sauce recipe' with the same measures as a multinational in the food industry. Proportionality and weighting up the interests of the parties, third parties such as consumers and public interest is established as a guiding principle for the competent judicial authorities to take into

[48] Recital 14, TS Dir.

[49] See Aplin et al, arguing that the mere fact that enforcement is sough is generally considered enough evidence that the secret has commercial value to its owner. Tanya Applin et al, *Gurry on breach of confidence: the protection of confidential information* (Oxford University Press, 2nd ed. 2012), p. 803.

[50] See F Dessementet, 'Protection of Trade Secrets and Confidential Information', in CM Correa and AA Yusuf (eds), Intellectual Property and International Trade: The TRIPS Agreement (2nd ed Wolters Kluwer 2008) 270,281. Arguing that false or misleading information may be protected; Cf Sousa Silva at n arguing that 'No information knwon to be false will enjoy (objective commercial values and is thus excluded from protection by virtue of that requirement.'

consideration when adjudicating in such matters.[51] General efforts to keep confidentiality will include technical and legal measures such as contracts and non-disclosure agreements, compartmentalization of information, dedicated servers, firewalls, encryption, and similar measures.[52]

In the specific context of AI and Big data again the directive's static approach to the nature of information, is likely to cause interpretation questions as to what measures will suffice to meet the due diligence threshold. The use of the wording 'under the circumstances' could open up for interpretation in light of the technical state of the art. A point of concern is what measures are technically possible to ensure that TS holders prevent access to the undisclosed information. Here the dynamic nature of big data is likely to aid a claim for reasonable steps. The question is whether it is technically and legally possible to prevent, delay or hinder the possibilities for independent discovery and reverse engineering.

The directive and its recitals show strong concern in ensuring that the scope of protection will not extend to the information in itself, but rather will focus on its secrecy.[53] It is the secret character that is protected against 'unfair methods of disclosure'. The information in itself is only protected as long as it has not been lawfully disclosed. As Drexl et al. point out the directive does not create a property right over information, but a system of liability for a specific tortious conduct – because it requires unlawful conduct in acquiring the information.[54] Reverse engineering and independent discovery are possible and incentivised, as long as these do not constitute unfair commercial practices.[55] Thus, reasonable steps would be any measure that, with respect for exceptions, result in making lawful acquisition impossible (e.g. contractual clauses), illicit (e.g. technical or physical barriers), or under the circumstances imply a considerable effort or are disproportionally onerous.

## 6. Can Trade secrets survive AI and big data analytics?

As previously mentioned, the protection of undisclosed know-how and business information has limitations and does not result in an absolute property right. Recent harmonization in the EU, still leaves open the debate on the nature of such right. Theoretical discussions subsist on whether trade secret rights are a new IP right conferring a right to exclude others; if it is still a right created under the framework of unfair commercial practices, or if instead such dichotomy should be abandoned and

[51] Recital 21, TS Dir; see also article 13 TS dir.

[52] Recent decisions for national courts point to different approaches. For example, Austrian Supreme Court, Decision No 4 Ob 165/16t of 25 October 2016, considered that under national law the trade secret holder had adequately demonstrated their intention to keep the information secret by maintaining a logging system with a username and protected by a password, and ensuring that only a limited number of identified persons knew the information; while in Spanish case Civil Judgment No 441/2016, Provincial Court of Madrid, Section 28,Rec 11/2015 of 19 December 2016, the court found the steps to avoid disclosure should be adequate and reasonable and directed both externally and internally. External steps should prevent third parties from gaining access to the secret and internal steps should limit the number of employees and collaborators who know or have access to the information. Established Case law from England and Wales point out that "reasonable steps" require establishing an obligation of confidence by providing notice that the information is confidential (see: Coco v A.N. Clark (Engineers) Ltd [1968] FSR 415), ant that secret holders should limit the dissemination of the secret or at least not encourage or permit widespread publication (See Lansing Linde v Kerr [1991] 1 WLR 251.

[53] Articles 3 and 4, recital 16, TS Dir.

[54] Josef Drexl et al, (2016) 'Data Ownership and Access to Data: Position Statement of the Max Planck Institute for Innovation and Competition of 16 August 2016 on the Current European Debate', p. ?

[55] Recital 17, TS Dir.

TS should be considered a new *sui generis* immaterial right.[56] The answer to this debate is bound to have procedural implications under the applicable national procedural rules and thus should not be underestimated because procedural rules are vital aspects of enforcement. Arguably, there might be some dissent in the implementation phase that slowly will have to be clarified by judicial intervention or further legislative harmonization efforts. However, it is undeniable that the TS dir. has created a specific framework for the protection of undisclosed know-how and business information, harmonizing legislations and introducing fundamental changes in the way information was protected.

Conventional wisdom would say that despite legal protection, in the future few secrets are expected to survive the growing use of machine leaning and artificial intelligence. Innovators will be confronted with the question of how to protect a TS from AI when the directive clearly points out to a scope of protection limited to protecting the information from unlawful and unfair acquisition. This section concerns the scope of protection of TS and will exam how far can the legal framework protect TS owners from the use of advanced information technologies, and whether technology may create challenges to the legal framework.

### 6.1. Scope of Trade secrets rights: lawful acquisition, lawful use and lawful disclosure

Lawfulness of acquisition, use or disclosure of a TS is possible if required or allowed by law.[57] While acquisition of TS is considered lawful if obtained by independent discovery or creation and reverse engineering, to exercise workers' rights to information, and by 'any other practice which, under the circumstances, is in conformity with honest commercial practices'.[58]

Generally, article 3 (2) TS Dir. considers lawful any acquisition, use or disclosure that is either required or allowed by either EU or national law. Reinforcing subject-matter exclusions and enforcement exceptions. This entails that although TS protection is harmonized and a single definition of the scope of protection now exists, it is possible that information acquired lawfully under one jurisdiction might not be able to be disclosed or used without infringement in another jurisdiction. TS protection will also have to be balanced against rules that impose duties to reveal information[59]and rules that provide for exceptions to enforcement.[60] First, article 2 TS Dir. creates limits to the subject-matter, restricting the right to TS protection in situations of conflict with fundamental rights protected by the EU charter, namely freedom of expression and information.[61] It also excludes from protection under the TS directive, regarding information that has been requested to be disclosed (to either the public or public authorities) under EU or national rules for reasons of public interest. Likewise, information submitted to EU or national public institutions may be disclosed in accordance to the law. The directive also distinguishes employee information, experience and skills from their employer information and know-how. Secondly, rticle 5 TS Dir. establishes exceptions to enforcement for exercise of free speech and right to information, whistleblowing activities, exercise of workers' rights and any other legitimate interest recognised by EU or national law.

---

[56] See above for an opinion on this matter.
[57] Article 3 (2), TS Dir.
[58] Article 3 (1), TS Dir.
[59] Article 2, TS dir.
[60] Article 5, Ts dir.
[61] Article 11, Charter of Fundamental Rights of the European Union, OJ C 326, 26.10.2012, p. 391–407, ELI: http://data.europa.eu/eli/treaty/char_2012/oj.

Article 3 (1) TS Dir. further restricts the scope of protection in line with the traditional idea of protecting competition and fair commercial practices. A first point to note is that article 3 (1) TS Dir. only mentions the acquisition of a TS and not its use or disclosure. However, the use and disclosure of TS lawfully acquired are not necessarily lawful (e.g., TS acquired lawfully under a confidentiality agreement).[62]

The question of independent discovery or creation and reverse engineering are of particular relevance to analyse in light of AI and big data analytics. Industrial espionage (including cyber threat and hacking) and competition by former employees are the traditional routes for trade secret misappropriation and misuse. However, a more insidious threat is lurking in the back. Data mining techniques and predictive algorithms are capable of revelling undisclosed personal and business information even when all possible efforts are made to keep it secret. For example, studies claim to be able to detect sexual orientation by analysis of photos posted on social media;[63] criminal tendencies;[64] political ideas;[65] suicide prevention algorithms;[66] pregnancy detection.[67] It is not farfetched to imagine that correlating information from multiple sources, might reveal valuable information concerning strategic market positioning decisions and on-going research projects. Information on pricing, client list, suppliers, distribution routes and networks, manufacturing capability and processes can also likely be inferred. Even extremely complex process or products are likely to be reversed engineered in a fast and low-cost manner. Making TS protection more difficult and less reliable.

Independent discovery throug the use of AI and big data analytics will be lawful if access to the data is not considered unlawful. Reverse engineering is a form of lawful acquisition only if it refers to 'observation, study or testing of a product or object' that is either made available to the general public (release in the market in some way) or that has been acquired without any clauses limiting the acquisition of TS. Namely, this condition will be fulfilled if the source data, considered in isolation would not reveal any protected information, for example: purchases, job advertisements, employees' professional profile and social networks, outsourcing research agreements and other outsourcings contracts, scientific publications and patent fillings. Multiple sources of data acquired lawfully, combined and subject to big data analytics will create inferences and predictions. These are likely to

---

[62] See Article 4 (3) and (4), TS Dir.

[63] Stanford University study Wang, Y., & Kosinski, M. (in press) 'Deep neural networks are more accurate than humans at detecting sexual orientation from facial images' Available: https://osf.io/zn79k/. Kosinski is known for his work with Cambridge University on psychometric profiling, including using Facebook data to make conclusions about personality. However, the science has been publicly contested both in the press and academia. See for example: https://www.dailymail.co.uk/sciencetech/article-5270365/Google-experts-debunk-sexuality-detecting-AI.html; https://www.theguardian.com/technology/2017/sep/07/new-artificial-intelligence-can-tell-whether-youre-gay-or-straight-from-a-photograph; Paper debunking these findings: https://medium.com/@blaisea/do-algorithms-reveal-sexual-orientation-or-just-expose-our-stereotypes-d998fafdf477.

[64] Xiaolin Wu and Xi Zhang's paper, "Automated Inference on Criminality Using Face Images", submitted to arXiv (a popular online repository for physics and machine learning researchers) in November 2016. These finding are also contested; Cf. paper debunking the findings: https://medium.com/@blaisea/physiognomys-new-clothes-f2d4b59fdd6a

[65] https://www.theguardian.com/technology/2017/sep/12/artificial-intelligence-face-recognition-michal-kosinski

[66] Torous, J., Larsen, M.E., Depp, C. et al. 'Smartphones, Sensors, and Machine Learning to Advance Real-Time Prediction and Interventions for Suicide Prevention: a Review of Current Progress and Next Steps' Curr Psychiatry Rep (2018) 20: 51. https://doi.org/10.1007/s11920-018-0914-y.

[67] For example, iIn 2012, news outlets advanced that a US based grocery store was using an algorithm to provide clients with relevant discount coupons, including uncovering client's pregnancy and predicted due date. Charles Duhigg 'How Companies Learn Your Secrets' The New York Times, FEB. 16, 2012.

reveal TS lawfully , because the source data is not unlawfully acquired or used. However, in some circumstances, it may be possible to argue whether such is an unfair commercial practice.

An important question of legal interpretation concerns the validity of anti-reverse engineering clauses. The TS Dir. Specifically mentions reverse engineering as a means for lawful acquisition of a TS.[68] MS cannot deviate from this norm and have to include reverse engineering as a lawful means of acquisition of a TS.[69] However, the TS dir. ais also quite clear in allowing contractual limitations to the disclosure or use of a TS, even when such information was obtained lawfully under rticle 3 TS Dir. It can be debated, whether the directive consecrates a general right to reverse engineer, and if so whether such right can be object of contractual waiver. Recital 16 TS Dir. clearly states, both that reverse engineering can be subject to contractual prohibition and that such contractual provisions may be limited by law. Neither the directive nor its recitals mention a right to reverse engineering, meaning that undertakings can unilaterally establish in their products or services terms and conditions of sale or use prohibitions to reverse engineering.

The text of the directive excludes anti-reverse engineering clauses in products or objects generally made available to the public, but distinguish situations where acquisition is subject to anti-reverse engineering clauses.[70] In practice reverse engineering cannot be restricted in mass sale consumer goods, but TS owners are allowed to insert anti-reverse engineering clauses in terms of service or use, and in contracts with other undertakings within the production chain. TS owners can also restrict the use and disclosure of TS revealed by lawful reverse engineering with contractual clauses, these are allowed and expressively mentioned under the directive[71]. However other areas of law have to be considered in this analysis. In particular, concerning digital technologies, reverse engineering (decompilation) of computers programs protected by copyright is only allowed for the purpose of achieving interoperability.[72] Since this rule has not been changed, it is a strong indication that restrictions to reverse engineering are not generally precluded.

Furthermore, the validity of such clauses might be found abusive under competition law or considered an unfair commercial practice under national law. Despite TS harmonization, MS remain free to maintain an additional layer of protection in the form of general clauses concerning unfair commercial practices and some very broad anti-reverse engineering clause might fall under such regulation. The TS Dir. interpreted in light of Recital 16 TS Dir. clearly allows member states to restrict anti-reverse engineering clauses. While, article 1(1) TS Dir. only allows MS to provide for more far-reaching protection to TS, it also limits this possibility to compliance with a most of the provisions of the TS Dir.[73] Article 3 TS Dir., interpreted in light of recital 16 TS Dir. does not create a right to reverse engineering, it merely allows it unless otherwise established by law or contractual disposition,

[68] Article 3 (1) (b), TS Dir.
[69] Article 1 (1), TS Dir.
[70] Article 3 (b), TS Dir..
[71] Article 4 (3) (b) and (c), TS Dir.
[72] Article 6, Directive 2009/24/EC of the European Parliament and of the Council of 23 April 2009 on the legal protection of computer programs (Codified version) (Text with EEA relevance) OJ L 111, 5.5.2009, p. 16–22, ELI: http://data.europa.eu/eli/dir/2009/24/oj.
[73] Member States may provide for more far-reaching protection against the unlawful acquisition, use or disclosure of TS under the condition that national legislation ensures compliance with Articles 3, 5, 6, Article 7(1), Article 8, the second subparagraph of Article 9(1), Article 9(3) and (4), Article 10(2), Articles 11, 13 and Article 15(3).

meaning that anti-reverse engineering clauses are allowed. Imposing in national law a complete prohibition of anti-reverse engineering clauses is precluded since these are allowed under article 3(1)(b) TS Dir., in fine.

However, as it will also be discussed in the next section, the provisions of article 4 TS Dir. can be subject to national deviation in order to create additional TS protection. Meaning that the scope of what is considered unlawful acquisition, use and disclosure of TS cannot be reduced, but can be expanded in national law. It follows from article 4(2)(b) TS Dir. that TS acquisition shall be considered unlawfully if carried out by 'any other conduct which, under the circumstances, is considered contrary to honest commercial practices.'[74] Because MS are allowed to provide for more far-reaching protection and in doing so are not limited to comply with article 4 TS Dir.[75] in theory, there is room for MS to consider certain types of broad anti-reverse engineering clauses prohibited for being unfair or contrary to honest commercial practices.

### 6.2. Scope of Trade secrets rights: Unlawful acquisition of TS

Unlawful acquisition of TS occurs when a TS is obtained without consent of the TS holder, and cumulatively it stems from 'unauthorised access to, appropriation of, or copying of any documents, objects, substances or electronic files, lawfully under the control of the trade secret holder, containing the trade secret or from which the trade secret can be deduced' [76]; or 'any other conduct which, under the circumstances, is considered contrary to honest commercial practices.'[77]

The first major point to consider and more susceptive to divergent interpretation, is to determine in practice what is the concept and scope of the requirement 'under the control of the TS holder' in a digital sense and how this will be interpreted by courts. It might be relatively strait forward to determine if a physical object or static information is under the control of a TS holder – f. ex. something is locked in a room and only a few persons have access to the key, or in similarity it is in a server protected by passwords, firewalls and encryption. Even so, advances in digital technology increase the possibilities to circumvent technical measures to prevent access. Such, includes also the danger that, when available, quantum computing will entail the practical end of all existing forms of encryption. These technical considerations are linked to the duty of diligence in keeping secrecy, in which the specific technical and factual circumstances are relevant. However, it is a separate question, because reasonable steps to maintain secrecy is not equivalent to the information being under the control of the TS holder. A question may arise on whether information stored in external servers or cloud services is in fact under the control of the TS owner. In similarity, if a company subcontracts documentation services or if a start-up uses shared facilities, digital services or communication networks. In order to keep an appropriate balance, the scope of the concept of control should be low, meaning that it should be interpreted as not imposing a duty to set in place disproportionably expensive or impractical solutions, as the purpose of this requirement is to ensure legal certainty. Contractual provisions imposing secrecy, labelling documents as secret, physical barriers and passwords should remain examples of measures to assert control, as long as these are effective in ensuring that access requires an active conduct on a third party that requires effort. The higher the

[74] Article 4 (2) (b), TS Dir.
[75] See Article 1 (1) second paragraph, TS Dir.
[76] Article 4 (2) (a), TS Dir.
[77] Article 4 (2) (b), TS Dir.

effort and investment in circumvention of protective measures installed by TS owners the easier will be to offer evidence of unauthorized access, as accidental or incidental findings cannot no longer be used as a defence. Furthermore, circumvention of protective measure would qualify as contrary to honest commercial practices.[78]

But what about situations where the trade secret was obtained by big data analytics harvesting public information, or information made public – social media profiles of consumers and companies? Can a TS secret be claimed over such information and if so, can it be maintained secret? It may be very difficult to argue that the right holder was in control of the data harvested and contained in a multitude of files in social media, or even in public repositories of statistic information.

Even when we are talking about internal company information, it might be difficult to assert control over the data from which the trade secret can be deducted. For example, real time information on transport of raw materials, goods delivery times to stores, sales volumes of each specific type of item, prices, expenses, profits and so on. The volume of data is so large that it has to be decentralised using for example cloud services. Data will also be taken from a multitude of sources: sensors on goods, shipping manifests, invoices, internal and external communications, etc. The data is known by multiple persons and entities, some internal, some external and imposing contractual secrecy obligations may not always be possible. Furthermore, the true potential of bid data analysis is to be able to establish correlations from a large number of variables. This means that such analysis will tend to include multiple sources of external data that is partially or completely outside the control of the TS owner, e.g. weather reports, traffic accidents, strikes, timetables and delays, post office reports, social media.

An argument, TS Owners may resort to, would be to interpreter the provision as focusing on control of the outcome of the big data analysis and not on control of every source of raw data. Meaning that control of an essential part of the raw data would be enough to prevent lawful deduction of the trade secret. However, such may result in expanding the scope of protection beyond what is contemplated by the directive legislator. Another option, as discussed above, is also to argue that the TS discovered by a third party using big data analytics was acquired through unfair or dishonest commercial practices, contrary to honest commercial practices. In particular in situations where the discovery cannot be said to be accidental and considerable investment is made in training an AI to discover one or more TS.

However, under article 4(3) TS Dir. the use and disclosure of a TS will only constitute an infringement if the secret was obtained in an unlawful manner, by breach of confidentiality agreement or any other duty not to disclosure, or breach of contractual obligations limiting the use of the TS.[79] Meaning that the directive opens the door to inter-part limitation of lawful means of TS acquisition by independent discovery or creation and reverse engineering, e.g. in licencing agreements or terms of use. Contractual dispositions can and will likely be used as a tool to prevent the use or disclosure of TS acquired by big data analytics.

Finally, a question may arise as to whether the use and disclosure of TS is unlawful if there is not actual knowledge that the information obtained is a TS. The directive imposes that civil redress must be available,[80] it does not textually establish any particular type of liability, however the unlawfulness

[78] Article 4 (3), TS Dir.
[79] Article 5 (3), TS Dir.
[80] Article 6 (1), TS Dir.

of the conduct does not depend on any subjective element on the part of the alleged infringer. TS holders can apply for provisional measures and injunction preventing or demanding the cessation of the infringing conduct based on objective unlawful conduct. However, measures, procedures and remedies should be applied following a proportionality principle,[81] and subjective elements will play a role. Likewise, concerning damages if the infringer knew or ought to know that the conduct, but specific intent is not necessary.[82]

7. Conclusion:

The use of Trade Secret protection for big data in light of the EU TS Dir. might face hurdles in some instances. The good news is that the dynamic nature of emerging digital technologies also offers technical options for trade secrecy by design. Confidentially agreements, non-disclosure clauses and other contractual obligations will remain necessary. Physical and technological measures to prevent disclosure and assert 'control' over valuable the information will also be advisable whenever possible.

TS protection operates against unlawfully acquisition, use or disclosure of commercially valuable information that had been maintained secret. TS protects only against unlawfully access and not against independent discovery and reverse engineering. The concept of unlawful access depends on maintaining control over actual or potential sources of information. In the age of AI, machine learning and big data analytics protection of TS will depend more and more on measures to control company data. Even trivial data can be a problem, multiple sources of trivial data may allow inferences and predictions concerning business information that gives competitive advantage to competitors. Compartmentalization of information to workers, collaborators, suppliers, clients and the market is no longer in isolation a sufficient measure of precaution. Confidentiality agreements and legal duties not to disclose and not to use the trade secret more than ever will constitute the cornerstone of TS protection. Moreover, such contracts and agreements should, whenever possible, extend their object to cover also all possible data sources.

In the age of big data and AI this means a higher standard of the duty of care. Reliable measures to keep the secrecy of valuable know-how and business information, need to cover also data from which the TS can be deducted. TS protection plans and strategies will be essential, including clear inventory of what information such be kept undisclosed and want data might reveal such information. Preventive measure will entail also reversing potential reverse engineering possibilities in order to implement technical measures to make it harder or even impossible, as well as legal measures in the form of non-reverse engineering clauses.

The TS dir. re-opens several issues of legal interpretation, including the EU interpretation of duty of diligence, scope of the TS right and the legal admissibility of clauses prohibiting reverse engineering. Given the growing importance of TS protection, and the issues raised by emerging information technologies, such as AI, and Big data analytics, these questions are likely to occupy courts and legal scholars in the coming years. The answers will tell us whether trade secret protection in the age of AI and Big data is or not a legal oxymoron.

[81] Article 7, TS Dir.
[82] Article 14, TS Dir.