Building the European Data Economy

Data Ownership

WHITE PAPER

1 January 2017

Bird & Bird

Benoit Van Asbroeck
Julien Debussche
Jasmien César
Legal Notice

This paper reflects the views only of the authors, and not those of the Commission, Bird & Bird LLP or the other consortium members. The Commission, Bird & Bird LLP and the other consortium members cannot be held responsible for any use that may be made of the information contained herein.

This paper is a report on the current state of the law and not a prospective study (although it may include some critical view and prospective ideas). Accordingly, this paper does not represent a final state of mind of its authors; it only intends to encourage discussion on the topics covered.

We wish to thank the persons who provided critical insights and guidance for various parts of the paper, in particular (in alphabetical order): Layla Bakker, Alexander Benalal, Anne Federle and Natalia Zebrowska.
Executive Summary

The EU Commission has voiced on multiple occasions the most important legal issues in a data environment. In its data-driven economy Communication of July 2014, but also more recently in the context of its 2016 free flow of data initiative, it highlighted that "barriers to the free flow of data are caused by the legal uncertainty surrounding the emerging issues on 'data ownership' or control, (re)usability and access to/transfer of data and liability arising from the use of data".

In this context, the present report examines the data ownership and data sharing aspects from a legal perspective, looking at the current EU acquis. In particular, the research has been performed in a commercial context, relying on the concrete circumstances of the TOREADOR big data analytics project and its four use-cases.¹ The results of our research are not only aimed at the partners of the aforementioned project, but also strive to provide additional evidence to policymakers on the emerging issues of data ownership.

This report provides an overview of the EU acquis in relation to data management and in particular on access to, ownership, exploitation or exchange of data. To this end, we provide a diagram with a snapshot of the current EU framework. Such mapping demonstrates the various categories of EU legislation related to data and potentially playing a role with respect to the possible "ownership" of data, but also with respect to the restrictions and requirements that may apply in relation to certain types of behaviours or certain types of data.

More specifically, our in-depth analysis has allowed concluding the following:

- There is at present no EU legislation that specifically regulates the question of ownership in data. Such absence of ownership-related legislation does however not exclude the fact that there are numerous legislations that have an impact on data or that may confer some kind of protection to certain types of data or on datasets (i.e., copyright, database rights and trade secrets).

- The case-law at EU level does not recognise explicitly an ownership right in data. However, according to some authors, the Court of Justice of the EU opened the door for a discussion on ownership in intangible assets in its UsedSoft judgment issued on 3 July 2012. Despite such ruling and the possible interpretation deriving from it, a high legal uncertainty remains.

¹ TOREADOR is a big data EU research and innovation project funded by the European Commission: http://www.toreador-project.eu/.
While it was found that at national level there are also no legislations relating to data ownership, the existing case-law in some Member States addresses, to some extent, the issue of data ownership. In the same vein, in some countries, legal scholars debate the question of ownership in data, suggesting a novel interpretation of existing civil law provisions.

While the protection of individuals' privacy has triggered a lot of attention in the past months with the adoption of the General Data Protection Regulation, this report does not examine in depth the issues related to privacy and data protection. This being said, this report takes an approach whereby ownership is examined in relation to data, be it non-personal or personal data. Although some scholars suggest otherwise, we take the position that personal data is not necessarily owned by the individual and thus that an "ownership" right in data for data controllers or processors cannot be excluded. Such ownership would however be subject to the individual's control over his/her personal data.

There has been a lot of attention from antitrust regulators and scholars for the various legal issues surrounding data and competition law. In particular, the difficulty lies in the fact that antitrust regulators still apply the legal principles of the nuts-and-bolts world to a reality in which data are the new corporate assets. Even though some solutions emerge in the framework of competition law, a lack of legal certainty remains in relation to key competition law notions and their practical implementation in the data-driven economy.

Numerous legislations that may impact a company's control of, the access to, or the rights in data were identified. Such legislations regulate data sharing obligations in various ways and depending on various factors, such as the sector concerned or the reasons of public interest that have led to the adoption of the instrument (e.g., public security, public health, consumer protection, etc.). However, such legislations remain mute on the ownership in the data concerned.

The ownership-like rights currently available are limited to intellectual property rights and trade secrets. However, none of these allow providing an adequate protection of (ownership in) data. With respect to copyright, while there are several features that can be seen as beneficial for the protection of data (e.g., long-term protection, broad exclusive rights, disclosure of data permitted), there are numerous disadvantages that hinder the protection of data by copyright (e.g., originality requirement, territoriality, exclusivity). As regards database rights, it is difficult for the current data economy to accommodate the dual protection in the EU (copyright and 
\textit{sui generis}). On the one hand, the copyright protection on the structure has become moot. On the other hand, the \textit{sui generis} protection awarded for the investment does not protect the data as
such. Trade secrets protection, finally, is not fit for purpose as it was created for other reasons than the blanket protection of all data and requires the information to remain secret.

On the basis of the abovementioned findings, it is established that the existing legal framework in the EU is not optimal and does not sufficiently facilitate operations on or including data. In the same vein, it is concluded that the cumulative implementation of the current maze of different possibly applicable legislations is a significant hurdle to the uptake of data analytics in the EU and is creating legal uncertainty in this fast-growing market. It results from this situation that those involved in the data value cycle may currently hold back on data sharing initiatives and presently have no choice but to rely on contractual arrangements to manage their rights in data.

While relying on contracts may seem to provide greater flexibility to the contracting parties, it was found that it nevertheless comes with various difficulties. In particular, the lack of harmonisation of contract law in the EU, but also the limits of contractual arrangements towards third parties and the issues related to the validity of data-related agreements, create a high legal uncertainty that affects the entire data value chain and all data flows. It is therefore concluded that such situation is not sustainable in a data-driven economy and with the fast-increasing development and adoption of data mining and analysis tools.

Against a background where the EU strives towards a data-driven environment in which both citizens and companies can reap the benefits of novel data technologies, but also against a background where the current legal framework does not sufficiently tackle all the issues related to data and where actors involved in the data value chain have no certainty as to the ownership of the data they have gathered, created, analysed, enriched or otherwise processed, we conclude that a more solid and legally secure solution is needed.

In such context, the last Chapter of this report suggests the creation of a non-exclusive, flexible and extensible ownership right in data(sets), with a data traceability obligation as a safeguard. Such Chapter discusses the specificities of said right and obligation, their interaction with the other existing rights in data, their incidence on civil law, and their possible reflection in contractual arrangements.
"Information Wants To Be Free. Information also wants to be expensive. Information wants to be free because it has become so cheap to distribute, copy, and recombine – too cheap to meter. It wants to be expensive because it can be immeasurably valuable to the recipient. That tension will not go away. It leads to endless wrenching debate about price, copyright, ‘intellectual property’, the moral rightness of casual distribution, because each round of new devices makes the tension worse, not better."

Stewart Brand, Hackers’ Conference 1984;
Contents

1. Introduction................................................................................................................. 10
   1.1 The EU Data Economy: Some Legal Issues .................................. 10
   1.2 Practical issues in a big data context .......................................... 12
   1.3 The Concept of Big Data ........................................................................... 13

2. Overview of the EU Legal Framework .............................................................. 15
   2.1 Introduction .............................................................................................. 15
   2.2 Ownership-like Rights ........................................................................... 17
      2.2.1 Concept of Ownership .................................................................. 17
      2.2.2 Actors in the Data Value Chain who Could Claim
            Ownership in Data ....................................................................... 18
      2.2.3 Lack of "Ownership" Right in Data ............................................. 22
      2.2.4 Intellectual Property Rights and Trade Secrets ................. 27
   2.3 Individuals’ Rights ............................................................................... 29
      2.3.1 Privacy and Data Protection ...................................................... 29
      2.3.2 Consumer Protection .................................................................. 32
   2.4 Competition Law Rights and Obligations ............................................. 35
      2.4.1 Characteristics and Key Competition Law Aspects of
            Big Data ...................................................................................... 36
      2.4.2 Dominance and Essential Facilities ........................................ 38
      2.4.3 Mergers and Acquisitions ......................................................... 40
      2.4.4 Conclusion on the Competition Law Aspects ....................... 42
   2.5 Data Sharing Obligations ................................................................... 43
      2.5.1 Introduction ................................................................................. 43
      2.5.2 Public Sector (open data) .......................................................... 48
   2.6 Other Data-related Legislation .............................................................. 56

3. Intellectual Property Rights and Trade Secrets ............................................. 59
   3.1 Copyright .................................................................................................. 59
      3.1.1 Legal Framework ......................................................................... 59
      3.1.2 Copyright Protection: General Overview ................................. 64
      3.1.3 Copyright Protection of Data ...................................................... 68
      3.1.4 Exclusive Rights and Copyright Exceptions ......................... 70
      3.1.5 Contractual Aspects .................................................................. 75
Chapter 1

Introduction

1.1 The EU Data Economy: Some Legal Issues

This report is drafted in the context of the EU-funded research and innovation TOREADOR project. It aims to support organisations lacking in Big Data expertise in managing Big Data Analytics ("BDA"), by providing a Model-based BDA-as-a-Service (MBDAaaS) framework.

The TOREADOR project falls within the Digital Single Market ("DSM") strategy of the European Commission (6 May 2015), which notably aims to clarify emerging issues of data ownership, access and liability.²

This builds on the Commission's Communication entitled "Towards a thriving data-driven economy" published in 2014 already.³ This Communication expected the Big Data market to grow worldwide to USD 16.9 billion in 2015 at an annual rate of 40%. The Commission nonetheless also indicated that the EU had been slow in embracing this revolution and that the complexity of the legal environment and the insufficient access to large datasets created entry barriers to SMEs and stifled innovation.

The 2014 Communication addressed the various challenges by sketching the features of the European data-driven economy of the future and drawing some conclusions to support and speed up the transition towards it. It notably concluded that to be able to seize the opportunities related to a data-driven economy and to compete globally in such economy, the EU must "make sure that the relevant legal framework and

³ Commission, 'Towards a Thriving Data-Driven Economy' (Communication) COM(2014) 442 final
the policies, such as on interoperability, data protection, security and IPR are data-friendly, leading to more regulatory certainty for business and creating consumer trust in data technologies”.

In a section dedicated to the regulatory issues, the Communication further highlighted the issues related to personal data protection and consumer protection, data mining and security. It also raised the concerns pertaining to the ownership and liability of data provision and data location requirements in various sectors that limit the flow of data.

In its 2015 Staff Working Document related to the DSM, the Commission reiterated the legal issues by putting forth problem drivers related to the data economy: “currently, collecting, processing, accessing and protecting data is a major challenge. This includes issues such as ownership of data, treatment of personal and industrial data, availability, access and re-use, contractual terms and conditions, data security, quality of data (e.g. timely updates), authentication of users, cybercrime, acceptance of electronic documents, liability for incorrect information, standardisation of languages and formats.”

Finally, in the context of the DSM, the EU Commission voiced its intentions to propose in 2016 a European ‘Free flow of data’ initiative that would notably address the restrictions on the free movement of data and the emerging issues of ownership, interoperability, usability and access to data in situations such as business-to-business, business-to-consumer, machine generated and machine-to-machine data.

Following a public consultation workshop in May 2016, the Commission has published in October 2016 an inception impact assessment. In such document, the Directorate General for Communications Networks, Content and Technology of the European Commission ("DG Connect") raises again that additional “barriers to the free flow of data are caused by the legal uncertainty surrounding the emerging issues on 'data ownership' or control, (re)usability and access to/transfer of data and liability arising from the use of data”.

---

4 Ibid 3
5 This Deliverable does not examine the issues related to liability, which remain nevertheless essential and ought to be researched further in a separate study.
1.2 Practical issues in a big data context

Although this report will examine the issues relating to data in general, it relies, where necessary, on the practical findings of the legal analyses of the TOREADOR big data project. It will therefore emphasise, where needed, some of the particularities of "big data" and will consider the legal issues having big data analytics technologies in mind.

More specifically, with the aim of providing a report that takes into account the realities of the market, the particularities of the four use-cases of the TOREADOR project were kept in mind. Several common characteristics can be drawn from use-cases, and in particular with respect to the aggregation of data (see Annex 1 for details about the use-cases):

- There is a vast volume of data being gathered from multiple sources and transferred onto the TOREADOR Platform for big data analytics purposes.
- The data is provided in variable formats.
- The data can be provided in unstructured forms.
- The data flow and analysis is fast and may be continuous.

Furthermore, the aggregated datasets at stake are being analysed in most instances in real-time. Such analysis is performed by dedicated BDA software and tools, with possible various kinds of analyses such as predictive analytics, correlation analysis or anomaly detection. The ultimate aim of the big data analysis performed in the context of TOREADOR is to increase the value of data and to enable European companies reaping the full benefits of BDA.

The common operations performed on the data are also worth highlighting:

- **Input**: data is gathered from multiple internal (e.g., IT systems, instruments and sensors) and external sources (e.g., clickstreams, open data, employee data, etc.).

- **Analysis**: data is processed and analysed externally via the TOREADOR data analytics platform using business intelligence and analytics applications but also data visualisation, machine learning and algorithms.

- **Output**: the analysed data is provided back to data provider for multiple purposes – internal (risk management, improvement of device functioning and maintenance, fraud detection, production process performance and improvement) and external (marketing, regulatory, etc.).
1.3 The Concept of Big Data

What is to be understood by "big data"? While there is no real consensus on a definition, the initial logical observation is that it is often described as a large dataset comprising different types of data that have grown beyond the ability to manage and analyse them with traditional tools.\(^9\) Hence, handling variable (un-)structured data in real-time requires the adoption and use of new methods and tools (e.g., processors, software, algorithms, etc.).\(^10\)

One could however not discuss the notion of big data without highlighting some of the key characteristics of big data, usually expressed with a series of "V's", and in particular:

- **Volume**: refers to the vast amount of data acquired, stored, searched, shared, analysed, visualised, generated and/or managed. Big data technologies have notably enabled the storage and use of large datasets with the help of distributed systems, where parts of the data are stored in different locations, connected by networks and brought together by software.\(^11\)
- **Velocity**: refers to the speed, which is of essence in a big data context. More particularly, it refers to the speed with which data is stored and analysed, as well as the speed at which new data is generated.\(^12\)
- **Variety**: refers to the heterogeneous types of data that can be analysed, combining structured but also unstructured datasets. There are unanimous findings that most of the data being generated and analysed today is unstructured.

In addition to these three key features, several authors also refer to "Veracity" which relates to the ability of analysing datasets that

\(^10\) Commission, 'Towards a Thriving Data-Driven Economy' (Communication) COM(2014) 442 final, 4
In this context, it shall be reminded that the TOREADOR Platform will also indirectly address issues related to novel data structures, algorithms, methodology, software architectures, optimisation methodologies and language understanding technologies for carrying out data analytics, data quality assessment and improvement, prediction and visualisation tasks at extremely large scale and with diverse structured and unstructured data.
In the context of TOREADOR, real time cross-stream analysis of very large numbers of diverse and multimodal data streams will notably be addressed.
comprise less controllable and accurate data. The "V" of "Value" has also been highlighted to refer to the ability of turning data into value.

Finally, taking a bit of distance with the concept of "big data", one shall also examine the notion of "data" itself. While we have a very good understanding of the meaning of personal data, there is no real definition of the term "data". This report will therefore rely on the definition provided by the joint technical committee of the International Organization for Standardization (ISO) and the International Electrotechnical Commission (IEC). In 1993 already, the ISO/IEC provided the following definition of data:

"Reinterpretable representation of information in a formalized manner suitable for communication, interpretation, or processing."

It is important to keep this broad-reaching definition in mind when considering the ownership and intellectual property aspects of data management, notably in a big data context. It is also important to remember the fact that data, which may be "personal data", can either be created by (legal or natural) persons or generated by machines or sensors, often as by-products.

---

13 Frank J. Ohlhorst, Big Data Analytics: Turning Big Data into Big Money (John Wiley & Sons 2012) 3
15 "Any information relating to an identified or identifiable natural person" (GDPR, art 4(1))
Chapter 2

Overview of the EU Legal Framework

2.1 Introduction

This second provides an overview of the EU acquis (i.e., the body of EU common rights and obligations) in relation to data management and in particular for access to or exchange of data. More particularly, it aims to highlight the current EU legal regime related to data, such as ownership, access and exploitation rights and obligations.

In such specific context, BIRD & BIRD has identified various categories of EU legislation related to data and potentially playing a role on the possible "ownership" of data, but also on the restrictions and requirements that may apply in relation to certain types of behaviours or data.

The diagram below provides a snapshot of the current EU legal acquis related to data. The following Sections detail the findings and draw some preliminary conclusions.
Ownership of Data

Figure 2.1: Overview of the EU legal framework related to data

- Copyright
- Database rights
- Trade secrets
- EU Data Protection Regulation (GDPR)
- Privacy (GDPR)
- Right to Access to Environmental Information
- Water Framework Directive 2000/60/EC
- Inland Waterways Safety Directive 2002/57/EC
- Maritime Transport Systems (ITS)
2.2 Ownership-like Rights

2.2.1 Concept of Ownership

There is often some kind of misunderstanding between legal practitioners and non-legal professionals on the meaning of the term "ownership".

Following the Oxford Dictionary of Law, the word "ownership" has the following meaning: "it is the exclusive right to use, possess, and dispose of property, subject only to the rights of persons having a superior interest and to any restrictions on the owner's rights imposed by agreement with or by act of third parties, or by operation of law."\(^\text{17}\) It is therefore something that implies certain rights over a property such as being able to enjoy, use, sell, rent, give away, or even destroy an item of property. Ownership may be corporeal (i.e., title to a tangible/material (im)movable object) or incorporeal (i.e., title to an intangible object, such as intellectual property, or a right to recover debt).

However, for businesses, the meaning of "ownership" may be different, especially in a data environment. It is often used to assign responsibility and accountability for specific databases, whereby reference to the "data owner" is made.\(^\text{18}\) In such particular context, the 'ownership' does not have a legal connotation but refers to other concepts such as assurance of data quality and security. There is thus no transfer of or licence over a property as such.

Facing such different meanings, an author suggested in 1998 already to use the term "data stewardship" as it would be more appropriate\(^\text{19}\), capturing the "responsibility that organisations are actually looking to promote with the ownership concept."\(^\text{20}\)

In this report, the term "ownership" will be used in its legal meaning. This nevertheless includes certain difficulties due to the particularities of data. Indeed, data is not like any other tangible or intangible "thing". It

\(^{17}\) Jonathan Law and Elizabeth A. Martin, A Dictionary of Law (7\textsuperscript{th} edition, Oxford University Press 2014)
\(^{18}\) OECD, Data-driven Innovation: Big Data for Growth and Well-being (OECD Publishing 2015) 195
\(^{20}\) OECD, Data-driven Innovation: Big Data for Growth and Well-being (OECD Publishing 2015) 195
has certain characteristics often put forth when discussing the data economy, such as the fact that data is limitless and non-rivalrous, that fit uneasily with the legal concept of "ownership".

### 2.2.2 Actors in the Data Value Chain who Could Claim Ownership in Data

The issue of data ownership is even more complicated by the data value cycle which can be rather complex and involves numerous stakeholders. This increases the difficulties in determining who could or would be entitled to claim ownership in data. Many of such stakeholders may attempt claiming ownership in data because, for instance, they create or generate data, or because they use, compile, select, structure, re-format, enrich, analyse purchase of, take a licence on, or add value to the data. Accordingly, in many instances, different stakeholders will have different powers depending on their specific role. Hence, no single data stakeholder will have exclusive rights.\(^{21}\)

The following Figure aims to depict the data value cycle.\(^{22}\)

[Image: Data value cycle diagram]

Looking at the data value cycle, one can distinguish various actors and determine their roles in the data economy, in particular in the "datafication" process, the analysis of data and the decision-making phase. It should however be kept in mind that certain organisations may

---


\(^{22}\) Ibid 33
play multiple roles. Also, the data value cycle does not reflect the cross-border flow of data and the legal intricacies related thereto.\textsuperscript{23}

There is a multitude of actors on the market actively reaping the benefits of the data economy. The relationships between such actors are an essential element of the data value cycle. Some of the most important actors and their central role are summarised below.

**Internet Service Providers**

Internet Service Providers ("ISPs") are at the heart of the data ecosystem through which data is exchanged.\textsuperscript{24} They play an important role at the beginning of the process, as they provide the necessary technical foundations to end-users (organisations or individuals), or to other ISPs. Certain ISPs also provide supplementary IT services, such as cloud computing and data analytics services. Consequently, ISP's play a fundamental part in big data analytics, including – for some of them – by offering specific data-related services and/or availing of such services for their own needs.

**IT Infrastructure Providers**

IT infrastructure providers make available to other companies the toolkit, including both software and hardware, to handle and analyse big data. They offer tools for data analytics, data management, critical computing, data storage and transport, cloud computing, software allowing database management and analytics, etc.\textsuperscript{25} A typical example is Hadoop, which has almost become a 'standard' technology allowing to deal with complex unstructured large volumes of data.

**Data Providers**

Various kinds of data (service) providers are active in the data environment.

- *Data brokers and marketplaces*

Data brokers and marketplaces compile and aggregate information (including personal data) obtained from a broad range of sources with the ultimate objective to sell, license or otherwise distribute such data to

\textsuperscript{23} See in this respect the recent Free Flow of Data Initiative of the EU Commission as part of the Digital Single Market
\textsuperscript{24} OECD, *Data-driven Innovation: Big Data for Growth and Well-being* (OECD Publishing 2015) 72
\textsuperscript{25} Ibid
companies, consumers or other data brokers. Possible data sources include:

- Data disclosed or provided by organisations or individuals;
- Data from sensors;
- Data mined or crawled on the Internet;
- Data obtained from not-for-profit organisations;
- Open data (see sub-Section 2.5.2 below);
- etc.

- **Individuals (such as data subjects, consumers, patients, etc.)**

Certain individuals play an active role in the data economy either by providing their data (be it personal or not) to organisations (including data brokers), or by assembling, storing and managing their own (personal) data; including in the cloud.

- **Public sector**

Public authorities have been active for several years in making certain sets of data 'freely' available – a concept which is also known as "open data" (see also sub-Section 2.5.2 below). In the EU, for example, the EU institutions adopted a Directive on the re-use of public sector information (government-held data), which aims at unlocking the potential of big data held and accumulated by government authorities.27

### Data Analytics Service Providers

The analysis of data is oftentimes performed by ISPs, IT infrastructure providers or data providers. Nevertheless, the data ecosystem still includes specific providers of data analytics services, including for the development of dedicated software and visualisation tools based on data analytics.28 The role of data analytics service providers tends to be assumed by start-ups or SMEs specifically active in the development of new techniques, such as predictive analytics, simulations, scenario development and advanced data visualisations.29

Peculiar, however, is the fact that data analytics service providers, contrary to data brokers, generally obtain their data directly from their customers, rather than from third party sources. This naturally has consequences for the identification of actors as data controllers or processors in a data protection context. Taking into account that particularity of their service, data analytics service providers usually

---

26 Ibid 82
29 Ibid
Ownership of Data

qualify as "data processors", rather than data controllers. Data brokers, on the contrary, are generally considered to be independent data controllers.

Data-driven Entrepreneurs

The last category of actors we will discuss covers those organisations developing cutting-edge products, services and technologies based on the use of data and data analytics for different purposes; the so-called data-driven entrepreneurs. These include start-ups and incumbents, but also innovative (ICT and non-ICT) companies and governments. Not only do they use data as the core enabler for their business operations; for a majority of them it can even be said to be the fundamental economic value behind the service they provide. Against such background, data becomes a valuable asset due to the transformation of data into know-how and intelligence, and thus it can be used for decision-making purposes.

A Layered Approach of the Key Roles of Actors

The actors as well as their roles briefly explained above can be depicted in layers, whereby the underlying layers supply the upper layers with goods and services:

Figure 2.3: The data ecosystem as layers of (key roles of) actors

---

2.2.3 Lack of "Ownership" Right in Data

Our researches have not enabled us to identify any EU legislation that would specifically regulate the question of ownership in data. This being said, such absence of ownership-related legislation does not exclude the fact that there are numerous legislations that have an impact on data (see Figure 2.1 above) or that may confer some kind of protection to certain types of data or on datasets (i.e., copyright, database rights and trade secrets).

According to some authors, the CJEU opened the door for a discussion on ownership in intangible assets in its *UsedSoft* judgment issued on 3 July 2012 (case C-128/11). In this ruling, the Court held that the commercial distribution of software via a download on the Internet is not only based on a licence, but on a sale of goods. Therefore, the owner of copyright in software cannot prevent a perpetual "licensee" from selling his software (understood as downloaded file). The decision implies that there is a specific ownership attributed to intangible goods like software downloaded via the Internet. Applicability of this model to other digital goods remains to be considered in future court decisions. Despite such ruling and the possible interpretation deriving from it, a high legal uncertainty remains.

To address these problems of legal uncertainty, the Commission announced an initiative for the free flow of data. According to the Commission Communication "Digitising European Industry: Reaping the full benefits of a Digital Single Market" 32, the Commission is due to propose in 2016 an initiative, i.e., allowing to examine in greater detail the emerging issues of data ownership, access and re-use rules, including as regards data in an industrial context and especially data generated by sensors and other collecting devices. However, according to the recent press reports 33, the EU efforts to regulate "data ownership" have been delayed to give officials more time to grapple with the legal issues surrounding who controls and can access (big) data.

The same issues apply when looking at the situation at national level. In such context, we examined the situation in various Member States (i.e.,

31 Thomas Hoeren, 'Big Data and the Ownership in Data: Recent Developments in Europe' (2014) 36(12) EIPR 751
Belgium, France, Germany, Italy, Spain and the United Kingdom). Excluding the national laws pertaining to intellectual property and trade secrets, the results of our cross-jurisdictional study are detailed in the following sub-sections.

Is there any Particular National Legislation relating to Data Ownership?

<table>
<thead>
<tr>
<th></th>
<th>Belgium</th>
<th>France</th>
<th>German</th>
<th>Italy</th>
<th>Spain</th>
<th>UK</th>
</tr>
</thead>
<tbody>
<tr>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
</tbody>
</table>

There clearly is no specific data-related legislation that explicitly recognises ownership in data in the various Member States examined in this report.

Having said that, some countries have in place legislation allowing to control the flow of data. One example would be France, where the civil code sets out mechanisms (based on both civil and criminal law measures) enabling the holder of data to prevent or restrain the misuse of data.

Is there Case-law in your Jurisdiction Addressing the Issues of "Ownership" of Data?

<table>
<thead>
<tr>
<th></th>
<th>Belgium</th>
<th>France</th>
<th>German</th>
<th>Italy</th>
<th>Spain</th>
<th>UK</th>
</tr>
</thead>
<tbody>
<tr>
<td>No</td>
<td>Yes (limited)</td>
<td>Yes (limited)</td>
<td>Yes (limited)</td>
<td>No</td>
<td>Yes (limited)</td>
<td></td>
</tr>
</tbody>
</table>

Certain issues that may be of particular interest when considering the "ownership of data" have been addressed by the German Courts.

The first landmark decision comes from the Higher Regional Court of Karlsruhe and concerns destruction of data. The Court considered that deletion of data stored on a data carrier may violate the ownership in the

---

34 OLG Karlsruhe, Urteil v. 07.11.1995 – 3 U 15/95 – Haftung für Zerstörung von Computerdaten
Ownership of Data

data carrier under the German Civil Law Code, extending the protection of the ownership in the data carrier to data stored on it. On the other hand, later decisions of German courts opposed the possibility to hold ownership over data as such, since data lacks the necessary material character\(^{35}\) and since it is not considered a ‘thing’ under the German Civil Law Code.\(^{36}\)

More recently, the Court of Appeal of Nuremberg\(^ {37}\) has built on the general principle adopted in Germany, according to which things that are neither rights nor goods may nevertheless be sold within a sale contract (Section 453 of the German Civil Act). To decide whether former employees were allowed to delete the data stored on their company-owned laptops, the Nuremberg Court made reference to the theory of the so-called "Skripturakt". According to this theory, the person who generates the data gets the right to the data, even if the data afterwards are used for the business or for the sake of the employer. In consequence, under criminal law, the employees were allowed to delete the data.\(^ {38}\)

The Nuremberg Court has however indicated that whilst the rule derived from the "Skripturakt"-theory also applies in the employment context, the situation may be different if the data have been already passed over to the employer; in such case the employer would become the owner of data. In addition, in the Court’s opinion, the data will originally belong to the employer if they were created completely according to his demands.\(^ {39}\)

One should bear in mind that the above case had a strong criminal law connotation; the employees who deleted the data without prior authorisation were accused of theft, with their employer asking for a conviction under Section 303(a) of the German Criminal Act (prohibiting unlawfully erasing, corrupting or altering computer data under penalty of imprisonment). It is unclear whether the same rule would be applied by German courts in a civil law matter.

The Labour Court of Appeal of Saxony (Landesarbeitsgericht) had to decide a similar case in 2007, from a civil law perspective.\(^ {40}\) This decision is however somewhat contradictory to the one issued by the Nuremberg Court later on. The Saxon Court claimed that because the employee installed software (Microsoft Outlook) on a company-owned laptop, the employer has obtained the property in the software.

\(^ {35}\) LG Konstanz, Urt. v. 10.05.1996 – 1 S 292/95 = NJW 1996,2662
\(^ {36}\) OLG Dresden, Besch. v. 05.09.2012 – 4 W 961/12 = ZD 2013,232
\(^ {37}\) OLG Nürnberg 1. Strafsenat decision of 23.01.2013, 1 Ws 445/12
\(^ {38}\) OLG Nürnberg 1. Strafsenat decision of 23.01.2013, 1 Ws 445/12, par. 14
\(^ {39}\) OLG Nürnberg 1. Strafsenat decision of 23.01.2013, 1 Ws 445/12, par. 16-17
\(^ {40}\) LAG Sachsen, decision of 17.01.2007, 2 Sa 808/05, MMR 2008, 416
In consequence, when the employee deleted the software from this laptop he destroyed the data of the employer, and, in consequence, could be dismissed.

While the question of "ownership" of data was also indirectly addressed by the courts in the United Kingdom, they did not set out clear rules on that matter. So far, the UK courts held that data is not property and therefore cannot be stolen, that data are not eligible to be the subject of a common law lien, and that there is no proprietary right in the content of an email.

Finally, the French Supreme Court ("Cour de cassation") has recently rendered a ruling that could open a way to recognising the ownership of "data". The Court found that downloading (remotely) computer data without taking away their support may amount to the offence of theft, acknowledging therefore indirectly that such independent data may be owned.

Are Scholars Debating the Issues related to Data Ownership?

<table>
<thead>
<tr>
<th>Belgium</th>
<th>France</th>
<th>German</th>
<th>Italy</th>
<th>Spain</th>
<th>UK</th>
</tr>
</thead>
<tbody>
<tr>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
</tr>
</tbody>
</table>

According to our research hitherto, it seems that although the problem of data "ownership" is recognised by numerous authors, only very few of them provide a deeper analysis of the related issues.

The current lack of clarity as to the status of data under English law was addressed for instance by Christopher Rees, who believes that data could be classified as property (based on a simple definition of property as the right to use something and exclude others from its use).

Most of the German academics argue that German law does not know a right for creating such right. There are however voices opposing this line

---

41 Oxford v Moss [1979] 68 Cr App Rep 183
42 Your Response v Datateam Business Media [2014] EWCA Civ 281
43 Fairstar Heavy Industries v Adkin, [2013] EWCA Civ 886
44 May 20, 2015 (No14-81336)
45 Christopher Rees, 'Who Owns our Data?' (2014) 30(1) Computer Law & Security Review 75
of thought, in view of the recent jurisprudence of the German Courts. In particular, Prof. Dr. T. Hoeren examined the issues of data ownership under the current German legal framework and jurisprudence\textsuperscript{47}, concluding that "in general, the property in data is attributed to the originator, creator, or producer of these data. However, in the case of data made for hire (to use the US copyright term), the data belong to the employer".

Other scholars seem to suggest that one may rely on the current wording of Section 950 of the German Civil Code to claim some kind of property right in data. Such Section stipulates that "A person who, by processing or transformation of one or more substances, creates a new movable thing acquires the ownership of the new thing, except where the value of the processing or the transformation is substantially less than the value of the substance. Processing also includes writing, drawing, painting, printing, engraving or a similar processing of the surface." Despite the legal uncertainty surrounding such theory, and notably its particular application to intangible assets such as data, certain undertakings have already relied on it in their general terms and conditions.

Having said that, the majority of German academics seems to agree that no right in data exists.\textsuperscript{48}

Commentators seem to be divided as to the ownership of data under French law. While some commentators indicate that data are not appropriable as such\textsuperscript{49}, others believe that in view of the abovementioned ruling of the French Supreme Court the ownership over data cannot be doubted.\textsuperscript{50} Having said that, most of the discussions on the recognition of ownership seem to focus on the individuals' ownership over their personal data\textsuperscript{51, 52}

\textsuperscript{47} Thomas Hoeren, 'Big Data and the Ownership in Data: Recent Developments in Europe' (2014) 36(12) EIPR 751
\textsuperscript{48} See also p.62 below for further details on the situation in Germany
\textsuperscript{50} Pierre Berlioz, 'Consécration du vol de données informatiques. Peut-on encore douter de la propriété de l’information?' (2015) 4 Revue des contrats 951
\textsuperscript{52} See also in that respect p.45 on the recent legislative changes in France
2.2.4 Intellectual Property Rights and Trade Secrets

Our study of the current EU legal framework has led us to confirm that the ownership-like rights currently available are limited to intellectual property rights and trade secrets (see Chapter 3 of this report for a more in-depth analysis of such legal regimes).

![Figure 2.4: Overview of ownership-like rights in data](image)

**Intellectual Property Rights**

"Intellectual property rights" refers to a broad notion that confers an exclusive right to a person. It allows the holder to exercise a monopoly on the use of the item for a specified period. By restricting certain acts, a monopoly power is therefore conferred, but the social costs of the monopoly power may be offset by the social benefits of higher levels of creative activity encouraged by the monopoly earnings.\(^53\)

More specifically, intellectual property refers to the legal rights which result from intellectual activity in various fields – i.e., industrial, scientific, literary and artistic.

Such exclusive rights, recognised in international treaties and harmonised to a certain extent at EU level, aim at safeguarding creators and other producers of intellectual goods and services by granting them certain time-limited rights to control the use made of those productions. It shall be borne in mind that those rights do not apply to the physical object in which the creation may be embodied but instead to the intellectual creation as such.\(^54\)

In the context of data, it can therefore not be excluded that intellectual property rights, and in particular copyright and database rights, provide some kind of protection to data and/or datasets.


Given the importance of ownership rights in the context of intellectual property, Chapter 3 of this report examines in detail the protection granted by the EU copyright and database legal frameworks.

**Trade Secrets**

In addition to intellectual property rights (see previous Section and Chapter 3), the protection granted to "trade secrets" may also be of particular relevance in a data context.

At EU level, a specific legal ground for the protection of trade secrets now exists, namely the new EU Trade Secrets Directive, which entered into force on 5 July 2016.\(^{55}\) The recognition of trade secrets and the associated legal protection also exist at international level where Article 39 of the Agreement on Trade-Related Aspects of Intellectual Property Rights ("**TRIPs**") refers to the protection of "undisclosed information":

"Natural and legal persons shall have the possibility of preventing information lawfully within their control from being disclosed to, acquired by, or used by others without their consent in a manner contrary to honest commercial practices so long as such information:

(a) 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; and

(c) has been subject to reasonable steps under the circumstances, by the person lawfully in control of the information, to keep it secret."

Despite such protection, numerous authors highlight the difficulties related to the protection of trade secrets:

"Despite their enormous importance, and our intuitive sense of what they are, trade secrets are the most enigmatic and least consistent form of intellectual property. Indeed, there are those purists who would deny that trade secrets are a form of property at all, let alone intellectual property, and such a controversy,

---

\(^{55}\) Directive (EU) 2016/943 of the European Parliament and of the Council on the protection of undisclosed know-how and business information (trade secrets) against their unlawful acquisition, use and disclosure [2016] OJ L 157/1
relatively unimportant though it may be when it comes to the real life practicalities of protecting trade secrets, is reflected in the widely different legal bases under which they are protected in different countries.\textsuperscript{56}

In a world of data, trade secrets are often invoked as the means to claim protection. Commercial but also technical data could benefit from the trade secrets protection in the EU if it fulfils the various criteria. Given the importance of such legal ground, Chapter 3 of this report examines in detail the protection granted by the trade secrets legal framework in the EU.

2.3 Individuals' Rights

Some rights are awarded to individuals with respect to certain types of data. Accordingly, natural persons may be in a position to restrict what can be done with data collected and processed by public or private organisations.

2.3.1 Privacy and Data Protection

The protection of individuals' privacy has triggered a lot of attention in the past months in the EU. Indeed, the most significant legislative framework in this respect was adopted in May 2016, i.e., the General Data Protection Regulation No 2016/679, and will be applicable as from 25 May 2018.

Accordingly, the issues related to privacy and data protection are undeniably at the centre of the current debates related to data across the EU.

Is Privacy and Data Protection the Main Issue relating to Data Discussed by Scholars in your Jurisdiction?

<table>
<thead>
<tr>
<th>Belgium</th>
<th>France</th>
<th>German</th>
<th>Italy</th>
<th>Spain</th>
<th>UK</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
</tbody>
</table>

Certain scholars in various Member States raise particular issues in relation to personal data. For instance, in the **United Kingdom**, several authors highlight the fact that there is an increasing erosion of the line between "personal" and other types of data. This is explained by the fact that data is collected into datasets which are then combined. Also, the improvements in (big) data analytics allow more easily linking data to individuals in such ways that could not have been imagined before. Accordingly, "non-personal data" can more easily become now or in the (near) future "personal data". It goes without saying that the development and improvement of the Internet of Things or smart cities will accentuate such finding.57

In Germany, the debates surrounding personal data also relate to the issues of ownership. Some scholars argue indeed that personal data should belong to the individual (i.e., the "data subject").58 However, most commentators agree that data subjects have no general rights in their data.59

In this context, it must be reminded that the rights and obligations under data protection law emanate from the fundamental right to privacy. Accordingly, such rights relate to a personality right recognised to individuals. The fact that the GDPR recognises important rights to data subjects does not as such regulate the question of data ownership and therefore does not recognise a "property" right of individuals in their data. In our view, the GDPR only regulates the relationship between the data subject and the data controller(s)/processor(s), without creating and regulating the issues of commercially exploitable rights in personal data. This being said, any "ownership" right that would be awarded in

---

57 Henry Pearce, 'Systems Approach to Data Protection Law and Policy in a World of Big Data?' (2016) 22 CTLR 90
59 E.g., Michael Dorner, Big Data und "Dateneigentum" (2014) 9 CR 617, 619 ff
personal data to persons other than data subjects, would be restricted by the application of the GDPR.\textsuperscript{60}

Some scholars highlight however the fact that the GDPR "recognises different levels of control rights to consumers in accordance with a 'proprietarian' approach to personal data."\textsuperscript{61} More specifically, some have emphasised the fact that in practice personal data is treated as property of the individual which the data relates to.\textsuperscript{62}

France
Recently, the French legislator has adopted the Digital Republic Act, which entered into force on 9 October 2016. The new framework aims to incorporate support for innovation and new business models, more open data, better personal protection, heightened platform fairness and extended rollout of digital access.\textsuperscript{63} The text notably imposes new requirements on businesses, in anticipation of the GDPR. In such context, it is particularly interesting to note that the Digital Republic Act incorporates a principle into the French Data Protection Act (Article 1 of Act No. 78-17 of 6 January 1978), according to which "every person shall have the right to decide and control the uses made of his or her personal data." Although its terms have been criticised, this provision anchors the protection of personal data into the sphere of fundamental rights. Some commentators argue that in practice, such provision could be used as an additional legal ground to invalidate terms and conditions vesting in service providers a property right over individuals' personal data.\textsuperscript{64}

The present specific report takes an approach whereby ownership is examined in relation to data, be it non-personal or personal data, which can be as a matter of fact owned by data controllers or processors but

\textsuperscript{60} See in the same vein, in the context of disclosure of chemical data, the Court Order of the EU General Court in case T-189/14 wherein the President examines in obiter dictum the question of privacy (Case T-189/14 R Deza a.s v Agence européenne des produits chimiques [2014] ECLI:EU:T:2014:686). More particularly, the President acknowledges the relevance of the question of privacy of legal entities but nevertheless reminds, on the basis of the decision of the Court of Justice of the EU in case C-450/06, that it may be necessary to prohibit the disclosure of information qualified as confidential in order to preserve the fundamental right to privacy of an undertaking (Case C-450/06 Varec SA v État belge [2008] ECLI:EU:C:2008:91).

\textsuperscript{61} Gianclaudio Malgieri, 'Property and (Intellectual) Ownership of Consumers' Information: A New Taxonomy for Personal Data' (2016) 4 PinG 133; Jacob M. Victor, 'The EU General Data Protection Regulation: Toward a Property Regime for Protecting Data Privacy' (2013) 123(2) Yale Law Journal 266

\textsuperscript{62} Nadezhda Purtova, 'The Illusion of Personal Data as No One's Property' (2015) 7(1) Law, Innovation and Technology 83


\textsuperscript{64} Willy Mikalef, 'New Data Protection Provisions in France: Are You Ready?' (2016) 16(45) WDPR 1
where data subjects maintain a control over their personal data within the limits of the GDPR.

2.3.2 Consumer Protection

In addition to privacy, individuals in their "consumer role" benefit from a specific legal framework aimed at protecting them. Certain consumer protection requirements relate to data and may affect the control an organisation may have on such data.

Although there is no uniform definition of "consumer", the EU legislation generally labels "natural persons acting outside of their trade or profession" as such.\textsuperscript{65} Simply put, in order for EU consumer protection legislation to apply, the party asserting the protection must be a consumer.

In the EU, an inclusive and multifaceted system of consumer protection guarantees European consumers a high level of rights. On that basis, consumers ought to be able to extend this expectation – that their rights are protected – to purchases made in the digital marketplace on a national and even a global level. Application of EU rules covers the entire "lifecycle" of a consumer contract, from the advertisement of a service to the contract offer and formation of a contract, through procedural and substantive issues regarding the content of terms, and finally setting the rules governing how and where disputes will be adjudicated if the need arises.\textsuperscript{66}

This system of protection is expressed visually in the following chart. Protections (as shown below) start with pre-contractual information at the bargaining/advertising/pre-contractual phase, continue protection throughout the contracting phase, and even provide post-contractual remedies for terms that are substantively or procedurally unfair, among other possible remedies.

\textsuperscript{65} See e.g. Article 2(1) of Directive on Consumer Rights defining "consumer" as "...any natural person who, in contracts covered by this Directive, is acting for purposes which are outside his trade, business, craft or profession" (Directive 2011/83/EU of the European Parliament and of the Council on consumer rights [2011] OJ L 304/64)

\textsuperscript{66} For further information on formation of cloud computing contracts, see D2.2 § 6.2, at 143-47
Pre-contractual protections

- Unfair Terms Directive
- Consumer Rights Directive
- Consumer Sales and Guarantees Directive
- e-Commerce Directive
- Unfair Commercial Practices Directive

Post-contractual protections

- Unfair Terms Directive
- Rome I Regulation
- Brussels I Regulation
- ADR/ODR Directive

Protection throughout the contracting lifecycle

Table 2.1: Overview of the consumer protection EU legal framework

In this context, the EU Commission published in December 2015 a Proposal for a directive on certain aspects concerning contracts for the supply of digital content (hereafter the "Digital Content Proposal"). It contains new rules in relation to the contractual aspects of the relationship between consumers and suppliers of digital content.

The Digital Content Proposal notably highlights that "information about individuals is often and increasingly seen by market participants as having a value comparable to money." Accordingly, the scope of such (proposed) Directive would also apply when the consumer actively provides counter-performance other than money in the form of personal information.

---

76 Ibid, Recital 13
data or any other data.\textsuperscript{77} It therefore provides for the commodification of (personal) data, following certain scholars who have considered the ‘propertisation’ of personal data and thus the creation of a property right \textit{(in rem)} in relation to data concerning an individual.\textsuperscript{78} Such reasoning would be particularly relevant in a big data context as it would allow a more widespread analysis of personal data in the framework of big data analytics.

The Digital Content Proposal however includes certain exceptions. It would apply only to contracts where the supplier requests and the consumer actively provides data. It should not apply to situations where the supplier collects data necessary for the digital content to function in conformity with the contract. Also, it should not apply to situations where the supplier collects information, including personal data, such as the IP address, or other automatically generated information such as information collected and transmitted by a cookie, without the consumer actively supplying it, even if the consumer accepts the cookie. It should also not apply to situations where the consumer is exposed to advertisements exclusively in order to gain access to digital content. In the same vein, Recital 17 of the proposed Directive discusses the issues related to the Internet of Things, affirming that it is opportune to address specific issues of liability related to the Internet of Things, including the liability for data and machine-to-machine contracts, in a separate way.

Finally, in the event of termination of the contract, the Digital Content Proposal takes over from the Consumer Sales and Guarantees Directive\textsuperscript{79} the idea of hierarchy of remedies. In such context, it includes specific rules on the consequences of termination, especially in the event of contracts concluded for consideration (“counter-performance”).\textsuperscript{80} It notably aims to regulate the further use of the consumer’s personal data and the supplier’s duty to stop using the consumer’s data and to take all measures in order to comply with data protection rules by deleting it or

\footnotesize{
}
rendering it anonymous\textsuperscript{81}, in light notably of the rights of data subjects under the GDPR.\textsuperscript{82}

2.4 Competition Law Rights and Obligations

Most competition law dates from an era when companies had tangible assets, and if they were not making nuts and bolts, they were providing a service with a more or less well defined purpose, value, and customer.

But 'data' has now emerged as a new commodity. In large quantities, data can define company value. It can be bought and sold. It can be used to leverage new product or service offerings.

That has put antitrust regulators in a difficult position, as they apply the legal principles of the nuts-and-bolts world to a reality in which anything from a person's date of birth to their holiday photos and Internet browsing habits are the new corporate assets.

And that is just personal data. Companies active in what has become known as big data – blocks of mass information that can be processed – are trading in so much more than one's online shopping details. Weather, pollution levels, traffic flow, energy use, prices, values, water levels, crop yields, exchange rates, trade flows and countless more provide data points that can be collated, compounded, correlated (or not), and marketed on.

That means that any exercise intended to explore how competition law applies to data must start where – appropriately – most antitrust cases start: product and market definition. What do we understand by the concept of 'data' and what are its main characteristics?

This Section will analyse the impact of big data on different aspects of EU competition law and will seek to create more clarity on when and how the ownership or use of (big) data can give rise to competition law concerns.

\textsuperscript{82} Ibid 30
2.4.1 Characteristics and Key Competition Law Aspects of Big Data

As already mentioned above, big data can be defined as a large amount of data sets (which may be raw or have been processed) which can be used for different purposes. Data can come from a large number of sources and can relate to very different types of information. However, many data share the following characteristics:

- **Data is not expendable**: the fact that one company has easy access to the data does not necessarily mean that other companies do not have the same or similar access;
- **Data is not rivalrous**: it can be obtained and used by multiple companies, sometimes even simultaneously; and
- **Data has a short shelf life**: data may quickly become outdated and therefore may lose its value within a short period of time.

These characteristics are of particular interest in the context of competition law, as they have an impact on

- whether a company with access to large volumes of data may or may not be considered dominant;
- what types of agreements between companies relating to the access to or the use of data could be viewed as anti-competitive; and
- whether an M&A transaction involving data-rich companies will be considered as problematic.

The European Commission as well as several national competition authorities have been showing a growing interest in big data. Although there are many ways in which competition law can affect the way in which big data is collected, used and shared, the primary areas of interest are currently the role of big data in merger control as well as the question whether the ownership of (or privileged access to) big data gives a company a dominant position.
In May 2016 the German and the French competition authority published a jointly developed report on "Competition Law and Data", which analyses the implications and challenges for competition authorities resulting from data collection in the digital economy and other industries.  

The German competition authority also highlighted the importance of data in a competition law context in its working paper "Market Power of Platforms and Networks" of June 2016. The working paper discusses in particular:

- What is the importance of data for Internet services?
- How can market power be calculated in free data-driven markets?
- When can data be considered an "essential facility"? and
- To what extent should data protection rules be allowed to affect competition law analysis?

In addition, on 23 May 2016, the French competition authority launched a sector inquiry into the relevance of data in the online advertising sector. The issues investigated by the authority include:

- Data and relevant markets: examining the degree of substitutability between different forms of targeted advertising, with a particular focus on advertising on social networking sites;
- Data and market power: analysing the weight and strategies of large players like Google and Facebook as well as media agencies and certain data providers; identifying competitive advantages and possible dominant positions, as well as the role of data in the exercise of market power; and
- Data and business practices: assessing to what extent certain business practices may restrict the development of merit-based competition, with a focus on conditions for access to data and the commercial supply of data, as well as access to the advertising ecosystem infrastructures for data providers.

---

83 Bundeskartellamt/Autorité de la concurrence, 'Competition Law and Data' (10 May 2016) [https://www.bundeskartellamt.de/SharedDocs/Publikation/DE/Berichte/Big%20Data\%20Papier.pdf?\_blob=publicationFile&v=2] accessed 15 November 2016

84 Bundeskartellamt, 'Arbeitspapier Marktmacht von Plattformen und Netzwerken' (Bundeskartellamt 2016) 91, an English summary is available at [http://www.bundeskartellamt.de/SharedDocs/Publikation/DE/Berichte/Think-Tank-Bericht-Kurzzusammenfassung_Englisch.pdf?\_blob=publicationFile&v=2]

Similarly, the Dutch competition authority has launched a market study\textsuperscript{86} into online platforms that is looking into:

- the terms and conditions, and businesses models of these platforms;
- the advertisement options, and revenues from advertisements;
- the collection and use of personal data for personalised advertisements; and
- the distribution of online videos.

2.4.2 Dominance and Essential Facilities

When a company holds a dominant position on a particular market, and uses this position to distort competition by, for example, trying to eliminate competitors, or by creating a barrier to entry for competitors, this will be considered an abuse of dominance, which is prohibited by competition law. To be dominant is not illegal; only the way in which a company uses its dominant position may constitute a breach of competition law.

To determine whether a company is dominant, one must first define the relevant product market and the relevant geographic market in which the company is active.

The simple fact that a company has access to large amount of data does not automatically provide it with a dominant market position. Important factors that need to be taken into account to determine the existence of dominance include:

- Do other competitors have access to the same data?
- Is there data which can substitute the data collected by the company?
- Does the company have the ability to analyse and monetise the collected data?
- Is the data held by the company raw data or fully analysed data?

The trend in current analysis seems to focus primarily on the amount of data, with limited attention being given to the aspects listed above. These aspects may lead to the conclusion that, in a given case, even access to a very large amount of data does not provide a company with market power.

The main criteria allowing to determine whether access to certain data gives market power include:

Ownership of Data

- **Quantity**: Once a certain volume of data has been gathered, the collection of additional data will not necessarily lead to any significant additional findings or benefits for the collecting company (so-called *diminishing returns* theory). The level above which the returns decrease will obviously differ between companies and industry sectors;
- **Quality**: Not all collected data has the same value. Raw data which cannot be processed and thus cannot be immediately monetised has a lower value than data which is ready for use and monetisation;
- **Availability**: As mentioned above certain data is readily available to multiple companies since consumers typically use their personal data in different manners for different purposes. (*multi-homing*).

The dating app Tinder and the photo/video social media app Snapchat are telling examples of how a company can enter an app-rich market without any data and, within a short period of time, become the leader in its field owning a very large amount of consumer data.

Furthermore, the existence of *data brokers* means that even if a company has difficulties collecting large amounts of data itself, it may be able to obtain data from a third-party provider. In this scenario, the cost to obtain the data may not create a considerable hurdle.

All of the above is also relevant for the question whether certain data can be categorised as being an *essential facility* and thus must be made available to competitors on non-discriminatory, fair and reasonable conditions. A facility or infrastructure is deemed essential when its use is indispensable for an activity in a market upstream or downstream of the market where the company holding the data is active. It is clear from the above that these requirements are likely to be satisfied only in very limited circumstances.

The German legislator has decided not to wait for the finalisation of the different market studies and public consultations regarding the impact of data on a company's market position. The 9th Amendment to the German Competition Act will therefore add "access to competitively relevant data" as well as network effects and multi-homing to the criteria for determining whether a company holds a dominant position.\(^{87}\)

---

\(^{87}\) See the proposed amendment of section 18 of the German Competition Act (inclusion of a new paragraph 3a) according to the reform bill ("Regierungsentwurf eines neunten Gesetzes zur Änderung des Gesetzes gegen Wettbewerbsbeschränkungen", available at <http://dip21.bundestag.de/dip21/btd/18/102/1810207.pdf>.)
2.4.3 Mergers and Acquisitions

There has been a marked increase of acquisitions in data-related sectors. This is not surprising given that a growing number of established market players are seeking to strengthen their position by acquiring either data-rich companies, or investing in start-ups collecting large volumes of data.

Whether or not the acquisition of a data-rich company raises competition law concerns will vary from case to case. The fact that one of the parties owns a large amount of raw data does not automatically give rise to concerns.

In the Google/Doubleclick\(^{88}\) and Facebook/Whatsapp\(^{89}\) transactions, neither the EU Commission nor the US Federal Trade Commission concluded that the transaction would give rise to competition law concerns. Although many believed that this would be different in the more recent Microsoft/LinkedIn\(^{90}\) case, Microsoft ultimately was able to avoid an in-depth investigation by the European Commission by offering commitments. The commitments (which are described in more detail below) did not relate to data as such and are rather reminiscent of the 2009 Microsoft Internet Explorer decision.\(^{91}\)

The Microsoft/LinkedIn transaction was also cleared by the US competition authority as well as multiple national competition authorities\(^{92}\). The EU Commission did not approve the transaction as quickly as the other authorities and extended the review period to discuss the proposed commitments.

According to Commissioner Margarethe Vestager\(^{93}\), the Commission’s analysis of the Microsoft/LinkedIn transaction focused on whether the data acquired by Microsoft "has a very long durability or might constitute a barrier for others". Salesforce.com had voiced concerns that the

---

\(^{88}\) Google/Double Click (Case M.4731) Commission Decision 11/03/2008 [2008] OJ C 927 final


\(^{90}\) Microsoft/LinkedIn (Case M.8124) Commission Decision pending

\(^{91}\) Microsoft/Tying (Case AT.39530) Commission Decision 06/03/2013 [2013] OJ C(2013) 1210 final

\(^{92}\) Flavia Fortes, 'Microsoft Given Go-ahead by US Antitrust Enforcers to Buy LinkedIn' (2016) MLex


transaction would have anticompetitive effects on the market. In an official statement\textsuperscript{94} made by the Chief Legal Officer and Chief of Corporate and Government Affairs, Salesforce stated that:

"If Microsoft gains ownership of LinkedIn, the company will have the ability and incentive to use LinkedIn’s one-of-a-kind data to enhance its own products, while preventing competitors from accessing and effectively utilizing that same data. The result will fundamentally change the marketplace in a way that will be harmful to consumers. Even more damaging, Microsoft could end up stifling future innovation in the market at large by further extending to the cloud the same monopolistic position upon which Microsoft has built its traditional franchise."

The EU Commission approved the Microsoft/LinkedIn deal on 6 December 2016 subject to the parties' compliance with a series of commitments aimed at preserving competition between professional social networks in Europe.

The Commitments include:

- ensuring that PC manufacturers and distributors will be free not to install LinkedIn on Windows and allowing users to remove LinkedIn from Windows should PC manufacturers and distributors decide to preinstall it;
- allowing competing professional social network service providers to maintain current levels of interoperability with Microsoft's Office suite of products through the so-called Office add-in program and Office application programming interfaces;
- granting competing professional social network service providers access to "Microsoft Graph", a gateway for software developers. It is used to build applications and services that can, subject to user consent, access data stored in the Microsoft cloud, such as contact information, calendar information, emails, etc. Software developers can use this data to drive subscribers and usage to their professional social networks.

The commitments will apply in the EEA for a period of five years and will be monitored by a trustee.\textsuperscript{95}


\textsuperscript{95} European Commission, 'Press Release. Mergers: Commission Approves Acquisition of LinkedIn by Microsoft, subject to Conditions' (2016)
The relevance of big data in the context of merger control has been discussed only in a handful of cases. This may, however, change in the very near future, following the conclusion, on 13 January 2017, of the EU Commission's public consultation on the functioning of the EU Merger Regulation.\(^{96}\) The consultation explores inter alia whether the purely turnover-based thresholds in the EU Merger Regulation result in an enforcement gap regarding acquisitions of data-rich companies that do not yet generate significant turnover but have a high market potential. One of the solutions considered by the Commission is the introduction of a complementary "deal size threshold" which would be based on the value of the transaction.

The German legislator has already decided to address this issue by complementing the turnover thresholds under the German merger control rules by a new threshold that is based on the agreed purchase price. According to the latest version of the proposed amendments to the German Competition Act, a transaction will also require prior merger control approval in Germany if the target does not meet the local turnover thresholds but if the purchase price is at least EUR 400 million and if the target has significant activities in Germany\(^{97}\).

2.4.4 Conclusion on the Competition Law Aspects

Assessing the market conduct of companies with access to large volumes of data raises complex issues under competition law. The difficulty of the exercise is compounded by the fact that the analysis also needs to take into account data privacy and consumer protection issues that are intimately linked to the questions under competition law.

Both the European Commission and various national competition authorities have already invested significant time and effort into the competition law analysis of big data and there is extensive legal literature on this topic. However, many issues remain unexplored and new issues will arise as a result of the on-going technological development. An effective response to these developments will require close cooperation in particular between the European competition and data protection authorities and the use of thorough economic analysis to avoid an over-enforcement that could stifle innovation and the emergence of new services and business models.


\(^{97}\) See the proposed amendment of section 35 para. 1 of the German Competition Act according to the reform bill ('Regierungsentwurf eines neunten Gesetzes zur Änderung des Gesetzes gegen Wettbewerbsbeschränkungen', available at <http://dip21.bundestag.de/dip21/btd/18/102/1810207.pdf>.)
2.5 Data Sharing Obligations

2.5.1 Introduction

In the context of the examination of the EU legal framework related to data, we identified numerous legislations that may impact a company’s control of, the access to, or the rights in data. Such legislations are sector-focused and provide for an array of rights and obligations in relation to specific types of data in particular circumstances.

The following diagram illustrates our findings, listing – albeit non-exhaustively – examples of EU legal instruments:

![Diagram of sector-specific EU data sharing obligations]

Our research has further led us to analyse the various legislations listed in the above diagram. In addition to the public sector and the particular issues related to open data (see sub-Section 2.5.2 below), it appears that the data sharing obligations may vary depending on various factors;

---

Ownership of Data 43 / 142
e.g., the sector concerned or the reasons of public interest that have led to the adoption of the instrument, such as for instance public security or public health. Furthermore, while most Directives and Regulations concerned aim to reach transparency by imposing data sharing obligations, some include mechanisms to protect and limit the disclosure of certain types of data, such as commercially confidential information.

Although the aim of this report is not to delve into the details and complexities of the above-listed legislative instruments at EU and national level, we examine succinctly below some of the most interesting findings.

**Limited Data Sharing to Authorities**

On the one hand, certain legislations have a narrow scope, requiring the disclosure of limited amounts of data for well-defined purposes. This is for instance the case in the aviation sector where air carriers must communicate information concerning passengers, and thus "personal data" to certain authorities. Such legislation has little to no impact from a commercial perspective, as the data is not made publicly available and competitors thus have no access to the collected and transmitted data.

**Informative Data Sharing to Consumers**

On the other hand, a range of legislations aims to impose data sharing obligations – and thus limiting the control over such data – on well-defined operators in order to provide information to the public.

This is the case in certain legislations relating to the information that must be clearly made available to consumers, free of charge – these include for instance:

- The **Food Information to Consumers Regulation** (No 1169/2011), which concerns the display of product information on product packaging and online stores; pertaining to food and beverages sold in the EU. The objective of this instrument is to standardise food labelling and provide greater clarity to consumers on ingredients, nutrition and allergens.
- The **Car Labelling Directive** (1999/94/EC), which aims to help consumers choose vehicles with low fuel consumption by requiring dealers in new passenger cars to provide potential buyers with useful information on these vehicles' fuel consumption and CO₂ emissions. This information must be displayed on the car's label, on posters and other promotion material, and in specific guides.
• The **Energy Labelling Directive** (2010/30/EU), which aims to help consumers choose energy-efficient products by imposing labelling requirements.\(^98\)

It follows that the above 'labelling' requirements aim to make information publicly available in a certain format, primarily for the benefit of consumers. They however do not regulate the issue of "data ownership" and thus do not prevent competitors or other service providers to collect, process or otherwise use such data for other purposes.

**Remunerated Data Sharing**

Other types of legislation, such as in the automotive sector, impose the sharing of information, but possibly against a fee. Indeed, under the **Vehicles Emissions Regulation** (No 715/2007), which aims to place specific obligations on manufacturers and open up access to repair and maintenance information, manufacturers must notably provide unrestricted and standardised information on vehicle repair and maintenance to independent operators. However, the Regulation specifies that in such event, the manufacturer is entitled to charge 'reasonable fees' for this service.\(^99\)

**Value-preserving Data Sharing**

In a series of other instruments, companies are required to disclose information to public authorities. They nevertheless provide for mechanisms, which are however limited, to preserve the value of such data or the exclusive rights that the disclosing company may have in such data. This is the case for instance in the following Directives:

• The **Public Access to Environmental Information Directive** (2003/4/EC), which aims to guarantee the public access to environmental information held by, or for, public authorities, both upon request and through active dissemination. Such Directive however provides that requests for access to information may notably be refused if the disclosure could be damaging one's intellectual property rights or commercial or industrial confidentiality. In this context, it shall be noted that the CJEU held in two recent judgments of 23 November 2016 that the confidentiality of

\(^{98}\) In the energy sector, the Energy Efficiency Directive (2012/27/EU) further provides for information sharing requirements, such as for instance by notably regulating the smart-metering information requirements to final customers.

commercial and industrial information may not be invoked to refuse the divulgation of information on biocides released into the air, water, soil and plants.\textsuperscript{100}

- The \textbf{INSPIRE Directive} (2007/2/EC), which lays down general rules setting up an infrastructure for spatial information in Europe for the purpose of EU environmental policies, provides for a similar restriction. Accordingly, Member States may limit public access to spatial data and services for various reasons, such as the confidentiality of commercial or industrial information\textsuperscript{101} or intellectual property rights.

- The \textbf{Internal Market in Electricity and in Gas Directives} (respectively 2009/72/EC and 2009/73/EC), which aim at introducing common rules for the transmission, distribution and supply of electricity and natural gas. In such context, electricity/gas providers are obliged to provide information to final customers. Moreover, both Directives provide for specific rules aiming for the preservation of the confidentiality of commercially sensitive information. They further include dedicated provisions in relation to retail markets, stipulating that "Member States shall ensure that the roles and responsibilities of transmission system operators, distribution system operators, supply undertakings and customers and if necessary other market parties are defined with respect to contractual arrangements, commitment to customers, data exchange and settlement rules, data ownership and metering responsibility."\textsuperscript{102}

It follows from the above Directives that there is an explicit acknowledgement of the issues of confidentiality of commercial or industrial information, as well as of a possible protection of such information by intellectual property rights, including to some extent of the issues of "data ownership".\textsuperscript{103}

\begin{flushleft}
\textsuperscript{100} Case C-44/14 Kingdom of Spain v European Parliament and Council of the European Union [2015] ECLI:EU:C:2015:554; Case C-673/13 P European Commission v Stichting Greenpeace Nederland and PAN Europe [2016] ECLI:EU:C:2016:889

\textsuperscript{101} Where such confidentiality is provided for by national or Community law to protect a legitimate economic interest (INSPIRE Directive, art 13(d)).


\textsuperscript{103} The Internal Market in Electricity and in Gas Directive does however not provide further details as to the concept of "data ownership", nor the practical implications of such ownership.
\end{flushleft}
Hence, while there is a general transparency obligation applicable to documents held by authorities\textsuperscript{104}, there are limits to such principles in favour of companies. Such complex system is particularly illustrative in the life science sector, notably under the Medicinal Products Directive (2001/83/EC) and the REACH Regulation (No 1907/2006). Such legislations indeed require companies to submit large amounts of information before being entitled to put medical or chemical products on the market.

Looking at the specific framework applicable to chemicals, as enshrined under the REACH Regulation and as implemented in practice by the European Chemicals Agency (hereafter "ECHA"), the assessment of confidentiality of business information is a regulatory process under REACH. Indeed, Article 119 of REACH defines the information that is held by ECHA on substances which are to be made publicly available, free of charge, over the Internet, as well as the information on substances for which an exception to publicity can be applied for by a party who would demonstrate that such publication would be potentially harmful to the commercial interests of the registrant or of any other party concerned.\textsuperscript{105} In accordance with Article 119 of REACH, some information can therefore be claimed to be confidential\textsuperscript{106}, provided however that a fee is paid and adequate justification is provided (such as a non-disclosure agreement, commercial interest, potential harm, etc.).

It follows from the above that the current legislation explicitly acknowledges in certain cases that certain types of information may have some value and thus require protection.


\textsuperscript{106} See for instance Article 119 of REACH, Article 67 of BPR, Articles 38 et seq. of CLP and Article 4 of ATD (see Annex I for excerpts such provisions).

\textsuperscript{106} Such as for instance Degree of Purity / Identity of Impurities and/or Additives; Certain information in the "Safety Data Sheet" or "Study Summary" or "Robust Study Summary".
2.5.2 Public Sector (open data)

Contextualisation

In the public sector, the adoption of specific rules under the notion of "open data" aims to regulate the control that public sector bodies have on their data. Accordingly, through legislative measures, the control of such bodies over data has been specifically regulated in order to impose an access right.

Pursuant to the ‘Open Definition’\(^{107}\), open data is "data that can be freely used, re-used and redistributed by anyone – subject only, at most, to the requirement to attribute and sharealike".

Its most important aspects are:

<table>
<thead>
<tr>
<th>Availability and access</th>
<th>The data must be available as a whole, at no more than a reasonable reproduction cost, preferably by downloading over the Internet. The data must also be available in a convenient and modifiable form.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Re-use and redistribution</td>
<td>The data must be provided under terms that permit re-use and redistribution including the intermixing with other datasets.</td>
</tr>
<tr>
<td>Universal participation</td>
<td>Everyone must be able to use, re-use and redistribute - there should be no discrimination against fields of endeavour or against persons or groups. For example, ‘non-commercial’ restrictions that would prevent ‘commercial’ use, or restrictions of use for certain purposes (e.g. only in education), are not allowed.</td>
</tr>
</tbody>
</table>

Table 2.2: Most important aspects of open data\(^ {108}\)

In the context of the Digital Single Market of the EU Commission, the concept of open data is defined as referring to "the idea that certain data should be freely available for use and re-use".\(^ {109}\)

More particularly, it refers to material produced, collected, paid for and/or held by public sector bodies at national, regional and local levels, such as ministries, agencies, municipalities, but also organisations mainly funded by or under the control of a public authority.\(^ {110}\)

---

\(^{107}\) Open Definition is a project of Open Knowledge International, available online at [http://opendefinition.org/](http://opendefinition.org/)

\(^{108}\) Open Definition 2.1, available online at [http://opendefinition.org/od/2.1/en/](http://opendefinition.org/od/2.1/en/)


\(^{110}\) European Commission, 'European Legislation on Reuse of Public Sector Information' (European Commission - Digital Single Market - Digital Economy &
The Open Data movement and governments around the world, including the EU, are committed to make data, and more particularly 'government data', publicly available and usable. Public entities generate and hold enormous amounts of data which in many cases are sensitive or confidential in nature. Government and public institutions have an inherent interest in managing carefully this large amount of data, both to improve their performance and generate savings that allow for much sought-after spending cuts, but also to be able to provide open data to their citizens and business entities.

The EU Commission claims to support open data for various reasons, including for the re-use thereof in new products and services, to face societal challenges, to achieve efficiency gains, and to increase transparency. Also, open data is considered to be crucial for EU research and development.

**Non-legislative Measures on Open Data**

The EU institutions have taken both legislative and non-legislative measures to encourage the uptake of open data.

On the non-legislative front, the EU Commission has been very active in the field of open data providing for soft measures facilitating access to data.

Its involvement includes:

- engaging with Member States through the Public Sector Information expert Group (PSI Group);
- funding the Open Data incubator, a 6-month incubator for open data entrepreneurs across Europe;
- funding the ePSIplatform;
- funding the Legal Aspects of Public Sector Information (LAPSI) network;
- commissioning studies related to the re-use of public sector information;
- developing the European Union Open Data Portal, which provides access to data from the EU institutions and bodies for re-use,

---


112 Commission, 'Open Data. An Engine for Innovation, Growth and Transparent Governance' (Communication) COM(2011) 882 final, 4

113 More information available online at <https://opendataincubator.eu/>

114 Accessible online at <http://data.europa.eu/euodp/en/data>

Ownership of Data 49 / 142
• contributing to the G8 Open Data Charter, which aims to open up government information.

**Legislative Measures on Open Data: Directive on the Re-use of Public Sector Information**

The EU legislative measures consist in the adoption of a Directive on the re-use of public sector information (government-held data). The EU Public Sector Information Directive (the “PSI Directive”)\(^{115}\), as revised by Directive 2013/37/EU, aims at unlocking the potential of (big) data held by public authorities. The EU Member States had to transpose the revised Directive into their national laws by 1 July 2015.

The PSI Directive is underpinned by two main principles of the EU internal market; *i.e.*, transparency and fair competition. As such, it stimulates the Member States to open up as much information for re-use as possible. It aims to regulate information held by public sector bodies in the Member States – at national, regional and local levels.\(^ {116}\) The targeted public sector bodies include ministries, state agencies, municipalities, organisations mainly funded by or under the control of public authorities, but also – since 2013 – museums, libraries and archives.\(^ {117}\) It however does not apply to the educational, scientific, and broadcasting sectors.\(^ {118}\)

The Directive targets "any content whatever its medium (written on paper or stored in electronic form or as a sound, visual or audiovisual recording) and any part of such content."\(^ {119}\) Put differently, it covers "any representation of acts, facts or information – and any compilation of such acts, facts or information" held by public sector bodies.\(^ {120}\) As such, it applies to different types and forms of information, such as written texts, databases, audio files, film fragments, etc.

The table below intends to provide an overview of the PSI Directive’s key aspects.

---


\(^{116}\) PSI Directive, art 1 and 2(1)

\(^{117}\) PSI Directive, art 1(2)(f)

\(^{118}\) PSI Directive, arts 1(2)(d) and 1(2)(e)

\(^{119}\) PSI Directive, art 2(3)

\(^{120}\) PSI Directive, Recital 11
<table>
<thead>
<tr>
<th><strong>Key aspects of the PSI Directive</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Reusability principle</strong></td>
</tr>
<tr>
<td><strong>Non-discrimination</strong></td>
</tr>
<tr>
<td><strong>Charging</strong></td>
</tr>
<tr>
<td><strong>Transparency</strong></td>
</tr>
<tr>
<td><strong>Prohibition of exclusive arrangements</strong></td>
</tr>
<tr>
<td><strong>Licences</strong></td>
</tr>
</tbody>
</table>

---

121 PSI Directive, art 3  
122 PSI Directive, art 10  
123 PSI Directive, art 6(1)  
124 PSI Directive, art 6(2)  
125 PSI Directive, art 7(1)-(3)  
126 PSI Directive, art 7(4)  
127 PSI Directive, art 11(1)  
128 PSI Directive, art 8(1)  
129 PSI Directive, art 8(2)
Ownership of Data

Public sector bodies shall make their documents available in any pre-existing format or language and, where possible and appropriate, in open and machine-readable format together with their metadata.  

Public sector bodies shall process the request for re-use and shall deliver the documents to the applicant or finalise the licence offer to the applicant within a timeframe of maximum 20 working days after its receipt.

More particularly regarding the licensing, the Commission published Guidelines in July 2014 to help the Member States implement the revised rules and to indicate best practices regarding recommended standard licences, datasets and charging for the re-use of public sector documents.

In these Guidelines, the Commission recommends using the Creative Commons CC0 public domain dedication, as it "allows waiving copyright and database rights on PSI, ensures full flexibility for re-users and reduces the complications associated with handling numerous licences".

The contents of the CC0 licence can be graphically summarised as follows:

<table>
<thead>
<tr>
<th>Commercial use</th>
<th>Yes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-commercial use</td>
<td>Yes</td>
</tr>
<tr>
<td>Modification</td>
<td>Yes</td>
</tr>
<tr>
<td>Distribution</td>
<td>Yes</td>
</tr>
<tr>
<td>Sub-licence</td>
<td>Yes</td>
</tr>
<tr>
<td>Responsibility for content</td>
<td>No</td>
</tr>
<tr>
<td>Attribution (credit)</td>
<td>No</td>
</tr>
<tr>
<td>Fees</td>
<td>No</td>
</tr>
<tr>
<td>Duration</td>
<td>Unlimited</td>
</tr>
</tbody>
</table>

If it is impossible to use the CC0 public domain dedication, the Guidelines encourage public sector bodies to use open standard licences appropriate to the Member State's intellectual property and contract law and that comply with the licensing provisions recommended in the

---

130 PSI Directive, art 5(1)  
131 PSI Directive, art 4(2)  
132 Commission Notice, Guidelines on recommended standard licences, datasets and charging for the reuse of documents [2014] OJ C 240/1  
133 Ibid 2
Commission’s Guidelines. The Guidelines further stimulate Member States to develop a suitable national open licence.¹³⁴

The licensing provisions recommended by the Commission in its Guidelines cover the following topics:

- **Scope** – temporal and geographical scope of the rights covered by the licence, the types of rights granted and the range of re-use allowed.
- **Attribution** (only when simple notices are not possible) – obligation on the re-user to acknowledge the source of the documents in a manner specified by the public sector body (licensor).
- **Exemptions** – indication of the datasets, if any, that are not covered by the licence.
- **Definitions** – concise definitions of the main terms of the licence.
- **Disclaimer of liability** – indication that the information is provided ‘as is’ and that the public sector body (licensor) assumes no responsibility for its correctness or completeness.
- **Consequences of non-compliance** – consequences of non-compliance with the licence terms.
- **Information on licence compatibility and versioning** – indication of the other licences with which the licence in question is compatible.¹³⁵

Neither the CC0 public domain dedication nor the Commission’s recommended licensing provisions for open data cover by any means the ownership of the data concerned. Indeed, the recommended licensing provisions are aimed at making the data available for re-use and excluding liability for any use that will be made of said data.

Therefore, under the PSI Directive, **the ownership of data is rather renounced than claimed**.

The same goes for the open data licensing strategies developed in the EU Member States pursuant to the revision of the PSI Directive and the Commission’s Guidelines. Such national open data licensing strategies are monitored by the EU Commission in the framework of the EU Open Data Portal mentioned above. Since 2015, the European Data Portal publishes the results of such monitoring in an annual report on the 'Open Data Maturity in Europe'.

¹³⁴ Ibid
¹³⁵ Ibid 3-4
As regards the Member States examined in the framework of this report, the results of the 2016 report\(^\text{136}\) on licensing norms can be summarised as follows:

<table>
<thead>
<tr>
<th></th>
<th>BE(^\text{137})</th>
<th>FR</th>
<th>DE</th>
<th>IT</th>
<th>ES</th>
<th>UK</th>
</tr>
</thead>
<tbody>
<tr>
<td>What type of licence does the national open data policy encourage using?</td>
<td>Standard licence</td>
<td>Etalab Open Licence (^\text{138})</td>
<td>Datenlizenz Deutschland(^\text{139})</td>
<td>Internationally recognised licences</td>
<td>Standard licence</td>
<td>Open Government Licence</td>
</tr>
<tr>
<td>Is all data free of charge?</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Is all data open licensed?</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
</tbody>
</table>

Table 2.5: Open data licensing norms in key Member States (2016)

In line with the Commission's recommended licensing provisions, none of the abovementioned national open data licences cover the ownership of the data. They do however extensively deal with the exclusion of liability of the public sector body.

By way of illustration, we provide some examples below of disclaimers of liability in the abovementioned national open data licences:

- **UK Open Government Licence**: "The Information is licensed 'as is' and the Information Provider and/or Licensor excludes all representations, warranties, obligations and liabilities in relation to the Information to the maximum extent permitted by law. The Information Provider or Licensor are not liable for any errors or omissions in the Information and shall not be liable for any loss, injury or damage of any kind caused by its use."

- **Etalab Open Licence (France)**: "The Information is made available as produced or received by the Producer, without any other express or tacit guarantee not specified in this licence. [...] It [the

\(^{136}\) Wendy Carrara, Margriet Nieuwenhuis and Heleen Vollers (European Data Portal), 'Open Data Maturity in Europe 2016. Insights into the European State of Play' (European Union 2016)

\(^{137}\) It shall be noted that the Belgian open data licensing framework is currently being reviewed.

\(^{138}\) A French free licence for open data published in 2012 by the French Republic.

\(^{139}\) Recommended terms of use for datasets by the German Federal Government
Producer] shall not be liable for any loss, injury or damage of any kind caused to third parties resulting from its re-use."

- **Datenlizenz (Germany):** "The provider does not assume any liability for the accuracy and completeness of data and contents and for permanent availability of services. Exempt from this are any claims for damage due to an injury to life, limb or health. Also exempt is damage based on wilfulness or gross negligence."

**Conclusion relating to Open Data**

Open data is one of the prominent areas in which licensing of data on (re)user-friendly terms is encouraged by regulatory measures, like the PSI Directive at EU level.

However, as demonstrated above, it would be wrong to argue that open data licence terms cover the issue of data ownership, as they merely regulate the liability related to said data. Pursuant to the licensing provisions recommended in the Commission's Guidelines, public sector bodies provide their data 'as is' and exclude all liability for any consequence of a further use of the data. In doing so, they renounce data ownership (be it indirectly) as opposed to claiming it. The concept of open data does therefore not address the issue of data ownership but only imposes access rights to certain types of data held by public bodies.
2.6 Other Data-related Legislation

In addition to the various legal instruments briefly analysed in the above sub-Sections and depicted in Figure 2.1 above, the legal framework related to data also comprises other legal grounds allowing companies to take action to protect their data or datasets. In the EU, such legal grounds are often not, or only to a limited extent, harmonised. A few examples of existing legislations or areas of law that may serve in a data context are presented below.

Criminal Law

Foremost, our study of the situation in various Member States has shown that criminal law may provide interesting solutions to protect data. National laws provide for criminal provisions in an intellectual property context, but also contain specific provisions applicable in well-defined circumstances (e.g., acts contrary to financial/monetary obligations or acts against the interests of a nation). In addition, national systems include ad hoc provisions applicable in an information technology environment to, for example, informatics forgery or fraud, data manipulation and hacking. In the same vein, specific investigation measures are available in such context, such as data seizure, network searching and the involvement of (IT) experts.

This being said, criminal grounds do not protect the data as such but rather offer a legal framework that may apply in a data context, depending on the circumstances at hand. The question arises however as to whether the "theft" of an intangible asset such as data can be envisaged. While the theft of intangible assets (such as water, gas or electricity) is recognised by law or jurisprudence, the theft of data remains a difficult issue. The issue remains therefore unclear in most Member States. The situations in France and Germany however offer interesting insights.

France

In a ruling rendered on 20 May 2015\(^{140}\), the French Supreme Court (the "Cour de cassation") held that a non-authorised data download, without the data carrier being stolen, may be qualified as theft. Pursuant to such case law, some authors have suggested that a property right in data may exist, but such view is not widely shared. Up until recently, French criminal courts used to dismiss the offence of data theft when the physical support of the data was not misappropriated.\(^{141}\)

\(^{140}\) Already mentioned on p. 32 supra

Germany

Some German academics claim that there exists, under German law, a right in data, as it would provide the basis for criminal law protection of data. The German Criminal Code deals with the protection of data as such in the following provisions:

- **Section 202a (data espionage):** "Whosoever unlawfully obtains data for himself or another that were not intended for him and were especially protected against unauthorised access, if he has circumvented the protection, shall be liable to imprisonment not exceeding three years or a fine."
- **Section 202b (phishing):** "Whosoever unlawfully intercepts data not intended for him, for himself or another by technical means from a non-public data processing facility or from the electromagnetic broadcast of a data processing facility, shall be liable to imprisonment not exceeding two years or a fine, unless the offence incurs a more severe penalty under other provisions."
- **Section 202c (acts preparatory to data espionage and phishing)**
- **Section 303a (data tampering):** "Whosoever unlawfully deletes, suppresses, renders unusable or alters data shall be liable to imprisonment not exceeding two years or a fine."

There is a consensus among scholars that these provisions aim to protect the authority to use the data.

Consequently, prominent German scholars have suggested creating a civil law ownership right in data on the basis of the data ownership concept developed in criminal law. They argue that the authority to use data protected by the Criminal Code amounts to an *erga omnes* right within the meaning of Section 823 of the German Civil Code, according to which "a person who, intentionally or negligently, unlawfully injures the life, body, health, freedom, property or another right of another person is liable to make compensation to the other party for the damage arising from this." They thus plead for the recognition under civil law of such *erga omnes* right, the beneficiary of which would be the creator of the data rather than the owner of the data carrier. The data creator would then be defined as the person storing the data, or the person responsible for storing the data, on a physical support.

**Unfair Competition and Tort Law**

In many Member States, it could be envisaged to initiate an action to (indirectly) protect data on the basis of unfair competition laws, which are designed to protect consumers and businesses against deceptive business practices. It is for instance possible in Belgium and France to

---

142 Hilgendorf, JuS 1996, 509 (511); Hoeren, MMR 2013, 486
143 Thomas Hoeren, 'Dateneigentum – Versuch der Anwendung von § 303a StGB im Zivilrecht' (2013) 8 MMR 486
bring, in certain circumstances, an action on the basis of unfair commercial practices or parasitism in the event of misappropriation of data outside a contractual framework.

One could also envisage relying on unfair commercial practices in a context of data to act against a company gaining a competitive advantage by not complying with the applicable legislation, such as for instance privacy or consumer protection law. In certain circumstances, this may also lead to antitrust issues.\(^{144}\)

Also, in some instances, tort law (liability outside a contractual relationship) may supplement unfair competition law. In such event however, strict conditions may apply. For instance, in Belgium and in France, it shall be demonstrated that the wrongdoer committed a fault and that there is a causal link between such fault and the damage suffered. It is however difficult to prove disloyal behaviour, such as the likelihood of confusion or the disruption of business.

**Contract Law**

Finally, it shall be noted that contract law may also serve as a basis to protect data. This being said, as the EU has taken limited initiatives in the field of contract law, the harmonisation is rather limited, in particular in a business-to-business context. The contractual arrangements in a data context are further examined in Chapter 4 below.

Chapter 3

Intellectual Property Rights and Trade Secrets

3.1 Copyright

It seems to be widely agreed amongst legal scholars that copyright can be invoked to protect, to a certain extent, non-personal and commercial data.

Below, we describe the current legal framework for copyright and explain the scope of the protection available. We also examine the relevant copyright aspects from a transactional point of view as well as the advantages and disadvantages of this type of protection in a (big) data context.

3.1.1 Legal Framework

The rules governing copyright protection have been established at international, regional and national level. In order to understand the protection granted to literary and artistic works, one needs to become familiar with this entire legal framework.

International Legal Framework

There exists no single international copyright instrument that would automatically confer uniform protection on literary and artistic works worldwide. However, international treaties, conventions and trade agreements were established as from the 19th century in order to ensure a minimal level of legal protection to creators of original works.

The international legal framework for copyright is based on the following principles:
The **territoriality principle** refers to the fact that copyright is of a territorial nature and that national laws can only rule on conducts occurring within national borders.\[145\]

According to the **national treatment principle**, a country must provide the nationals of other countries, party to the same international instruments, with a treatment no less favourable than the one it accords to its own nationals with regard to such rights.\[146\] There are however certain exceptions to this principle.

**Reciprocity** is the negation of the national treatment principle as it refers to making protection, or the extent of protection, in a given country (A) of copyright or related rights of nationals of another country (B) conditional on the existence of the same (or at least similar) extent of protection granted in that other country (B), to the nationals of the country concerned (A).\[147\]

This Report does not aim at analysing issues in relation to territoriality, national treatment, jurisdiction and conflicts of laws. It is however important to understand that the above principles are necessary in the field of copyright because copyright laws are not identical from country to country.

It is important to remember that the international treaties provide for minimum standards only and individual countries may therefore provide for additional protection. Also, treaties do not cover some important issues like ownership of data and transfer of rights.

---

\[145\] This has been confirmed by the Court of Justice of the European Union in *Lagardère*, wherein it states that "it must be emphasised that it is clear from its wording and scheme that [the Rental and Lending Directive] provides for minimal harmonisation regarding rights related to copyright. Thus, it does not purport to detract, in particular, from the principle of the territoriality of those rights, which is recognised in international law and also in the EC Treaty. Those rights are therefore of a territorial nature and, moreover, domestic law can only penalise conduct engaged in within national territory" Case C-192/04 *Lagardère Active Broadcast v Société pour la perception de la rémunération équitable and others* [2005] ECLI:EU:C:2005:475, para 46

\[146\] World Intellectual Property Organization, 'Guide to the Copyright and Related Right Treaties Administered by WIPO and Glossary of Copyright and Related Rights Terms', (WIPO 2003) 297

\[147\] Ibid 306
The main international instruments of copyright law are the following:

- The Berne Convention\(^{148}\)
- The Universal Copyright Convention\(^{149}\)
- The TRIPS Agreement\(^{150}\)
- The World Copyright Treaty\(^{151}\)

**European Union Legal Framework**

In addition to the international treaties, to which the European Union and the 28 Member States are contracting parties, a series of EU Directives was also adopted to harmonise various discrepancies between the copyright laws of the Member States, notably among civil law and common law jurisdictions.

In spite of these Directives, there currently is no common and fully harmonised legal framework for copyright within the EU. Since copyright laws are to a great extent territorial in each Member State, the international treaties and national legislations remain important sources of copyright law.

The most important EU instrument related to copyright is the Information Society Directive (the InfoSoc Directive)\(^{152}\), which can be of particular importance in the context of the issues related to the ownership of data. It aims at (i) adapting the legislations on copyright and related rights to reflect the technological developments; and (ii) transposing into EU law the main international obligations arising notably from the World Copyright Treaty.

---

148 The Berne Convention for the Protection of Literary and Artistic Works of 9 September 1886
149 Universal Copyright Convention adopted in Geneva on 6 September 1952, as revised in Paris on 24 July 1971
150 The Agreement on Trade-Related Aspects of Intellectual Property Rights, adopted in Marrakech on 15 April 1994 and which corresponds to Annex 1C to the Agreement establishing the World Trade Organization
In addition to the InfoSoc Directive, the EU has adopted other instruments particularly important in the context of copyright, such as in particular:

- The Database Directive;\(^{153}\) (see Section 3.2 below for further details)
- The Rental and Lending Directive;\(^{154}\)
- The Term Directive;\(^{155}\)
- The Software Directive;\(^{156}\)
- The Orphan Works Directive.\(^{157}\)

The Court of Justice of the European Union (the "CJEU") has played an important role in the harmonisation of copyright (and database rights) by interpreting the various Directives listed above, and in particular the InfoSoc Directive and the Database Directive. Although legal systems of most of the EU Member States are based on continental law, which means that they do not pay so much importance to case law as common law countries, the judgments of the CJEU play an important role in providing a binding interpretation of EU law.

**The EU Copyright Reform**

On 14 September 2016, the Commission published several legislative proposals aiming to modernise the existing EU copyright rules.\(^{158}\) The so-called Copyright Package consists of two Directives and two Regulations.

What seems to be at the core of the reform, from the point of view of digital services providers, are the Proposals for a Directive on copyright in the Digital Single Market\(^{159}\) and for a Regulation laying down rules on the exercise of copyright and related rights applicable to certain online transmissions of broadcasting organisations and retransmissions.\(^{160}\)

---


\(^{160}\) Commission, ‘Proposal for a Regulation of the European Parliament and of the Council laying down rules on the exercise of copyright and related rights applicable
According to the Commission's announcement, these Proposals aim to ensure better choice and access to content online and across borders, improved copyright rules on research, education and inclusion of disabled people, as well as a fairer and sustainable marketplace for creators, the creative industries and the press.

More specifically the Proposal for a Directive on copyright in the Digital Single Market introduces a new related right for press publications, which is supposed to give the press industry a stronger bargaining position to protect their investments, explore new business models and eventually complete its transition to the digital environment.

The Proposal also further harmonises the copyright exceptions by introducing three new mandatory exceptions covering:

- digital uses of works or other protected content for the purposes of illustration for teaching;
- text and data mining by research organisations, acting in the public interest, of copyright protected content to which they have lawful access (this exception is limited to non-commercial purposes) (see sub-Section 3.1.4 below for further details); and
- copying of works by cultural heritage institutions in a way that is suitable for the digital environment (this exception will cover works that were created directly in digital form as well as the digitisation of works in analogue formats, and will help audiences to access them for longer).

This Proposal for a Directive on copyright in the Digital Single Market is consistent with the existing EU copyright legal framework, and is thus based upon, and aims to complement the rules laid down in various copyright directives, including the InfoSoc and Database Directives.

The four legislative Proposals presented by the Commission were submitted to the European Parliament and to the Council for adoption. It shall nevertheless be noted that the Proposals do not aim to clarify the protection of data under copyright law nor provide for new rules relating to the development and increased use of digital tools such as big data and the Internet of Things. They however include, as indicated above, a new – yet narrow – exception concerning text and data mining (see sub-Section 3.1.4 below for further details).

\[\text{COM(2016) 594 final}\]

161 Once the proposals are adopted, the Member States will have two years to implement the directives into the national law.
National Legal Framework

As indicated above, international treaties lay down the core principles of copyright protection. Various EU Directives, as interpreted by the CJEU, provide for a certain degree of further harmonisation in the EU.

However, although the copyright rules applicable in the Member States are similar, the threshold of protection, the exceptions, the practical implementation and the enforcement proceedings and remedies differ substantially. It is therefore of utmost importance to take into consideration the national legal traditions, examining both the applicable national legislation and its interpretation by national courts.

3.1.2 Copyright Protection: General Overview

Scope of Copyright Protection

To understand to what extent copyright may be used to protect non-personal data, one must first understand what types of creations can be protected, and what are the terms and scope of such protection.

The Berne Convention presents a broad non-exhaustive list of works protected under copyright¹⁶²:

"The expression "literary and artistic works" shall include every production in the literary, scientific and artistic domain, whatever may be the mode or form of its expression, such as books, pamphlets and other writings; lectures, addresses, sermons and other works of the same nature; dramatic or dramatico-musical works; choreographic works and entertainments in dumb show; musical compositions with or without words; cinematographic works to which are assimilated works expressed by a process analogous to cinematography; works of drawing, painting, architecture, sculpture, engraving and lithography; photographic works to which are assimilated works expressed by a process analogous to photography; works of applied art; illustrations, maps, plans, sketches and three-dimensional works relative to geography, topography, architecture or science."

It derives from that list that copyright protection has a broad scope but, at the same time, requires an intellectual human intervention and the consciousness of achieving a result. Therefore, raw data such as weather forecasts, stock quotations or sports scores would in principle be excluded from copyright protection.

¹⁶² Article 2(1)
The EU legal framework does not provide for a list of protected works like the Berne Convention does. Member States have however implemented Article 2(1) of the Berne Convention directly in their national legal frameworks. This approach implies that, in principle, any type of work can enjoy copyright protection as long as it meets the legal requirements for such protection.

Because they do not meet the fundamental requirements for copyright protection, copyright statutes and treaties (particularly the TRIPS Agreement and the World Copyright Treaty) exclude mere ideas from copyright protection. However, the expression of such ideas may be protected.

In the same vein, because their subject matter is considered as being outside the scope of copyright protection, mathematical concepts, methods of operation, gambling procedures and other intellectual tools are also excluded from copyright protection.

**United Kingdom**

UK law places particular emphasis on the formal expression of an idea as being at the heart of copyright protection. Hence, certain forms may not be protected by copyright; e.g., technical features such as the functionality, programming language and interfaces (such as data file formats) of computer programs are not themselves protected by copyright although the software's source code which creates them is.

Furthermore, for a work to be protected, it must fulfil two cumulative conditions:

- it shall be fixed in some material (concrete) form;
- it shall be original, meaning that it is the author's own original creation and reflects his/her personality, where he/she has been able to express his/her creative freedom by making free and creative choices and thus stamping his/her personal touch onto the work.

---

162 Article 9.2  
164 Article 2  
165 In *SAS Institute Inc v World Programming Limited* [2013] EWCA Civ 1482, Lewison LJ found that both the Software and the InfoSoc Directives incorporated the underlying principle from the Berne Convention that it was the form of expression rather than the underlying idea which was protected. The Court of Justice of the European Union found that whether it applied the Software Directive or the InfoSoc Directive, the functionality of the software in issue was not protected given that the functionality was the idea, but the source code was the expression in which that idea was embodied.
The originality criterion implies that some categories of data will not be protected by default. Having said that, the threshold for originality is rather low in the EU Member States, and even more in some of them (this is for instance the case in the Netherlands, in France and in Belgium).

The right owner (or the right holder in case of transfer of rights) of copyright protected works will enjoy various exclusive economic rights; i.e., to reproduce, communicate to the public and distribute the work. Accordingly, save where copyright exceptions apply, the author's consent is necessary to perform any of these activities (see sub-Section 3.1.4 below for further details).

Moreover, authors are also granted so-called "moral rights". The concept of "moral rights" is the consequence of the predominant view in (continental) European copyright law that a work is not a mere staple commercial object, but also the expression of the personality of the author. Moral rights are recognised by the Berne Convention. Its Article 6bis provides for minimum standards in this respect: the author has the right, even after the transfer of the economic rights, to claim authorship of the work and to object to derogatory action (distortion, mutilation or other modification) to the works which would be harmful to the author's honour or reputation. By contrast, the EU Directives explicitly exclude moral rights from their scope. More particularly, Recital 19 of the InfoSoc Directive stipulates that moral rights remain outside the scope of the Directive and that they should be exercised according to the legislation of the Member States and the provisions of the international treaties. It follows from such situation that moral rights suffer many discrepancies between Member States. Indeed, while some countries provide for a high level of protection of moral rights, others recognise such rights only within the minimum protection imposed by the Berne Convention. Some Member States even foresee additional moral rights.

One of such moral rights is the French "droit de repentir", which can be particularly relevant in a big data context since it allows the author to take his work back from the commercial circuit, making further exploitation of such work impossible. Even if the author needs to compensate the person who acquired the economic rights to the work for such a withdrawal, exercising the "droit de repentir" could prejudice an entire (big data) project based on the withdrawn work.

---

**France**

One of such moral rights is the French "droit de repentir", which can be particularly relevant in a big data context since it allows the author to take his work back from the commercial circuit, making further exploitation of such work impossible. Even if the author needs to compensate the person who acquired the economic rights to the work for such a withdrawal, exercising the "droit de repentir" could prejudice an entire (big data) project based on the withdrawn work.

---

166 At international level, moral rights are also recognised by Article 5 of the WIPO Performances and Phonograms Treaty of 1996 and Article 5 of the Beijing Treaty on Audiovisual Performances adopted in 2012.
167 Article L.121-4 of the French Intellectual Property Code
The general eligibility for copyright protection differs to a certain degree from one EU Member State to another. Indeed, the abovementioned abstract concepts have usually been specified in detail through the case law of each national state. Having said that, the numerous CJEU decisions in copyright-related matters lead to a gradual unification of the EU legal framework for copyright, with the notable exception of moral rights.

Ownership of Rights

In general, the copyright belongs to the (physical) author of the work. In case of works created by two or more persons, the copyright would be awarded to these persons jointly.

In the jurisdictions (such as France) recognising collective works (such as encyclopaedia or periodic reviews), the economic rights to such works will normally belong to their producer or publisher.

The question of ownership is usually more complex in case of works created by an employee. In some countries, the economic rights to works created by an employee in result of carrying out his contractual employment duties will be automatically transferred to his employer (unless otherwise agreed in the contract), whereas in other countries such presumption does not exist.

Works in the Public Domain

Authors of protected works benefit from copyright during their entire life, and these rights are maintained for a period of 70 years after their death (or the death of the last author), before falling into the public domain. In the EU Member States, the initial length of protection was 50 years after the author's death (as it is still prescribed by Article 7(1) of the Berne Convention) but the Term Directive increased the protection term to 70 years in the EU.168

Once a work falls into the public domain, it means that it can be freely exploited, reproduced or executed. No authorisation is needed and no royalties must be paid for the use of the work. However, some Member States have established a system of "domaine public payant". This is the case in Italy and the question is currently highly debated in France. Such system implies the payment of a royalty, for the use of a work comprised within the public domain, which will be bestowed to cultural purposes.

168 With the notable exception of France, where moral rights are perpetual, in most jurisdictions the 70-year term of protection after the death of the author applies to both economic and moral rights.
3.1.3 Copyright Protection of Data

As mentioned in the introductory part of this report, from the business perspective, the protection of (big) data is needed to secure the economic investment made in obtaining, verifying, storing, presenting and analysing data.

Copyright enables protection of non-personal and commercial data to a certain extent. However, it is crucial to distinguish different elements used in the operations on big data that can benefit from copyright protection.

Individual data, understood as pieces of information, can be protected by copyright as long as they fulfil the conditions set out in the relevant legislation (in particular fixation in a tangible form and originality – see above for further details).

As the trend to modernise the existing legal framework confirms, the traditional copyright laws have struggled to deal with new technologies and digital content distribution methods. In practice, it means that the current laws do not contain provisions that would directly address the use of protected works when applying new technologies and in the new digital context, in particular for cloud computing, text and data mining\textsuperscript{169} or big data projects. Nevertheless, the copyright rules will still apply, providing protection for those materials used that could be classified as works and would fulfil the protection requirements described in sub-Section 3.1.2 above.

In the context of big data projects, it is crucial to understand to what extent the data used can be copyright protected. Unfortunately, there is no unequivocal answer as to what types of data fall under such protection, and thus, the eligibility for protection needs to be examined on a case-by-case basis and in light of the particular rules and case-law in each country.

Also, given that the copyright legal framework does not provide for a registration system (unlike trademarks and patents), copyright protection will only be confirmed \textit{a posteriori} by a court. Such characteristic of copyright is particularly problematic in a context of data, posing issues of a lack of legal certainty.

Having said that, there are some objective criteria that can facilitate analysing whether or not specific data is protected. In particular, the data need to fulfil the two basic requirements for copyright protection –

\textsuperscript{169} See however sub-Section 3.1.4 below which discusses the proposed new exception to cover text and data mining.
they need to be fixed in some material (concrete) form and they need to be original.

In this context, ‘fixation’ of data means that the specific information needs to be saved in a tangible form. The form of saving the data can differ from handwritten notes (files), through photographic documentation (image) or recorded testimonies (sound) to digitised archives (digital files), as long as it remains concrete, can be easily identified and described. Results that have not yet been produced (future data), or results that cannot yet be described (e.g., because there are no means yet to express them) cannot benefit from copyright protection for as long as they have not materialised. This can present some difficulties in a big data context, given that big data tends to involve dynamic datasets and notably relies on cloud computing services.

The originality requirement can bring even more difficulties, since the evaluation of a work's originality leaves some room for discretion and requires, in any event, a human intervention in the creation process, whereby he/she can stamp the work with his/her personality.

In general, in order to be considered original, the data should represent a level of sophistication suggesting that no one else than the author could have created the same work even if based on the same raw data (e.g., summary texts). Having said that, it is indeed so that the originality threshold for copyright protection in most of the EU Member States is rather low. Even a low level of creativity can therefore prove to be sufficient to claim protection. Such low threshold is however criticised by numerous scholars.

On that basis, we can attempt to identify what types of data will be more (and less) likely to benefit from the copyright protection. However, such identification should only be seen as indicative. In practice, every piece of data would have to be evaluated on a case-by-case basis in order to determine whether it can be copyright protected in a particular country.

The following data, in our view, could more easily attract copyright protection:

- data in the form of free text;
- data presented in graphic form.

On the other hand, for the following types of data it would, in our view, be particularly difficult to demonstrate originality:

---

If we were to ask ten photographers to take a picture of the same object on the same day and time, we would most probably still obtain ten different photographs, each of them embedding an individual view of the artist, and reflecting his/hers artistic effort.

---

Ownership of Data
• raw numbers and other purely quantitative information;
• measurements results (e.g., measurements of temperature, pressure, other natural phenomena);
• financial results, prices of products and similar market data;
• sport results, competition results;
• demographic data;
• results of automated processes (e.g., video recording from security cameras, statistics on the use of electricity, water, use of the telephone (number of calls, use of data transfer), Internet (e.g. use of the browser)).

On the basis of the concrete analysis performed in the context of the TOREADOR project, it follows that most of the data collected and processed in the context of the various use-cases will not benefit from copyright protection.

Having said that, it cannot be excluded that the individual data presented above can gain originality once they are connected with other information or presented in an original way (by means of different possible forms of expression).

3.1.4  Exclusive Rights and Copyright Exceptions

In the event a particular piece of data is protected by copyright, the rightowner (or rightholder) will be granted several exclusive rights. Accordingly, when a work is protected, seeking authorisation is a requirement in order to reproduce, communicate or make available to the public, distribute, rent, lend, adapt, translate or alter such work. However, copyright laws include various exceptions (limitations) where, under specified conditions, such authorisation is not required.

**Exclusive Rights**

On the basis of the InfoSoc Directive, Member States are required to implement the following set of exclusive rights:

<table>
<thead>
<tr>
<th>Reproduction</th>
<th>Communication to the public</th>
<th>Distribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exclusive right to authorise or prohibit direct or indirect, temporary or permanent reproduction by any means and in any form, in whole or in part.(^{171})</td>
<td>Exclusive right to authorise or prohibit any communication to the public of their works, by wire or wireless means, including the making available to the public of their works in such a</td>
<td>Exclusive right to authorise or prohibit any form of distribution to the public by sale or otherwise.(^{173})</td>
</tr>
</tbody>
</table>

\(^{171}\) InfoSoc Directive, art 2  
\(^{173}\) InfoSoc Directive, art 4
way that members of the public may access them from a place and at a time individually chosen by them.\footnote{InfoSoc Directive, art 3}

<table>
<thead>
<tr>
<th>Table 3.1: Overview of the author's exclusive rights</th>
</tr>
</thead>
</table>

From a technical perspective, (big) data analysis generally requires the reproduction of the data, in part or in whole.\footnote{This is for instance the case in the context of the TOREADOR project which relies on the Apache Kafka stream processing platform. Such process implies data partitioning and replication.}

It can therefore not be excluded that a (big) data analysis process leads to the reproduction of copyright-protected data, requiring thus the authorisation of the rightowner/rightholder (except if one may rely on one or more copyright exceptions – see below).

Indeed, the reproduction right under the InfoSoc Directive has an extensive meaning, as a broad definition is needed to ensure legal certainty within the internal market.\footnote{InfoSoc Directive, Recital 21} Such view was confirmed by the CJEU:

- Infopaq judgment (C-5/08) of 16 July 2009: "an act occurring during a data capture process [e.g., creation of TIFF file by scanning, transferred to an OCR server to be translated in order to be processed digitally], which consists of storing an extract of a protected work comprising 11 words and printing out that extract, is such as to come within the concept of reproduction in part within the meaning of Article 2" (Infopaq, operative part).
- Premier League (C-403/08) of 4 October 2011: "the reproduction right extends to transient fragments of the works within the memory of a satellite decoder and on a television screen" (Premier League, operative part).

It follows that even technical reproductions or the transformation of files into other file formats will qualify as a "reproduction" within the meaning of EU copyright law. In our view, such conclusion is criticisable in the context of the digital era. Indeed, in many instances, the (partial) reproduction of protected works is performed for merely technical reasons (e.g., cache copies, temporary copies in the RAM memory, back-up reproductions to prevent data loss, copies made for availability and high performance purposes such as by CDNs, etc.).

As for the communication to the public right, it is less problematic in our view in a (big) data analysis context. Such conclusion applies despite
the far-reaching concept of communication to the public, as interpreted by the CJEU, which notably concluded the following:

- Establishing whether a communication to the public took place requires an individual assessment (*Phonographic Performance*, C-162/10, judgment of 15 March 2012).
- For there to be a communication, the user’s intervention needs to be of a deliberate character, and ‘public’ refers to an indeterminate number of potential viewers, implying a fairly large number of people (*SCF*, C-135/10, judgment of 15 March 2012).
- The concept of ‘communication to the public’ includes two cumulative criteria, i.e. an ‘act of communication’ of a work and the communication of that work to a ‘public’ (*Svensson and Others*, C-466/12, judgment of 13 February 2014).
- To be categorised as a ‘communication to the public’, a protected work must be communicated using specific technical means, different from those previously used or, failing that, be directed to a ‘new public’, i.e. a public that was not already taken into account by copyright holders when they authorised the initial communication to the public of their work (*Svensson and BestWater International*, C-348/13, order of 21 October 2014).
- The making available of audio-visual content through a technology other than that previously authorised by rightholders entails *ipso facto* a new communication to the public that requires a new authorisation of the rightholders. (*ITV Broadcasting and Others*, C-607/11, judgment of 7 March 2013).

Finally, with respect to the **distribution** right, it may apply in the context where the recipients of the data are not only those involved in the (big) data analysis process. Hence, where the dataset is copied and distributed to the public (e.g., if it is sold), the exclusive distribution right may apply. This being said, the distribution right finds an important limit in the exhaustion (also known as “first sale”) doctrine. It allows the resale of copies of works (or related subject matter) without authorisation once the protected work has been put in the market with the consent of the rightowner/rightholder.

**Copyright Exceptions and Limitations**

The InfoSoc Directive also provides for certain exceptions (or limitations) to the exclusive rights. Its Article 5 contains one mandatory exception for temporary acts of reproduction, and an exhaustive list of optional exceptions that Member States may implement into their national law.

Foremost, a mandatory exception to the right of reproduction is introduced with respect to certain temporary acts of reproduction which
Ownership of Data

are integral parts to a technological process. Generally, such exception concerns transient copies with a merely technical function and without any independent economic significance, in order to cover issues related to caching and Internet browsing.

The CJEU has had the opportunity to examine such mandatory exception in several cases - i.e., Infopaq I and II, and Premier League. Apart from stating that a copyright exception must be interpreted restrictively, it also indicated that several cumulative conditions must be met in order to benefit from such exception:

- The temporary copy must be transient or incidental. This means that the copy may only be ephemeral or at least non-permanent.
- The copy must be an integral and essential part of a technological process. In such context, the CJEU confirmed that the concept of the 'integral and essential part of a technological process' requires the temporary acts of reproduction to be carried out entirely in the context of the implementation of the technological process and, therefore, not to be carried out, fully or partially, outside of such a process. It further stated that this concept also assumes that the completion of the temporary act of reproduction is necessary, in that the technological process concerned could not function correctly and efficiently without that act.
- The sole purpose of the copy must be to either enable a transmission in a network between third parties and an intermediary, or a lawful use of a protected work.
- The temporary copy must have no independent economic significance. Accordingly, the temporary reproduction cannot enable the generation of an additional profit, going beyond that derived from lawful use of the protected work. Also, the acts of temporary reproduction cannot lead to a modification of the work.

---

176 InfoSoc Directive, art 5.1
177 InfoSoc Directive, Recital 33
179 Case C-302/10 Infopaq International A/S v Danske Dagblades Forening [2012] ECLI:EU:C:2012:16 ("Infopaq II")
180 Joined Cases C-403/08 and C-429/08 Football Association Premier League Ltd and Others v QC Leisure and Others (C-403/08) and Karen Murphy v Media Protection Services Ltd (C-429/08) [2011] ECLI:EU:C:2011:631 ("Premier League")
182 Infopaq II, para 30
183 Infopaq II, para 54
It results that the acts performed on data in the data value cycle, by various stakeholders, may hardly rely on the exception for temporary acts of reproduction. This is notably supported by the fact that the acts performed in a big data analytics process can have a great economic value contrary to the last condition set under Article 5.1 of the InfoSoc Directive.

Accordingly, in our view, there is little to no legal certainty with the current exception for temporary acts of reproduction. Hence, authorisation remains a key requirement without which any use of a copyright protected work would amount to copyright infringement and would give rise to possible enforcement proceedings.

In addition to the mandatory exception discussed above, the InfoSoc Directive provides for several optional exceptions. Discretion is left to the Member States with regard to the transposition of such exceptions (e.g., with regard to the conditions and practical arrangements of such exceptions); consequently, the scope of exceptions differs largely between the Member States.

As part of the optional exceptions, there is one related to the use of a copyright protected work for the sole purpose of illustration for teaching or scientific research, for non-commercial purposes. Such exception however also presents some limits and is not entirely satisfactory as to provide sufficient certainty in case a stakeholder of the data value cycle wishes to rely on it. Also, its non-mandatory nature has led to having discrepancies between Member States, which is certainly not desirable in a (big) data analysis context.\(^{184}\)

Finally, it should be mentioned that the EU Commission has made a Proposal in the context of the EU Copyright reform (see above) to include a new mandatory exception to cover "text and data mining". The definition of such concept is however broader and appears to cover the more general idea of "data analysis".\(^{185}\)

Such new exception would allow reproductions and extractions made by research organisations in order to carry out text and data mining of works or other subject matter to which they have lawful access for the purposes of scientific research. Unfortunately, such exception would be

---

\(^{184}\) See the study of Jean-Paul Triaille 'Study on the Legal Framework of Text and Data Mining (TDM)' for further details
<http://ec.europa.eu/internal_market/copyright/docs/studies/1403_study2_en.pdf>

\(^{185}\) Proposal for a Directive in the DSM, art 2.2: "text and data mining' means any automated analytical technique aiming to analyse text and data in digital form in order to generate information such as patterns, trends and correlations"
very narrow, similar to what has been adopted in the United Kingdom, where "text and data mining" is excluded for commercial purposes.

Although research organisations should also benefit from the exception when they engage into public-private partnerships\textsuperscript{186}, it is regrettable that the Proposal provides for such a narrow opening to permit the reproduction of data in an analysis context. It is indeed short-minded to restrict acts of reproduction for the mere purpose of data analysis only to the scientific sector.

In our opinion, there is a missed opportunity to rethink the essence of copyright and introduce a new infringement test through the creation of a new requirement: the use of the copyright-protected work as a "work".\textsuperscript{187} Indeed, in many instances, when a protected work is used for (big) data analysis, it is not reproduced with the same aim as the original purpose. Accordingly, the rightholder should not be in a position to hinder the reproduction of (part of) its work which is to be used for a different objective, especially in case such other objective fulfils the main conditions of the three-step test; \textit{i.e.} it does not conflict with a normal exploitation of the work and does not unreasonably prejudice the legitimate interests of the rightholder.\textsuperscript{188}

\subsection*{3.1.5 Contractual Aspects}

What is crucial in the context of big data projects is for the data to be operational and transferable. In particular, one needs to have the possibility to aggregate, reproduce, filter, enrich, merge, partition, share, etc. the data, and use them as a source of information. To that end, the ownership of data needs to be identified and access rights secured. It is also necessary to ensure a possibility to convey data to a third party \textit{(i.e., make the data available on the basis of transfer or licence agreements)}, without risking that the acquirer will have to face property claims from the (alleged) data owner.

\textsuperscript{186} Proposal for a Directive in the DSM, Recital 10 \textsuperscript{187} Such kind of test already exists to a certain extent in the field of trademarks. See also Alain Strowel, 'Reconstructing the Economic Rights: Taking Copyright Seriously' (2016), presented in the context of the research project "Reconstructing rights: Rethinking copyright's economic rights in a time of highly dynamic technological economic change", led by IVIR & CREATe. \textsuperscript{188} The latter two conditions derive from the three step test enshrined under Article 5.5 of the InfoSoc Directive. The latter provides the exceptions and limitations permitted by the InfoSoc Directive are to be applied \textit{(i)} in certain special cases, \textit{(ii)} which do not conflict with the normal exploitation of the work or other subject matter and \textit{(iii)} which do not unreasonably prejudice the legitimate interests of the right holder (or other right holders).
Therefore, in case data benefits from copyright protection, it is important to understand the rules governing the conclusion of transfer and licence agreements.

**Transfer of Copyright**

In order to permanently assign the ownership of a right to another person, a so-called transfer agreement should be concluded. Although the rights can also be donated or inherited, the transfer agreement (that can be compared to a sale) is the most common tool used in a commercial context.

When concluding a transfer agreement, it is very important to precisely describe the subject matter of the rights that will be transferred (i.e., the work), and, if possible, attach a copy of such work (or its specification) as an annex to the agreement. One should however note that in certain Member States, it is not possible to transfer rights to "all works" of a specific author or even "all works of a specific type"; if such provisions are included in the copyright transfer agreement, they may be declared null and void.

For practical reasons, it is also crucial to indicate a precise date of the copyright transfer. That allows establishing as of when one can use the right, and is important in case of infringements, as well as determining tax obligations.

Distinction needs to be made between transferring the ownership of a physical copy of the work (e.g., a movie on DVD, piece of furniture, or a sculpture) and transferring the copyright in relation to that work (intangible rights to the movie, the design of the furniture or sculpture). Selling a physical copy of the work will in principle not transfer the copyright to such work.

Since the moral rights are not attached to the economic rights, they will not be transferred along with the economic rights. Therefore, depending on whether the national jurisdiction allows for the transfer of moral rights, one should include in the copyright transfer agreement additional provisions, such as:

- provisions transferring the moral rights of the author to another party (if possibility to transfer such rights is recognised by the jurisdiction);
- provisions authorising the acquirer of the economic rights to exercise moral rights on behalf of the author (most common solution); or
- provisions imposing an obligation on the author not to exercise his moral rights against the acquirer of the economic rights.
In addition, inclusion of a separate clause is necessary to allow the acquirer to be able to create derivative works. Otherwise the acquirer will only be able to use the original work 'as is', and will not be able, on the basis of the transfer agreement, to modify the works acquired. This is a very important element of the agreement, since even the slightest adaptation of the work, even in its technical sense, may be seen as a modification that leads to the creation of a new, derivative work.

This is particularly relevant when considering the transfer of copyright in a big data context (provided such protection applies). Indeed, datasets are usually combined, processed and altered in some way. Also, the ultimate goal of big data analytics is to perform an analysis and thus create "derivative" information on the basis of the initial dataset(s).

Most jurisdictions require concluding the copyright transfer agreement in writing. Where this is the case, a transfer concluded with the omission of the written form will be considered null and void.

Also, in some jurisdictions it is necessary to indicate the so-called fields of exploitation of the work (e.g., in Poland it is an obligatory element of every copyright transfer agreement – if the fields of exploitation are not mentioned, the agreement risks to be considered null and void189).

**Licence Agreements**

By concluding a licence agreement, the right holder authorises a third party to use the work within the limits indicated in the licence, usually in return for remuneration (licence fees). The ownership of the copyright remains with the right holder.

There exist different types of licence agreements, depending on the territory, for how long and with what level of autonomy the licensee will be able to use the work.

Since the copyright protection remains territorial due to differences in the national law of the Member States, the licence agreement will often not cover a territory larger than one country. In practice, this is quite problematic – in order to be able to use a given work on the territory of the whole EU one will need to conclude 28 separate agreements.

A licence can be exclusive, guaranteeing the licensee to be the only entity able to use the licensed work in a given territory. This provides the licensee with an obvious competitive advantage, but also impacts the (higher) level of the licensor’s remuneration.

189 Article 41 of the Polish law on copyright and related rights of 1994
In most countries licence agreements, contrary to transfer agreements, do not have to be concluded in writing.

**Free and Open Licences**

There are many initiatives aiming at facilitating the use and re-use of works to stimulate creativity, facilitate expression and enable fast sharing of information. These initiatives usually rely on what is popularly called "open" or "free" licences. Such licences can be characterised by two common features – they have predefined terms and conditions and are available for all the interested parties (everyone can enter into such agreement).

Also, since in most jurisdictions it is impossible for the author to effectively waive his copyright, the authors who want to freely spread the use of their work will rely on open licences.

The most popular types of open licences are the licences designed by Creative Commons. Creative Commons is a non-profit organisation that enables the sharing and use of creativity and knowledge through free legal tools. Creative Commons licences are based on standardised licensing terms embedded in open licence agreements. The popularity of these licences mainly results from the fact that their creators were able to describe the terms and conditions of the licences in a graphic, easily accessible form. Because this system is simple and easy to use, it is frequently recommended, especially in relations between the public and private sector.

The importance of open licences in the context of big data projects comes from the fact that the use of such licences significantly improves the possibility to allow access to data. Not only does it facilitate conducting the clearance of rights, but it also removes the necessity of negotiating the conditions of the agreement with the copyright holder. This being said, similarly to what was concluded in sub-Section 2.5.2 above in relation to "open data" (which also relies on an "open" system), the associated licences do not address the issue of data ownership.

### 3.1.6 Importance for Big Data projects

**Enablers (advantages)**

There are several features of copyright protection that can be seen as beneficial in the context of big data projects.

---

190 More information on the Creative Commons licenses can be found on the following website: <http://www.creativecommons.be/node/3>
First, copyright ensures protection of various types of works, awarding protection to individual data as long as they are original and can be expressed in a material, concrete form. The broad understanding of these protection requirements facilitates extending protection to different types of data.

Second, the long duration of copyright protection secures the possibility for the author to compensate the investment and effort put in the creation of the work.

Third, the copyright holder is granted several exclusive rights that allow controlling the protected work's use and facilitate enforcement in case a third party uses the work without an authorisation. The reproduction, communication to the public and distribution right are indeed a useful toolkit which, balanced by the copyright exceptions, allows for an optimal protection of rightholder's interests.

Finally, copyright law provides for a wide scope of measures securing the rights of the author in case of dissemination of his work and the use of these works by third parties. The rules governing copyright protection aim at enabling further use of the works, securing at the same time the legitimate interests of the author.

**Bottlenecks (disadvantages)**

The most important hindrance resulting from copyright protection is the necessity to obtain authorisation from the copyright holder of each individual data. In the context of big data projects, this may mean identifying authors of hundreds (if not hundreds of thousands) works. In many cases, it might be difficult to identify or find the right holder and/or understand whether he has given his authorisation for use of the work. In practice, this means that time-consuming analyses need to be performed before the data gathered can be used.

Also, even if the originality threshold of works is relatively low, some of the data used in the context of big data projects will not be considered original. It can thus not be assumed that all of the data will benefit from copyright protection. Consequently, there is a need to identify separate measures allowing the use of the non-original data.

Moving to more general characteristics of copyright, it is important to stress that since the legal framework for copyright does not provide for a registration system, the eligibility for protection (and its scope) can only be confirmed *a posteriori* by a court, leading to a lack of legal certainty in the meantime.
Moreover, as regards the possibility to acquire copyright in data, the exclusivity of this type of right constitutes a hindrance, since it does not allow acquiring copyright in the same data "in parallel". The copyright protection foresees for the work to have one author or several co-authors (meaning respectively sole or joint ownership of rights), but excludes the possibility that different entities acquire the same right independently under a different title (e.g., if the data were collected independently or on the basis of different sources). The latter may however often be the case in a big data context, in particular where parties will be independently collecting the same or similar data, leading to the creation of convergent datasets.

Looking from a transactional angle, moral rights of authors can also be seen as a hindrance. Since at least in some Member States there is no possibility to validly assign moral rights, additional measures need to be taken to guarantee that the acquirer of the economic rights is free to use and modify data protected by copyright, to the extent necessary for big data projects.

Finally, the lack of full harmonisation of copyright protection at EU level can also have a chilling effect on EU-wide big data projects, since it requires a separate protection assessment for data originating from different Member States.

Table 3.2 below aims to provide an overview of some of the main advantages and disadvantages of copyright in a big data context.

<table>
<thead>
<tr>
<th>Features enabling protection of data</th>
<th>Features hindering protection of data</th>
</tr>
</thead>
<tbody>
<tr>
<td>✓ Enables protection of all types of data (text, images, sounds), as long as they are original</td>
<td>✓ Only original works can be protected – need for case-by-case evaluation</td>
</tr>
<tr>
<td>✓ Long term of protection</td>
<td>✓ Territorial: different scope of protection in different countries</td>
</tr>
<tr>
<td>✓ Rights of reproduction, communication to the public and distribution</td>
<td>✓ Need to prove that the data was copied and not created by a third person in parallel</td>
</tr>
<tr>
<td>✓ Protection measures available even after the disclosure of data</td>
<td>✓ A posteriori recognition of protection (no registration)</td>
</tr>
<tr>
<td></td>
<td>✓ Exclusivity of rights</td>
</tr>
<tr>
<td></td>
<td>✓ Moral rights</td>
</tr>
</tbody>
</table>

Table 3.2: Copyright: advantages and disadvantages in a big data context
3.2 Database Rights

Apart from individual data, collections of data (databases) are another element important to consider when examining the protection of data, including in a big data context. In this Section, we will therefore discuss the protection awarded to databases.

When considering such protection, a distinction needs to be made between, on the one hand, the database’s contents (individual data), and, on the other hand, its structure and the investment made in its creation. The protection of the latter elements is analysed in the sub-Sections below.

Finally, it shall be borne in mind that computer programs, including those used to obtain, verify, store, present and analyse data, can also be protected by copyright as literary works, as set out i.a. in the Software Directive.\textsuperscript{191} The Directive also guarantees the right to create interoperable products\textsuperscript{192}, which is particularly important in the context of big data projects. Having said that, this report does not aim to discuss the issues related to the protection of computer programs.

3.2.1 Legal Framework

Similarly to copyright, the rules governing database protection have been established at international, regional and national level. While international law provides only some underlying principles for database protection, the actual measures have been harmonised at EU level and implemented into national laws.

**International Legal Framework**

The protection of databases is anchored in international agreements.

Firstly, the Berne Convention explicitly provides in Article 5(2) related to protected works that:

"collections of literary or artistic works such as encyclopaedias and anthologies which, by reason of the selection and arrangement of their contents, constitute intellectual creations,


\textsuperscript{192} According to Article 6 of the Software Directive on decompilation and under conditions specified therein, the authorisation of the right holder is not required where reproduction of the code and translation of its form are indispensable to obtain the information necessary to achieve the interoperability of an independently created computer program with other programs."
shall be protected as such, without prejudice to the copyright in each of the works forming part of such collections”.  

Secondly, the TRIPS Agreement and the World Copyright Treaty extend the database protection to compilations of data or other material which by reason of the selection or arrangement of their contents constitute intellectual creations. Since the wording refers to "compilations of data or other material", protection of databases which do not contain copyrightable elements is also allowed. Consequently, both the TRIPS Agreement and the World Copyright Treaty stipulate that database protection does not extend to the data or the material itself and is without prejudice to any copyright subsisting in the data or material contained in the compilation.

**European Union Legal Framework**

EU law provides for a specific protection of databases, which goes beyond other international legal instruments. The already mentioned Database Directive was adopted with the objective of harmonising the protection of databases in all Member States:

- Its adoption has been driven by the need to secure the investment made in the creation of databases and to create a level playing field between the creators and the makers of databases.
- It applies to both electronic and non-electronic databases, while it however excludes computer programs and moral rights from its scope.
- It establishes in substance a dual system of protection of databases (see sub-Sections 3.2.2 to 3.2.4 below for further details).

---

193 The 1908 Berlin Act of the Convention introduced the protection of collections, which were protected as a category of "derivative works", and were mentioned (in Article 2(2)) along with translations, adaptations, etc. The 1948 Brussels revision conference transferred such protection in a separate paragraph.

194 TRIPS Agreement, art 10(2); World Copyright Treaty, art 5

195 Also note that the WIPO Diplomatic Conference on certain Copyright and Neighboring Rights Questions held in December 1996 had among its document a Basic Proposal for the Substantive Provisions of the Treaty on Intellectual Property in Respect of Databases (available at www.wipo.int/edocs/mdocs/diplconf/en/crnr_dc/crnr_dc_6.pdf) to be considered by the Diplomatic Conference. Although agreement was not reached, the Conference adopted a Recommendation Concerning Databases (available at www.wipo.int/edocs/mdocs/diplconf/en/crnr_dc/crnr_dc_100.pdf).

3.2.2 General Principles of Database Protection in the EU

Definition of Database

The definition of database provided in Article 1(2) of the Database Directive is rather broad\(^{197}\) and provides that a database should be understood as a collection of independent works, data or other materials which are:

- arranged in a systematic or methodical way; and
- individually accessible by electronic or other means.\(^{198}\)

The CJEU further clarified the requirements of independence and individual accessibility providing that the term database, as defined in the Database Directive, refers to any collection of works, data or other materials, that are separable from one another without the value of their contents being affected, including a method or system of some sort for the retrieval of each of its constituent materials.\(^{199}\) Having said that, a database does not need to be created for the purpose of retrieving individual elements of information in order to be protected.

Staying in line with the obligations deriving from international instruments, the database definition encompasses databases that include copyrighted and non-copyrighted elements, and databases in an electronic and non-electronic format.\(^{200}\) Also, a database will be protected by itself (if it fulfils the conditions of protection) without affecting the rights of third parties to the individual pieces of information, which are contained in the database.\(^{201}\)

The definition of "database" provided in the Database Directive is deliberately broad in scope, but it is not open-ended and in some

---

\(^{197}\) According to the CJEU, a wide scope is offered to the concept of "database" for the purposes of the Database Directive (Case C-30/14 Ryanair Ltd. v PR Aviation BV [2015] ECLI:EU:C:2015:10, para 33 (Ryanair)).

\(^{198}\) Database Directive, art 1(2). Also, Recital 17 of the Preamble stipulates that "the term 'database' should be understood to include literary, artistic, musical or other collections of works or collections of other material such as texts, sound, images, numbers, facts, and data (...)."

\(^{199}\) Case C-444/02 Fixtures Marketing Ltd v Organismos prognostikon agonon podosfairou AE [2004] ECLI:EU:C:2004:697, para 32

\(^{200}\) Article 1(1) stipulates that the Database Directive "concerns the legal protection of databases in any form"; according to Recital 15 protection under the Database Directive should be extended to cover non-electronic databases.

\(^{201}\) Article 3(2) of Database Directive provides that "the copyright protection of databases provided for by this Directive shall not extend to their contents and shall be without prejudice to any rights subsisting in those contents themselves".
instances, it has been subject to judicial scrutiny. On the one hand, national courts have excluded from protection random collections of independent data, tourist bus routes, standard contract forms, computer programs used in the operation of a database, a system of indexation for pharmaceutical products, or an algorithm for sport betting and lottery games. On the other hand, national courts recognised protection of inter alia telephone directories, collections of legal material, real estate information websites, radio and television guides, bibliographies, encyclopaedia, address lists, company registries, exhibition catalogues, tourism websites, collections of hyperlinks, and hit parades. In conclusion, to verify the scope of the database definition for a particular dataset, one needs to examine national case law.

**Discussions on Exclusion of Public Sector and Sole-source Databases**

Scholars have discussed the issues around exclusive rights on public sector databases (PSDs) and so-called sole-source databases, notably in view of antitrust issues. Due to their particularities, there are indeed some arguments that would lead to excluding PSDs and sole-source databases from database protection.

The historical evolution of the Database Directive is relevant for that matter. The initial Proposal of the Directive provided for a distinction between (i) databases that collect information from external and various sources; and (ii) "sole-source databases", i.e., when the creator of the database is the only source of the information. The concept of sole-source databases aimed notably at public authority databases. They were subject to a specific regime: licences for the commercial re-exploitation and re-utilisation of the information had to be granted on fair and non-discriminatory terms. That initial distinction was however abandoned in the final version of the Directive. Notwithstanding the abandonment of that proposal, scholars have continued to plead for the

---

202 Maurizio Borghi and Stravoula Karapapa, ‘Contractual Restrictions on Lawful Use of Information: Sole-source Databases Protected by the Back Door?’ (2015) 37(8) EIPR 505
203 *MIDI-Files*, Landgericht München I, Germany, March 30, 2000
204 Court of Appeal of Brussels (8th Ch.), Belgium, June 5, 2007
205 *Schutzfähigkeit von Musterverträger*, Landgericht Stuttgart, Germany, March 6, 2008
206 Haju County Court, Estonia
207 *Symposium Terapeutico*, Court of Appeal of Lisbon, Portugal, December 16, 2008
208 Z.S. v *Sportbetting*, Supreme Court of Croatia, November 17, 2010
exclusion of sole-source databases from the *sui generis* database protection.\footnote{210} This is merely a position advocated by certain legal scholars. To our knowledge, it has not been endorsed at this stage by the CJEU. Having said that, a few first instance national decisions have been issued – yet criticised by scholars – that have denied the existence or the ownership of the *sui generis* right to PSDs.\footnote{211} It is unlikely, however, that these national decisions remain established law in their jurisdictions after the more recent decision of the CJEU in case C-138/11 of 12 July 2012 which concerned the interpretation of Article 102 of the TFEU in relation to the making available of data from the companies register (Firmenbuch) stored in a database. In that decision, the CJEU has indeed not excluded the *sui generis* right on PSDs.\footnote{213}

Although this position has thus far not been embedded in the EU legislation on databases, the approach advocated for may seem interesting in the context of big data projects. Providing more favourable access conditions in case of sole-source databases could be a particularly interesting course of further analysis. It could also be examined whether some of the outstanding access to data issues could be solved by using

\footnote{210}The term "*sui generis* right" is a generic one and means “the right of its own kind".
\footnote{211}Bernt Hugenholtz, 'Abuse of Database Right. Sole-source Information Banks under the EU Database Directive', in François Lévêque and Howard Shelanski (eds), *Antitrust, Patents and Copyright: EU and US Perspectives* (Cheltenham: Edward Elgar, 2005): "Admittedly, such investment will not be difficult to achieve, and producers of sole-source databases will be quick to realize this. However, the Court's strict interpretation of the 'substantial investment' test is important in that it prevents the database right from being abused to convert the natural monopoly of a public service (utility) provider into a near-perfect legal monopoly in derivative information markets.", cited in Maurizio Borghi and Stavroula Karapapa, 'Contractual Restrictions on Lawful Use of Information: Sole-source Databases Protected by the Back Door?' (2015) 37(8) EIPR 505, 509.
\footnote{212}Cristiana Sappa, 'Public Sector Databases - the Contentions between *sui generis* Protection and Re-use' (2011) 17(8) CTLR 217, and case-law cited – i.e.: Dutch High Administrative Court, 29 April, 2009 and Italian Court of Rome, IP Chamber, order June 5, 2008, in AIDA 2010 692ff.
\footnote{213}Case C-138/11 Compass-Datenbank GmbH v. Republik Österreich, [2012] ECLI:EU:C:2012:449, para 47 ("In that regard, it must be held that a public entity which creates a database and which then relies on intellectual property rights, and in particular the abovementioned *sui generis* right, with the aim of protecting the data stored therein, does not act, by reason of that fact alone, as an undertaking. Such an entity is not obliged to authorise free use of the data which it collects and make available to the public. As observed by the Republik Österreich, a public authority may legitimately consider that it is necessary, or even mandatory in the light of provisions of its national law, to prohibit the re-utilisation of data appearing in a database such as that at issue in the main proceedings, so as to respect the interest which companies and other legal entities which make the disclosures required by law have in ensuring that no re-use of the information concerning them is possible beyond that database.").
Ownership of Data

open licences allowing for commercial re-exploitation and re-utilisation of the information on fair and non-discriminatory terms.\textsuperscript{214}

**Types of Protection**

Databases, within the meaning of the Database Directive, are protected in the EU by copyright (Chapter II of the Database Directive, see below), where such copyright protection echoes the one recognised in the international treaties, and a *sui generis* right (Chapter II of the Database Directive, see below). These two rights are independent, and can be applied separately. They will however apply cumulatively if the conditions for both regimes are simultaneously met.

Coming back to the distinction between the protection of individual data (addressed mostly in Sections 3.1 and 3.3 above), the structure of the database and the investment made in the creation of a database, we observe that while the (original) structure of the database is protected by copyright, the investment made in its creation is covered by a *sui generis* right.

<table>
<thead>
<tr>
<th>Individual data</th>
<th>Structure of the database</th>
<th>Investment in the creation of the database</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Civil ownership (?)</td>
<td>• Copyright (InfoSoc Directive)</td>
<td>• <em>Sui generis</em> right (Database Directive)</td>
</tr>
<tr>
<td>• Copyright (InfoSoc Directive)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Trade secrets (Directive 2016/943\textsuperscript{215})</td>
<td>• Copyright (Database Directive and InfoSoc Directive)</td>
<td></td>
</tr>
<tr>
<td>• Contractual arrangements</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 3.3: Databases: multiplicity of protection

### 3.2.3 Copyright Protection for Databases

**Protection Requirements**

According to the Database Directive, copyright protection is granted to databases which, as such, by reason of the selection or arrangement of their contents, constitute the "author's own intellectual creation"; no

\textsuperscript{214} See also in a similar context the discussions related to standard-essential patents and so-called "FRAND" (Fair, Reasonable and Non-Discriminatory) licensing (cf. Osborne Clarke LLP, Legal study on Ownership and Access to Data (European Union bookshop 2016), 25).

other criteria shall be applied to determine the eligibility of databases for that protection.\textsuperscript{216}

The CJEU clarified that the concept of the "author's own intellectual creation" refers to the copyright law criterion of originality (described in sub-Sections 3.1.2 and 3.1.3 above). Applied to databases, the criterion of originality is satisfied when, through the selection or arrangement of the data which it contains, its author expresses his creative ability in an original manner by making free and creative choices and thus stamps his "personal touch".\textsuperscript{217}

More specifically, the following principles apply according to the CJEU guidance:\textsuperscript{218}

- the intellectual effort and skill of creating the underlying data are not relevant in order to assess the eligibility of the database for protection by copyright;
- the originality criterion is not satisfied when the setting up of the database is dictated by technical considerations, rules or constraints which leave no room for creative freedom;
- it is also irrelevant, for that purpose, whether or not the selection or arrangement of the underlying data includes the addition of important significance to that data; and
- the significant labour and skill required for setting up the database cannot as such justify copyright protection if in the end they have not conferred any originality in the selection or arrangement of the underlying data.

Consequently, a database structure may be protected under copyright even if the elements contained therein are in the public domain or are otherwise not protected by copyright.

It also follows from the previous considerations that the originality criterion might be more difficult to fulfil in case of automatically created electronic databases that contain data selected by software, without the actual involvement of an author. In such situations it seems more likely to award copyright protection to the underlying software (algorithm written in a way allowing for selection of specific data/types of data), than to the database itself.

This is particularly relevant in a big data context. Indeed, the development of technology has enabled data analytics of unstructured

\textsuperscript{216} Article 3(1)
\textsuperscript{217} Case C-604/10 Football Dataco Ltd and others v Yahoo! UK Ltd and others [2012] ECLI:EU:C:2012:115, para 38; Case C-5/08 Infopaq International A/S v Danske Dagblades Forening [2009] ECLI:EU:C:2009:465, para 45
\textsuperscript{218} Case C-604/10 Football Dataco Ltd and others v Yahoo! UK Ltd and others [2012] ECLI:EU:C:2012:115, operative part
Ownership of Data

Accordingly, while a protection of datasets is particularly relevant, the protection of the database structure has become less relevant and more difficult when confronted to new types of databases, unforeseen by the Database Directive.

Ownership of Rights

The copyright database protection is generally granted to the creator (author) of the database. More precisely, Article 4 of the Database Directive ('database authorship') provides for the following explicit rules:

- the author of a database shall be the natural person or group of natural persons who created the database or, where the legislation of the Member States so permits, the legal person designated as the right holder by that legislation;
- where collective works are recognised by the legislation of a Member State, the economic rights shall be owned by the person holding the copyright;
- in respect of a database created by a group of natural persons jointly, the exclusive rights shall be owned jointly.

The above principles are generally reflected in the national laws on the copyright database protection in the Member States.

As with traditional copyright, the question is generally much more complex when considering works made within an employment relationship.

Belgium

Belgian copyright law contains a peculiarity with regard to works of employees. In contrast with other ordinary copyright works, the copyright on a database created by employees in the course of their employment contract will directly and exclusively belong to their employer, unless otherwise agreed upon. For databases created in the course of an employment (or service) contract, economic rights will be therefore directly held by the employer. Such presumption is however rebuttable, and concerns only the author's economic rights. It does not concern databases created in the cultural industry. Collective agreements (at the level of the enterprise or at the level of a sector, for instance) may determine the scope and practical arrangements of such presumption.219

By contrast, other Member States apply to databases the same rules as related to copyright in general.

Ownership of Data

France
In France, an employment contract does not have an incidence on the ownership of copyright on a database. The individual author remains the sole owner of his creation. Employers will thus need to conclude specific agreements granting rights except when the database was created under the regime of collective works, pursuant to Article L. 113-2 of the French Code of Intellectual Property.

Germany
Similarly, in Germany, when the database was created as part of the fulfilment of obligations resulting from an employment or service relationship, the provisions of the subsection relating to the allocation of the exploitation rights of the German Copyright Act apply unless otherwise provided in accordance with the terms or nature of the employment or service relationship. If the database is established under a contract to produce a work, it must be ensured through contractual arrangements, that the necessary rights to use the database are granted.

United Kingdom
Finally, the usual rules in the United Kingdom regarding ownership also apply to databases, including the rule that an employer shall be deemed to be the owner of a database created by an employee during the course of his employment, and that a database can be jointly owned where the contribution of each author is indistinguishable from the contribution of co-authors.

Copyright on Database

The prerogatives awarded to the database's author echo the exclusive rights for "general" copyright. They are spelled out in Article 5 of the Database Directive, listing the following so-called 'restricted acts' that require consent of the right holder:

- temporary or permanent reproduction of the database by any means and in any form, in whole or in part;
- any form of distribution to the public of the database or of copies thereof;
- any communication, display or performance to the public;
- translation, adaptation, arrangement and any other alteration of a database, as well as any reproduction, distribution, communication, display or performance to the public of the results of these acts.

Acts that cannot be performed without author's consent
The first sale in the Community of a copy of the database by the right holder or with his consent shall exhaust the right to control resale of that copy within the Community.
The moral rights of the natural person who created the database belong to the author and should be exercised according to the legislation of the Member States and the obligations resulting from the Berne Convention. However, they remain outside of the scope of the Database Directive.\footnote{Recital 28}

**Exceptions**

Exceptions to the exclusive rights of database authors are listed in Article 6 of the Database Directive.

Firstly, the Directive states that the authorisation of the right holder should not be required for a "lawful user" of a database to copy the database, or perform any other 'restricted act' which is necessary for the purposes of access to the contents of the databases and normal use of the contents.

Secondly, in addition to the exceptions traditionally authorised under general copyright law (see sub-Section 3.1.4 above), Member States have the possibility to provide for exceptions in case of:

- reproduction for private purposes of a non-electronic database;
- where there is use for the sole purpose of illustration for teaching or scientific research; and
- where there is use for the purposes of public security or for the purposes of an administrative or judicial procedure.

The three-step test also applies when relying on exceptions under the database legislation\footnote{Database Directive, art 6(3)} (see sub-Section 3.1.4 above).

### 3.2.4 **Sui Generis** Protection of Databases

The second type of protection introduced by the Database Directive is the protection awarded on the basis of a *sui generis* right\footnote{The term "sui generis right" is a generic one and means "the right of its own kind".}, rewarding the substantial investment of the database maker in creating the database. It was developed in order to prevent free-riding on somebody else's investment in creating the database and exists in parallel to the copyright protection on the structure of the database.

The term of the *sui generis* protection is much shorter than that of the copyright protection. It is limited to 15 years as from the first of January of the year following the date of completion of the database. However, such protection may in practice be much longer. According to the

\footnote{Recital 28}
Database Directive, any substantial change to the contents of the database, that could be considered to be a new investment, will cause the term of protection to run anew.\textsuperscript{225}

In practice, should such protection be applied in a big data context, this could result in providing an indefinite protection, given that the databases are usually dynamic, hence, leading in all likelihood to "substantial changes to the contents of the database".

**Protection Requirements**

In order for a database maker to benefit from the \textit{sui generis} protection it shall demonstrate that an investment was made to obtain, verify or present the contents of the database.\textsuperscript{226} The CJEU has had the opportunity to provide guidance on the scope and conditions of the aforementioned terms in several cases (see in particular British Horseracing Board\textsuperscript{227} and Fixtures Marketing I to III cases\textsuperscript{228}, all relating to databases of sport information). Said terms should therefore be interpreted as follows:

- **Obtaining:** "The expression 'investment in...obtaining...of the contents' of a database in Article 7(1) of the directive must be understood to refer to the resources used to seek out existing independent materials and collect them in the database. It does not cover the resources used for the creation of materials which make up the contents of a database. It does not cover the resources used for the creation of materials which make up the contents of a database."\textsuperscript{229}

\textsuperscript{225} Article 10(3) of the Database Directive stipulates indeed that "any substantial change, evaluated qualitatively or quantitatively, to the contents of a database, including any substantial change resulting from the accumulation of successive additions, deletions or alterations, which would result in the database being considered to be a substantial new investment, evaluated qualitatively or quantitatively, shall qualify the database resulting from that investment for its own term of protection".

\textsuperscript{226} We note \textit{en passant} that the French Code of Intellectual Property does not transpose the notion of "obtaining". Article L 341-1 indeed stipulates the following: "Le producteur d'une base de données, entendu comme la personne qui prend l'initiative et le risque des investissements correspondants, bénéficie d'une protection du contenu de la base lorsque la constitution, la vérification ou la présentation de celui-ci atteste d'un investissement financier, matériel ou humain substantiel".

\textsuperscript{227} Case C-203/02 British Horseracing Board Ltd and others v William Hill Organization Ltd [2004] ECLI:EU:C:2004:695


\textsuperscript{229} Case C-203/02 British Horseracing Board Ltd and others v William Hill Organization Ltd [2004] ECLI:EU:C:2004:695, para 42.
• **Verification**: "The expression 'investment in...the...verification...of contents' of a database in Article 7(1) of the directive must be understood to refer to the resources used, with a view to ensuring the reliability of the information, contained in that database, to monitor the accuracy of the materials collected when the database was created and during its operation."\(^{230}\)

• **Presentation**: "The expression 'investment in ... the ... presentation of the contents' of the database concerns, for its part, the resources used for the purpose of giving the database its function of processing information, that is to say those used for the systematic or methodical arrangement of the materials contained in that database and the organisation of their individual accessibility."\(^{231}\)

In addition, such investment needs to be qualitatively and/or quantitatively 'substantial'. It may consist in the deployment of financial resources and/or the expending of time, effort and energy.\(^{232}\)

Because of this obligation for the investment to be made in the creation of the database and not the data as such, fulfilling the protection requirements may become more difficult. This is because the processes of obtaining, verifying and/or presenting the data will happen more and more automatically, as they will be normally conducted with the use of an algorithm. In many cases, it might be true that the investment in creating the raw material exceeds the investment made in segmenting and aligning that pre-existing raw material. In those cases, it might be more difficult to rely on the *sui generis* protection.

That being said, there is no automatic exclusion from *sui generis* protection when the database's creation is linked to the exercise of a principal activity in which the person creating the database is also the one creating the materials that are processed in the database. It is however always the responsibility of that person to demonstrate a substantial investment (qualitative/quantitative, in the obtaining, verification or presentation of the content) independent from the resources used to create these materials.\(^{233}\) In practical terms, this means that companies should consider whether they need to restructure their organisation so that the functional aspects of creating the data and creating the database are kept separate and are separately budgeted.

It is in our view regrettable that the Database Directive, which was drafted in the 90s, does not accommodate for the technical evolution and thus everything that is possible with data and databases today. For instance, it is unclear how techniques of enrichment, partitioning,

\(^{230}\) Case C-203/02 British Horseracing Board Ltd and others v William Hill Organization Ltd [2004] ECLI:EU:C:2004:695, para 42.

\(^{231}\) Case C-444/02 Fixtures Marketing Ltd v Organismos prognostikon aganon podosfairou AE [2004] ECLI:EU:C:2004:697, para 43.

\(^{232}\) Database Directive, Recital 40 of the Preamble

\(^{233}\) Horseracing Board Ltd and others v William Hill Organization Ltd, para 35
harmonisation, homogenisation, etc. of data would fit within the criteria of obtaining, verification or presentation of the database contents. Moreover, the criterion of 'verification' may become less and less pertinent, especially in a big data context which allows analytics of unstructured data. Also, as mentioned in Section Error! Reference source not found. Error! Reference source not found., it shall be kept in mind that several authors refer to the ability of analysing datasets that comprise less accurate data.

Ownership of Rights

The *sui generis* right will benefit the maker (the producer) of a database, i.e., the person who takes the initiative and bears the risk of the investments that are at the origin of the database's creation.

Such approach excludes the possibility to grant *sui generis* protection to subcontractors\(^{234}\): if the work is subcontracted, it is the commissioner of the sub-contracted work that will be granted the *sui generis* protection.

Such narrow view may be particularly problematic in a big data context which includes numerous actors. Indeed, this poses either the issue of assigning rights to a sole actor of the data value cycle, or granting exclusive rights to a multitude of persons.

*Sui Generis* Rights to Database

The maker of a database is granted in substance two (exclusive) economic rights in relation to the *sui generis* protection, i.e. the right to prevent extraction and reutilisation of the whole or of a substantial part, evaluated qualitatively and/or quantitatively, of the contents of that database. The contents of these rights are rather similar to the economic rights of an author in the copyright context (respectively to the reproduction right and to the right of communication to the public).

The extraction is defined as "*the permanent or temporary transfer of all or a substantial part of the contents of a database to another medium by any means or in any form*".\(^{235}\) The CJEU has held that this concept needs to be interpreted broadly, as encompassing any unauthorised act of appropriation (via a physical copy or not) of the whole or part of the contents of a database.\(^{236}\) Neither the purpose of this extraction

\(^{234}\) Database Directive, Recital 41 of the Preamble

\(^{235}\) Database Directive, art 7(2)

\(^{236}\) Horseracing Board Ltd and others v William Hill Organization Ltd, para 51
(commercial or non-commercial) nor the technique of extraction (copying by hand or electronically) is of relevance in this regard.\textsuperscript{237}

The right of reutilisation is defined as "any form of making available to the public all or a substantial part of the contents of a database by the distribution of copies, by renting, by on-line or other forms of transmission."\textsuperscript{238} This means, for instance, that incorporating the data from a database into a catalogue or a website without permission from the right holder amounts to a 'reutilisation' (this resembles the notion of communication to the public under the Infosoc Directive).\textsuperscript{239}

These two exclusive rights are limited to the extraction and reutilisation of 'substantial' parts of databases. In this regard, 'substantial' can mean both qualitatively substantial (a small part of the database that represents a substantial part of the investment\textsuperscript{240}) or quantitatively substantial (a large part of the database).

According to the Database Directive, taking parts of the database that are 'insubstantial' does not amount to an infringement, unless (i) it occurs repeatedly and systematically; and (ii) it conflicts with a normal exploitation of that database or unreasonably prejudices the legitimate interests of the maker of the database.\textsuperscript{241}

\textsuperscript{237} Regarding the notion of "extraction", see particularly Case C-304/07 Directmedia Publishing GmbH v Albert-Ludwigs-Universität Freiburg [2008] ECLI:EU:C:2008:552, para 36: "The decisive criterion in this respect is to be found in the existence of an act of 'transfer' of all or part of the contents of the database concerned to another medium, whether of the same nature as the medium of that database or of a different nature. Such a transfer implies that all or part of the contents of a database is to be found in a medium other than that of the original database".

\textsuperscript{238} Database Directive, art 7(2)

\textsuperscript{239} The CJEU has had the opportunity of clarifying such notion in Case C-203/02 British Horseracing Board Ltd and others v William Hill Organization Ltd; Case C-173/11 Football Dataco and others (Football Dataco II) [2012] ECLI:EU:C:2012:642; and in Case C-202/12 Innoweb BV v Wegener ICT Media BV and others [2013] ECLI:EU:C:2013:850. See particularly Innoweb, para 37: In the light of that purpose, the concept of 're-utilisation' as used in Article 7 of Directive 96/9 must be construed as referring to any act of making available to the public, without the consent of the database maker, the results of his investment, thus depriving him of revenue which should have enabled him to redeem the cost of the investment(...)".

\textsuperscript{240} British Horseracing Board Ltd and others v William Hill Organization Ltd, para 78: "(...) the intrinsic value of the data affected by the act of extraction and/or re-utilisation does not constitute a relevant criterion for assessing whether the part in question is substantial, evaluated qualitatively (...)". In other words, it is the value of the investment which must be taken into account.

\textsuperscript{241} Database Directive, art 7(5)
In any event, if the database maker renders the contents of its database (or a part of it) accessible to the public, its *sui generis* protection does not allow it to prevent third parties from consulting that database.242

**Rights and Obligations of "Lawful Users"**

In the context of the *sui generis* protection, Article 8 of the Database Directive introduces the concept of 'lawful users'. Such users are granted specific privileges (rights); in particular, the database's producer may not prevent them from extracting and/or re-utilising insubstantial parts of the database contents, evaluated qualitatively and/or quantitatively, for any purposes whatsoever (where such authorisation is granted, it only applies to such insubstantial part of the database).

However, a lawful user may not (i) perform acts which conflict with the normal exploitation of the database or unreasonably harm the legitimate interests of the maker of the database; nor (ii) cause prejudice to the holder of a copyright or related right in respect of the works or subject matter contained in the database.

The concept of "lawful users" has been implemented differently across the Member States.

<table>
<thead>
<tr>
<th>Belgium</th>
<th>France</th>
<th>Germany</th>
</tr>
</thead>
<tbody>
<tr>
<td>The law refers to lawful users (&quot;utilisateurs légitimes&quot; – Articles XI.311 et seq. of the Belgian Code of Economic Law).</td>
<td>French law refers to the person who has lawful access (&quot;la personne qui y a licitement accès&quot; – Article L. 342-3 of the French Code of Intellectual Property).</td>
<td>The legislator has deliberately not used the term of lawful user as in the Directive, so that all consumer groups are recognised.243</td>
</tr>
</tbody>
</table>

**Exceptions to the Sui Generis Protection**

The Database Directive proposes exceptions that Member States may transpose under their national laws. These exceptions cover:

242 *British Horseracing Board Ltd and others v William Hill Organization Ltd*, para 55; *Directmedia Publishing GmbH v Albert-Ludwigs-Universität Freiburg*, para 51

243 Bundestag-Drs. 13/7934, 54. Furthermore, according to Article 87e of the German Copyright Act a contractual agreement by which, inter alia, the owner of a lawful copy of the database undertakes vis-à-vis the producer of the database to refrain from reproducing, distributing or communicating to the public quantitatively or qualitatively insubstantial parts of the database shall be ineffective insofar as these acts neither run counter to any normal utilisation of the database nor unreasonably impair the legitimate interests of the producer of the database.
Ownership of Data

- cases of extraction for private purposes of the contents of a non-electronic database
- cases of extraction for the purposes of illustration for teaching or scientific research; as well as
- cases of extraction and/or re-utilisation for the purposes of public security or an administrative or judicial procedure.

The exception related to "scientific research" may prove interesting, to a certain extent only, in the context of big data projects. Under Article 9(b) of the Database Directive:

"Member States may stipulate that lawful users of a database which is made available to the public in whatever manner may, without the authorization of its maker, extract or re-utilize a substantial part of its contents: (…) in the case of extraction for the purposes of illustration for teaching or scientific research, as long as the source is indicated and to the extent justified by the non-commercial purpose to be achieved (…)."

Member States are provided with the possibility of limiting such exception to certain categories of teaching or scientific research institutions. As a result, some Member States have not implemented such exception, and those that have, have done so in diverging ways and notably by providing additional conditions.

<table>
<thead>
<tr>
<th>Belgium</th>
<th>France</th>
<th>Germany</th>
</tr>
</thead>
<tbody>
<tr>
<td>Belgian law requires that the name of the database maker and the title of the database are mentioned, and thus not only the source (Article XI.310 of the Belgian Code of Economic Law).</td>
<td>The legislator has adopted a rather restrictive approach: it excludes from the benefit of the exception some databases and certain use and it limits the list of beneficiaries. Also, the user must pay a compensation. French law also does not specify that the research must be non-commercial.</td>
<td>Article 87(C) of the German Copyright Act refers to &quot;personal scientific use&quot;.</td>
</tr>
</tbody>
</table>

---

244 Database Directive, Recital 51 of the Preamble
245 Databases created for educational purposes and databases created for a digital written edition.
246 "To the exclusion of entertainment or recreational activity".
247 "So far as the public to whom the extraction and the re-utilization are intended is mainly composed of pupils, students, teachers or researchers directly involved".
248 "The use of the extraction [...] is compensated by a remuneration negotiated on a lump sum basis".

Ownership of Data 96 / 142
3.2.5 Possibility to Protect Data under Database Rights

In view of the rules described above it seems that there is very limited to no possibility to secure individual data by means of database protection.

It is true that the *sui generis* protection forbids extraction of all or a substantial part of the database contents to another medium, preventing thus also the copying of the individual data collected in a database. However, once the database maker renders the contents of its database accessible to the public, it cannot prevent third parties from consulting that database. The public is therefore aware of these data (information), and may use them without necessarily having to copy the database contents. Also, the current legal regime seems difficult to reconcile with developments in technologies such as big data or data mining that do not necessarily require data to be reproduced in order to perform analytics or mining processes.

In consequence, the ownership of rights to a database does not confer the rights to the individual data as such. In this context, database protection (both by copyright and the *sui generis protection*) should rather be seen as a complementary measure to protection granted to individual data under other titles such as traditional copyright or trade secret protection (see more in Section 3.3 below).

Having said that, it is important to observe that employing specific technical measures to block access to the database’s content may ensure a *de facto* protection of individual data, preventing the possibility to subject them to data mining or other types of automatic filtering initiated by third parties. The database protections do provide however an incomplete and unsatisfactory protection of the data as such in the event a third party possesses the datasets.

3.2.6 Contractual Aspects

Databases protected by copyright can be transferred or licensed under the same general rules as those described for works in Section 3.1 above. The rights of lawful users should nevertheless be preserved. The latter is particularly important, since according to the Database
Ownership of Data

Directive any provisions that would limit the rights of lawful users would be considered null and void.

As for the *sui generis* right, the Database Directive indicates that such right may be transferred, assigned or granted under contractual licence. Hence, in practice, it will normally be conveyed under a transfer or licence agreement. Also, in case of the *sui generis* protection, the prerogatives given to lawful users cannot be limited, and contractual clauses contrary to the Directive's provisions on the rights of lawful users will be considered null and void.

Finally, the judgment of the CJEU of 15 January 2015 in the *Ryanair v. PR Aviation* case is of significant importance in the context of databases and in particular in relation to contractual terms related to databases.

The facts were in substance as follows. PR Aviation operated a website on which consumers could search through the flight data of low-cost airlines companies, compare prices and, subject to the payment of a commission, book a flight. PR Aviation obtained the necessary data to respond to an individual query by automated means, inter alia, from a dataset linked to the Ryanair website, which was also accessible to consumers.

The use of the Ryanair's website was subject to a "tick-the-box" acceptance of the website terms and conditions. Under said terms and conditions, (i) Ryanair is the exclusive seller of Ryanair flights; (ii) use of Ryanair's website is permitted only for certain defined private and non-commercial purposes; and (iii) use of automated systems or software to extract data from the Ryanair website for commercial purposes is prohibited.

The case ended up before the Dutch Supreme Court, which referred a question to the CJEU as to whether the limits on contractual freedom as set out in Article 15 of the Database Directive apply to databases which are not protected by database copyright or the *sui generis* right.

The CJEU held that where a data set falls within the general definition of a 'database' under Article 1(2) of the Database Directive, but when at the same time the database does not qualify for protection as database

---

249 Article 15
250 Article 7(3)
251 Article 8
252 C-30/14 Ryanair v PR Aviation BV [2015] ECLI:EU:C:2015:10
253 Ibid para 15
254 Article 15 of the Database Directive on the 'Binding nature of certain provisions' provides that "Any contractual provision contrary to Articles 6(1) and 8 shall be null and void."
copyright and/or the sui generis right, the provisions governing database copyright and the sui generis right do not apply to that database.

As a consequence, the provisions of the Database Directive on lawful use (Articles 6(1) and 8) and on the limits to contractual freedom (Article 15) do not apply to such databases. Accordingly, the author/producer of such a database has the right to lay down contractual provisions on the use of the database of its choosing, subject to compliance with any applicable national laws.\textsuperscript{255}

Legal scholars have concluded that “non-protected sole-source database benefit from the full scope of contractual protection”\textsuperscript{256} and that “national laws on unfair competition and contract have the ability to defeat the objectives of the Directive and to upset the efficiency of the internal market”.\textsuperscript{257}

In result, in such circumstance the database authors/makers receive strong protection by having the lawful option to unilaterally exclude third parties from making free use of their databases’ content.\textsuperscript{258} The practical outcome of the interpretation given by the CJEU in the Ryanair judgment is that contractual terms may bring significant limitations to the possibility to use databases and that in certain cases, one may be better off not to benefit from the protections laid down under the Database Directive in order to be able to set strict contractual terms. Indeed, if, and to the extent that, a particular dataset does not qualify for protection as database copyright and/or the sui generis right under the Database Directive, the conclusion of the Ryanair case may have consequences on the contractual terms related to the use of said dataset. If one enjoys no intellectual property rights on its databases, in light of the Ryanair judgment, it would be entitled, at least from an intellectual property rights perspective, to be stricter in its contractual terms and to prohibit any kind of extraction or re-utilisation of its databases, and even of unsubstantial parts of said databases.

3.2.7 Importance for Big Data projects

Database protection presented in the above sub-Sections embeds several features that may be considered useful in the context of big data projects. However, some disadvantages need to be pointed out as well.

\textsuperscript{255}In this context, it shall be noted that the CJEU did not examine the validity or the enforceability of the contractual terms concerned.

\textsuperscript{256}Maurizio Borghi and Stravoula Karapapa, ‘Contractual Restrictions on Lawful Use of Information: Sole-source Databases Protected by the Back Door?’ (2015) 37(8) EIPR 505, 513

\textsuperscript{257}Poorna Mysoor, “Protecting the Unprotected Database” (2015) 131 LQR 556, 561

\textsuperscript{258}Maurizio Borghi and Stravoula Karapapa, ‘Contractual Restrictions on Lawful Use of Information: Sole-source Databases Protected by the Back Door?’ (2015) 37(8) EIPR 505
**Enablers (advantages)**

Firstly, the protection established by the Database Directive is dual, and supplements the possible protection granted to the data as such. In particular, the *sui generis* right provides an interesting protection of the investment made in obtaining, verifying or presenting the contents of the database. However, as demonstrated above, such right may have reached its limits in the current data- and technology-rich landscape.

Another feature that should be considered interesting is the possibility for the *sui generis* protection to run anew in case a substantial change to the contents of a database amounts to a substantial new investment. This rule allows securing reward in the investment made, for instance, in updating or upgrading the data collected. However, as demonstrated above, this may lead in a big data context, to having an unlimited protection in case of dynamic databases.

Finally, the database protection secures interests of the right holder even in situations where the database is made available to the public, and its content disclosed.

**Bottlenecks (disadvantages)**

First, the main disadvantage of relying on database protection in case of big data projects is the absence of protection it awards to individual data.

Second, the eligibility for protection needs to be evaluated on a case-by-case basis in order to verify whether the criterion of originality (copyright protection) or of the substantial investment (*sui generis* Protection) has been fulfilled in case of a specific database.

In addition, and similarly to the traditional copyright protection, the copyright protection for databases entails the necessity to account for the following features:

- the eligibility for protection and its scope can only be confirmed *a posteriori* by a court, leading to lack of legal certainty in the meantime,
- the exclusivity of copyright protection does not allow to acquire a right to the same or similar database, even if the data were collected independently or on the basis of different sources,
- in case the moral rights cannot be transferred, it is necessary to:
  - guarantee that the author will not exercise these rights against the entity who acquired the economic rights, and
  - ensure that the entity who acquired the economic rights is entitled to modify the data and/or database as required.
Also, in case of the *sui generis* protection of databases, the Database Directive prevents only extraction or re-utilisation of "substantial" parts of data, or "insubstantial parts of data" if they are repeatedly and systematically extracted, and not any individual data. Even if the database protection is complemented by the traditional copyright protection, it would still leave the non-original individual data unsecured.

Also, the level of protection ensured across the Member States, especially concerning the copyright to database, is significantly different. This particularly hinders the possibility to manage pan-European projects, since it implies the necessity to examine multiple national legislations in order to have clearance on the possibility to use data, or secure the investment made in a database containing data originating from different territories.

Focusing on the enforcement, in some cases, it might also be difficult to demonstrate that data used by third parties, identical to the ones forming the database content, were actually copied, and not created or collected by this third person in parallel. Without possessing sufficient evidence that the data was actually copied, the database maker may have serious difficulties in preventing the use of its database content.

Table 3.5 below aims to provide an overview of some of the main advantages and disadvantages of database in a big data context.

<table>
<thead>
<tr>
<th>Features enabling protection of data</th>
<th>Features hindering protection of data</th>
</tr>
</thead>
<tbody>
<tr>
<td>✓ Combination of copyright &amp; <em>sui generis</em> right</td>
<td>✓ Individual data not fully protected</td>
</tr>
<tr>
<td>✓ Possibility to extend the term of protection by undertaking substantial effort to change the database’s content</td>
<td>✓ For copyright protection: minimum standard of originality needs to be ensured</td>
</tr>
<tr>
<td>✓ Data can be disclosed</td>
<td>✓ Different scope of protection in different countries</td>
</tr>
<tr>
<td></td>
<td>✓ Need to prove that the data was copied and not created by the third person in parallel</td>
</tr>
<tr>
<td></td>
<td>✓ Case-by-case analysis required</td>
</tr>
<tr>
<td></td>
<td>✓ Right conferred to an &quot;author&quot; (individual)</td>
</tr>
<tr>
<td></td>
<td>✓ Moral rights on the structure protected by copyright</td>
</tr>
<tr>
<td></td>
<td>✓ Territoriality</td>
</tr>
<tr>
<td></td>
<td>✓ A posteriori protection</td>
</tr>
<tr>
<td></td>
<td>✓ Exclusivity of rights</td>
</tr>
</tbody>
</table>

Table 3.5: Database protection: advantages and disadvantages in a big data context
3.3 Trade Secrets and Confidentiality

While the mechanisms described in Sections 3.1 and 3.2 provide measures enabling control over the diffusion and use of works (including data that fulfil the originality criterion) and databases, the objective of trade secret protection is to keep commercially valuable information confidential or secret. Protecting undisclosed know-how and business information enables its creator to transform the effort invested in generating this know-how and information into a competitive advantage.

3.3.1 Legal Framework

Similarly to databases, only general rules requiring protection of trade secrets have been embedded in international law. Specific measures implementing this protection have been laid down at national level across the Member States. The importance of trade secret protection is growing thanks to recent harmonisation efforts undertaken at EU level.

International Legal Framework

The first step towards general recognition of trade secret protection was ensured by the TRIPS Agreement, which introduced a definition of "undisclosed information". Pursuant to that provision, information qualifies for protection if:

- 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;
- it has commercial value because it is secret; and
- it has been subject to reasonable steps under the circumstances, by the person lawfully in control of the information, to keep it secret.

More specific measures complementing this general principle were introduced in the national law of the contracting parties. Traditionally, relevant provisions were introduced within employment law (such as the obligation of an employee to preserve in confidentiality the information relating to the business activity of his employer) and/or legislation providing measures against unfair competition.

Legal Framework in the European Union

Trade secret protection has been established in EU legislation only very recently. In June 2016, the European Parliament and the Council adopted Directive 2016/943 on the protection of undisclosed know-how and business information (trade secrets) against their unlawful acquisition.

259 Article 39
use and disclosure ("Trade Secrets Directive"). The Directive aims to standardise the national laws of the Member States as regards the unlawful acquisition, disclosure and use of trade secrets.

The Directive harmonises the definition of trade secrets in accordance with existing internationally binding standards. It also defines the relevant forms of misappropriation and clarifies that reverse engineering and parallel innovation must be guaranteed (since trade secrets are not a form of exclusive intellectual property right).

The Member States need to implement the Trade Secrets Directive by 9 June 2018.

**National Legal Framework**

Currently, the approach towards the protection of trade secrets strongly differs amongst the Member States. This is best illustrated in Table 3.6 below published as part of the Impact Assessment prepared for the Trade Secrets Directive:

<table>
<thead>
<tr>
<th>Selected measures</th>
<th>Belgium</th>
<th>Germany</th>
<th>Spain</th>
<th>France</th>
<th>Italy</th>
<th>UK</th>
</tr>
</thead>
<tbody>
<tr>
<td>Definition of trade secret in civil law legislation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Availability of injunctions against third party in good faith</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Injunctions not limited in time</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

---


Table 3.6: The fragmentation of the legal protection (selected measures)


Table 3.6 above clearly demonstrates that the current national laws ensuring protection measures for trade secrets differ significantly.

Germany, Italy, and Spain have legislation on the misappropriation of trade secrets, although Germany and Spain fail to define what trade secrets are.

In Belgium, France, and the UK there are no specific provisions on trade secrets in civil law. The protection of trade secrets against misappropriation depends on judicial interpretation of the general provisions on extra-contractual liability or on traditional common law. In Cyprus, trade secrets are only protected by contract. In France misappropriation of certain types of trade secrets (namely, manufacturing secrets) are criminally punished if committed by employees.

This situation is expected to change once the Member States undertake necessary steps to implement the Trade Secrets Directive.
3.3.2 Possibility to Protect Data as Trade Secrets

Since the Member States are already under obligation to adjust their national laws to the protection standards introduced in the Trade Secrets Directive, this sub-Section will further refer to its text in order to describe the scope of trade secret protection and to evaluate the possibility to rely on this protection in case of individual data.

Definition and Scope of Protection

According to the definition provided in the Trade Secrets Directive, a ‘trade secret’ is an information which meets all of the following requirements:

- 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;
- it has commercial value because it is secret;
- it has been subject to reasonable steps under the circumstances, by the person lawfully in control of the information, to keep it secret.\(^{262}\)

Such definition follows the one provided in Article 39 of the TRIPS Agreement. Since the definition is anchored in an international agreement binding all Member States, it is already (to a certain extent) reflected in the national laws.

Trade secrets should be seen as complementary to intellectual property rights. They are heavily used in the creative process leading to innovation and the creation of IPR. Therefore, trade secrets are often at the origin of patents (new inventions), copyright (a new novel or a song), trademarks (a new branded product), and designs (a design of a new car model). Trade secrets are also used in relation to commercially valuable information for which there is no intellectual property rights protection, but for which investment and/or research are nevertheless required and which are important for innovation.\(^{263}\) Moreover, some may prefer to opt for a trade secret protection rather than an intellectual property right, as this may allow them to have an everlasting protection (as long as the conditions for trade secret protection remain fulfilled).

\(^{262}\) Trade Secrets Directive, art 2
Who Owns Trade Secrets?

The Trade Secrets Directive grants protection to the benefit of a trade secret holder, i.e. any natural or legal person who is lawfully controlling a trade secret.

Rights Conferred

As such, a trade secret holder has no private or exclusive rights to its use. Trade secrets are thus different from intellectual property rights that are safeguarded through an exclusive right that is legally enforceable. Instead, the Directive provides rights enabling to oppose the misappropriation of trade secrets.

Consequently, the holder of a trade secret cannot prevent competitors from copying and using the same solutions – reverse engineering (the process of discovering the technological principles of a device, object or system through analysis of its structure, function and operation) is entirely lawful. Trade secrets are only legally protected in instances where someone has obtained the confidential information by illegitimate means (e.g., through spying, theft or bribery).  

Thus, the differentiation between lawful and unlawful acquisition, use and disclosure of trade secrets are at the core of protection ensured by the Trade Secrets Directive.

The acquisition of a trade secret will be considered lawful when the trade secret is obtained by any of the following means:

- independent discovery or creation;
- observation, study, disassembly or testing of a product or object that has been made available to the public or that is lawfully in the possession of the acquirer of the information who is free from any legally valid duty to limit the acquisition of the trade secret;
- exercise of the right of workers or workers' representatives to information and consultation in accordance with the EU law and national laws and practices;
- any other practice conforming with honest commercial practices.

On the other hand, the acquisition of a trade secret without the consent of the trade secret holder will be considered unlawful, whenever carried out by:

---

unauthorised access to, appropriation of, or copying of any documents, objects, materials, 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;

- any other conduct which, under the circumstances, is considered contrary to honest commercial practices.

As to the use or disclosure of a trade secret, it will be considered unlawful whenever carried out, without the consent of the trade secret holder, by a person who:

- acquired the trade secret unlawfully;
- is in breach of a confidentiality agreement or any other duty not to disclose the trade secret;
- is in breach of a contractual or any other duty to limit the use of the trade secret.

The acquisition, use or disclosure of a trade secret shall also be considered unlawful whenever a person knew or ought, under the circumstances, to have known that the trade secret had been obtained directly or indirectly from another person who was using or disclosing the trade secret unlawfully.

In addition, the production, offering or placing on the market of infringing goods, or the importation, export or storage of infringing goods for those purposes, shall also be considered an unlawful use of a trade secret where the person carrying out such activities knew, or ought, under the circumstances, to have known that the trade secret was used unlawfully.

The Directive also harmonises the civil means through which victims of trade secret misappropriation can seek protection, such as:

- stopping the unlawful use and further disclosure of misappropriated trade secrets;
- the removal from the market of goods that have been manufactured on the basis of a trade secret that has been illegally acquired;
- the right to compensation for the damages caused by the unlawful use or disclosure of the misappropriated trade secret.

**Exceptions**

The Trade Secrets Directive introduces four obligatory exceptions to the rights it confers. The Member States need to ensure that an application for the measures, procedures and remedies provided for in this Directive is dismissed where the alleged acquisition, use or disclosure of the trade secret was carried out in any of the following cases:
Ownership of Data

- for exercising the right to freedom of expression and information as set out in the EU Charter of Fundamental Rights, including respect for the freedom and pluralism of the media;
- for revealing misconduct, wrongdoing or illegal activity, provided that the respondent acted for the purpose of protecting the general public interest;
- disclosure by workers to their representatives as part of the legitimate exercise by those representatives of their functions in accordance with EU or national law, provided that such disclosure was necessary for that exercise;
- for the purpose of protecting a legitimate interest recognised by EU or national law.

Data Protected as Trade Secrets

The protection established for trade secrets will expand to every piece of information, as long as it fulfils the protection requirements (mentioned above).

It is important to note that data can be protected as trade secrets as long as they remain secret. Once the dataset is published, or disclosed in any other way, the protection can no longer be claimed. This is particularly relevant in a big data context, as it follows from the foregoing that data used for big data analytics, and made publicly available will not qualify as trade secrets. Therefore, when considering to outsource big data analytics, any company should carefully assess whether its datasets comprise trade secrets that are valuable to the company and which cannot be disclosed for that reason.

3.3.3 Confidentiality: Contractual Aspects

The transfer of confidential information should be contractually regulated. Specific provisions imposing obligations on the acquiring party to ensure that the information will remain confidential, should be stipulated.

Confidentiality of data is important not only as an integral part of trade secret protection, but also as a security measure used at different stages of cooperation between entities.

Since trade secret protection is based on securing confidentiality of information, it is required for the information to be subject to reasonable steps under the circumstances.

Such reasonable steps can be based on the following technical measures backed by legal and policy measures:
<table>
<thead>
<tr>
<th>Technical measures</th>
<th>Legal and policy measures</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Software limiting access to data, blocking the possibility to copy, send or otherwise disseminate;</td>
<td>• Establishing in-house rules on data access and management;</td>
</tr>
<tr>
<td>• Login &amp; password to the computer used by an employee/other contractor;</td>
<td>• Introducing policy on data disclosure to third parties;</td>
</tr>
<tr>
<td>• Requirement of using access cards to enter the building/floor/data room.</td>
<td>• Including confidentiality clauses in the employment and cooperation contracts;</td>
</tr>
<tr>
<td></td>
<td>• Concluding non-disclosure agreements (&quot;NDA&quot;) for the purpose of business negotiation, at an early phase of cooperation (if the parties engage in pre-contractual discussions, confidentiality clauses play an essential role in allowing for an open dialogue and give a measure of security in the event that the parties do not reach a final agreement).</td>
</tr>
</tbody>
</table>

Table 3.7: Reasonable measures to secure confidentiality of information

In any instance where a company or an individual needs to disclose information to its/his partner (in business relations, research collaboration etc.), it needs to make sure that the information will not be further transmitted. It is therefore necessary to put in place appropriate technical and legal measures to ensure a sufficient level of protection. Confidentiality therefore plays an important role at different points of the contracting process and throughout the life of any big data service.

Safeguard procedures should be implemented and complied with, and the necessary filters should be applied in order to avoid that data the publication of which would potentially be harmful to the commercial interest of any party concerned, is made publicly available.

This can however pose several practical issues in the framework of big data projects, as the several actors of the data value cycle should be compelled to keep confidential the valuable data used. This would lead to a multitude of agreements with different actors and, most likely, with different terms and conditions. Such myriad of contracts would render the practical implementation of any big data project extremely burdensome and risky for the preservation of trade secrets.

A popular measure to ensure that an appropriate level of confidentiality is guaranteed is by concluding NDAs and including confidentiality clauses in the employment, cooperation and other types of similar agreements.

An NDA is a legal contract between at least two parties that outlines confidential material, knowledge, or information that the parties wish to
share with one another for certain purposes, but wish to restrict access to or by third parties. It is a contract through which the parties agree not to disclose information covered by the agreement. An NDA creates a confidential relationship between the parties to protect any type of confidential and proprietary information or trade secrets. As such, an NDA protects non-public business information.

A confidentiality clause is a contractual provision imposing the obligation not to disclose certain information that a party acquires while performing its tasks under the contract. It is usually secured by means of contractual penalties.

3.3.4 Importance for Big Data projects

**Enablers (advantages)**

The most interesting feature of trade secret protection in view of big data projects is that it allows for protection of individual pieces of information regardless of their originality. It also does not differentiate between the types of data that might be protected. Moreover, the protection is unlimited in time, as long as the information has not been disclosed.

It is also interesting that the definition of a trade secret holder refers to a person lawfully controlling a trade secret. Similarly, to avoid legally complicated discussions on the ownership of data, one could refer to a person lawfully controlling data.

**Bottlenecks (disadvantages)**

Although the specific type of protection established for trade secrets seems interesting to be used in view of the protection of data for the reasons explained above, it has one fundamental disadvantage – it requires the data to remain secret.

It does not exclude the possibility to use the system. It seems that at least in some jurisdictions it is possible to rely on confidentiality agreements to ensure that the requirement of secrecy of the data under the Trade Secrets Directive is maintained even after the transfer of data has been exercised. This is however yet to be confirmed by the courts.

Also, it may be difficult to demonstrate that an individual data has commercial value because it is secret. Many data will be considered valuable only if they are part of a bigger dataset.

Table 3.8 below aims to provide an overview of some of the main advantages and disadvantages of trade secrets in a big data context.
### Table 3.8: Trade secrets: advantages and disadvantages in a big data context

<table>
<thead>
<tr>
<th>Features enabling protection of data</th>
<th>Features hindering protection of data</th>
</tr>
</thead>
<tbody>
<tr>
<td>✓ Protection for individual data</td>
<td>- Protection disappears once the data are disclosed</td>
</tr>
<tr>
<td>✓ Protection for all types of data</td>
<td>- Difficulty in demonstrating that an individual data has commercial value because it is secret</td>
</tr>
<tr>
<td>✓ Protection unlimited in time (as long as not disclosed)</td>
<td>- The more parties are concerned, the more NDAs are needed</td>
</tr>
<tr>
<td>✓ Definition of trade secret holder</td>
<td></td>
</tr>
</tbody>
</table>
Chapter 4

Contractual Arrangements

4.1 Unsatisfactory Legal Framework

As becomes apparent from the previous Chapters, the existing legal framework in the EU is not optimal and does not sufficiently facilitate operations on or including data. More specifically, it can be concluded that the current legal regime is unsatisfactory due to several cumulative factors, such as the lack of data ownership, the possible multitude of rights but incomplete means to protect data, but also the fragmented national legal frameworks.

Moreover, the cumulative implementation of this maze of different possibly applicable legislations is a significant hurdle to the uptake of data analytics in the EU and is creating legal uncertainty in this fast-growing market.

Lack of Data Ownership: Multitude of Rights and Incomplete Protection

As examined in Chapters 2 and 3, there is currently no EU legislation that specifically regulates the ownership in data. The existing framework allows the protection of limited types of data under legislations the main purposes of which were not to protect raw data.

Indeed, the copyright, database and trade secrets legislations are not fit-for-purpose as they aim to protect literary and artistic works, the structure of databases, the investment made in databases or secret information. Similarly, the sole existing legal framework that aims to specifically protect data in the EU relates to "personal data". This being said, as discussed in sub-Section 2.3.1 above, such legislation does not explicitly confer on individuals an "ownership" right in their personal
data. It is limited to the protection of one's privacy by recognising rights over such information and giving a high level of control to individuals.

Accordingly, any data that is not covered by the protection available under the existing framework can hardly be protected against unauthorised use of such data by third parties.

This issue is not purely theoretical. For instance, in the TOREADOR project, it can reasonably be assumed that the big data analytics processes will be performed on all kinds of data, which may be protected under overlapping, yet unsatisfactory, systems – as depicted in the below diagram. However, it is expected that a very large amount of data will not benefit from the existing statutory protections.

![Figure 4.1: Overview of possible types of data used in the TOREADOR project](image)

What's more, the rights one may possibly rely on are not fully harmonised. As the applicable rules differ across the Member States, even if the data are protected, they are not protected in exactly the
same way, since for example thresholds to obtain protection may differ, as well as the available remedies. Given that one needs to rely on a patchwork of rights, it might be difficult to define the scope of protection, present/evidence the rights, but also to license or transfer such rights.

Such fragmented legal framework is particularly relevant when confronted with emerging technologies, including big data and cloud computing, as the aggregated, analysed, stored and otherwise used data may be flowing through different Member States, and may thus be subject to diverging protections.

In consequence, the many actors involved in the data value cycle do not have the necessary guarantees and protection over the data they collect, process, share, analyse or generally use. Hence, the investment made in collecting, generating or processing the data is not fully secured.

**Conclusion**

As a result of the above-depicted unsatisfactory legal framework, those involved in the data value cycle may currently hold back on data sharing initiatives. Also, they presently have no choice but to rely on contractual arrangements to manage their rights in data. Indeed, "in contrast to other intangibles, data typically involve more complex assignments of different rights across different data stakeholders, requiring of some stakeholders, 'the ability to access, create, modify, package, derive benefit from, sell or remove data, but also the right to assign these access privileges to others'". 265

Given the current reliance on contracts, it is important to examine whether contractual arrangements provide an efficient legal framework for managing rights attached to data, including on exclusivity, exchange, exploitation or access to data.

In the following Section, we examine the possible difficulties related to the reliance on contracts when considering data.

---

4.2 Data-related Contractual Arrangements

Under the current state of law, when considering the exploitation of data, the various stakeholders active in the data value cycle have no choice but to rely on contractual arrangements.

It can be expected that the reliance on contracts to regulate the use of data will increase in the future given the current legal framework, but also due to the greater adoption of big data analytics tools. Also, the interpretation given by the CJEU in the Ryanair judgment discussed in sub-Section 3.2.6 above may convince some persons to rely on contracts rather than on exclusive database rights conferred under the Database Directive. Indeed, according to such decision, one may be better off not to benefit from the protections laid down under the Database Directive in order to be able to set strict contractual terms, and restrict certain uses, including to "lawful users".

Although the current situation may seem to provide greater flexibility to the contracting parties, it nevertheless comes with various difficulties, including a great legal uncertainty on the market. Some of the most important identified difficulties are detailed below.

4.2.1 Lack of Harmonisation of Contract Law

The EU has taken several initiatives in the past to harmonise certain aspects of contract law, or at least to attempt increasing such harmonisation. This being said, the current regime is a patchwork of 28 national systems. Such situation presents a significant barrier to cross-border trade, but also leads to legal uncertainty, unnecessary costs and high complexity. Such situation is particularly true in a business-to-business (B2B) context where the EU has taken less action than in situations where consumers (seen as a weak party) are involved.

The differences in the rules governing contracts concern numerous key aspects; such as how an agreement is concluded and terminated, the applicable liability regime and the available remedies in the various Member States, but also some key principles applicable before, during and after the contractual relationship. The differences between common law and civil law countries that coexist in the EU demonstrate the current lack of harmonisation – as depicted in the table below. There are however also numerous discrepancies between countries applying the same legal system (civil or common law).
Ownership of Data

<table>
<thead>
<tr>
<th>Statutory regime governing B2B contracts</th>
<th>Is there a legally binding principle of ‘good faith’?</th>
<th>Contract drafting style</th>
</tr>
</thead>
<tbody>
<tr>
<td>Civil law (in Europe)</td>
<td>Generally yes, governing</td>
<td>Shorter, relying on default framework in the Civil Code (But English law approach is becoming common in larger B2B contracts)</td>
</tr>
<tr>
<td></td>
<td>• Pre-contractual phase</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Performance of obligations</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Post-contractual phase</td>
<td></td>
</tr>
<tr>
<td>English law</td>
<td>Generally not at the pre-contractual phase</td>
<td>Longer, detailed style of prescriptive drafting for commercial B2B agreements of all types</td>
</tr>
<tr>
<td></td>
<td>• The principle is steadily developing during contract performance phase</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Generally not at the termination phase</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Unfair Contract Terms Act 1977 on exclusion clauses</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Third Party Rights Act 1999 on third party rights</td>
<td></td>
</tr>
</tbody>
</table>

Table 4.1: High-level overview of key differences between civil and common law contracts

Despite the EU Commission's efforts to provide clarity and present initiatives, especially in certain sectors such as sales law, insurance contract law or cloud computing, it follows from the foregoing that the current situation does not provide the necessary level of harmonisation.

4.2.2 Limits of Contractual Arrangements

Both intellectual property rights and contractual arrangements present key characteristics that are particularly relevant:

<table>
<thead>
<tr>
<th>Intellectual property rights</th>
<th>Contractual arrangements</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Erga omnes</em>(^{266}) rights: refers to rights or obligations that are owed toward all. A (intellectual) property right is an <em>erga omnes</em> entitlement, and therefore enforceable against anybody infringing that right.</td>
<td>Principle of privity of contract: relates to the relationship that exists between the parties, whereby only the parties may enforce the contract. Third parties cannot derive rights from, nor have obligations imposed on them by someone else's contract.</td>
</tr>
</tbody>
</table>

Table 4.2: Overview of the difference between *erga omnes* rights and the principle of privity of contract

It follows that while *erga omnes* rights can be enforced against any infringer, rights based on contracts are only enforceable against the contracting party.

\(^{266}\) Latin: towards all
Such finding is predominantly problematic in the data economy, especially where the industrial data is not protected by intellectual property rights and thus only subject to contractual arrangements. Accordingly, a licensor may be in a position where it has limited to no recourses against a third party that has gained access to its industrial data.

In such situation, the various stakeholders of the data value cycle are compelled to carefully map – and anticipate – their data flows and put in place all necessary contractual arrangements with all those involved in the data processing. This also requires integrating complex back-to-back warranty clauses in the numerous contracts as regards the upstream source of data and the downstream uses of such data.

### 4.2.3 Content of Data-related Agreements

Given the current importance of contractual arrangements in a data context – for lack of a better legal framework –, it is crucial to carefully consider the content of such contracts.

One of the important aspects to consider is the inclusion of an ownership clause in a particular agreement. Indeed, in many instances, a party may attempt to stipulate that it owns the data which is at the heart of the contract. In our view, in the event the data is not protected by intellectual property rights, such claim may be problematic. Legally speaking, as there currently is no recognition of ownership rights in data, the enforceability of any clause claiming such ownership is questionable and requires an in-depth assessment of national civil law principles and case law.

In the same vein, it may be inadvisable for a party to contractually claim that the data is protected by intellectual property rights and/or by confidentiality agreements. Such type of provision may not sufficiently portray the actual situation of the licensed industrial data and could thus be challenged by the other party (e.g., if the data at stake is not protected under copyright) and/or by third parties to the contract. Similarly, limiting the contractual stipulation to database rights may not sufficiently regulate the aspects related to the data as such.

It follows from the foregoing that parties are strongly recommended to determine in detail the rights and obligations of each party. This notably requires detailing the following elements:

- The permitted processing acts
- The prohibited processing acts
• The purposes of the permitted processing acts (including for commercial and non-commercial uses, which ought to be detailed and defined)
• The data quality
• The possible intellectual property at stake (if any)
• The faith of any "database" rights (if any)
• The faith in case of merging / co-mingling of datasets
• The faith of the results of the data analysis
• The faith of derivative works of data mining or analysis processes
• The enhancement, modification or other revisions of (parts of) the datasets
• The communication of (parts of) datasets (including results) to third parties
• The possible co-ownership issues on datasets, derivative works and mining/analysis results
• The anonymisation and/or pseudonymisation obligations (including possibly the (software) licence on the key)
• The required technical and organisational security measures, including the access restrictions on the datasets
• etc.

In addition, the parties will need to agree on more classical provisions related to the duration, liability, warranties, indemnification, fees/royalties, privacy, audit rights, etc.

Detailing the above elements may prove to be a difficult exercise. Indeed, the parties do not benefit from a statutory framework and must therefore anticipate numerous elements in order to compensate the legislative deficiencies. This will often lead to lengthy verbose agreements, which are costly to negotiate as well as time consuming. Also, it requires anticipating future processing acts but also foreseeing possible issues that may arise during and after the contractual relationship.

4.2.4 Validity of Data-related Agreements

As already mentioned in the previous paragraphs, there remains uncertainty on the validity of clauses claiming ownership in data. Hence, in the event an agreement explicitly claims or even assumes ownership in data, the validity of such clause is possibly challengeable.

Moreover, even in the absence of any ownership claim, the mere fact that the parties seek to provide for a licence on data is challengeable. Indeed, from a legalistic point of view, one may wonder to what extent it is permitted to give a licence on something of which the ownership is not explicitly recognized under civil law. Given that the current legal framework does not recognise any ownership right in data (other than
possible intellectual property rights), there may be a lack of legal basis to allow for a valid and enforceable licence agreement.

The legal validity of the many existing agreements related to data will need to be assessed by courts, on a case-by-case basis. However, as demonstrated above, there is a relatively high risk of seeing diverging views across the EU due to the lack of harmonisation of civil contract law.

The high legal uncertainty deriving from the current situation affects the entire data value chain and the whole data flow. Such situation is not sustainable in a data-driven economy and with the fast-increasing development and adoption of data mining and analysis tools.
Chapter 5

Moving Forward on Data Ownership

It has been amply demonstrated in the previous Chapters that the current legal framework relating to data ownership is not satisfactory.

No specific ownership right subsists in data and the existing data-related rights do not respond sufficiently or adequately to the needs of the actors in the data value cycle. Up until today, the only imaginable solution is capturing the possible relationships between the various actors in contractual arrangements.

Nevertheless, as emphasised in Chapter 4, filling the data ownership gap with contractual arrangements is far from ideal. It would be practically burdensome – and probably even impossible – to regulate with full legal certainty by means of contracts the ownership issues in large-scale data undertakings (such as the TOREADOR project) where there is a multitude of data sources, storages, analyses and thus a myriad of actors who would want to claim ownership in the data concerned. On top of all that, comes the issue where contracts are in principle nonbinding, and therefore unenforceable, vis-à-vis third parties.

Against a background where the EU strives towards a data-driven environment in which both citizens and companies can reap the benefits of novel data technologies, but also against a background where the current legal framework does not sufficiently tackle all the issues related to data and where actors involved in the data value chain have no certainty as to the ownership of the data they have gathered, created, analysed, enriched or otherwise processed; a more solid and legally secure solution is needed.
We believe such solid solution consists in the creation of a non-exclusive, flexible and extensible ownership right in data(sets), with a data traceability obligation as a safeguard ("garde-fou"). The present Chapter shall therefore discuss the specificities of said right and obligation, their interaction with the other existing rights in data, their incidence on civil law, and their possible reflection in contractual arrangements.

5.1 Creation of a Non-exclusive Ownership Right in Data

Our analysis of the existing legal framework related to data as well as our observation that contractual arrangements do not satisfactorily cover the issue of ownership have led us to conclude that a new non-exclusive ownership right in data should be created to respond to the EU data economy’s demands.

Such new right would have the following specific characteristics, which will be further detailed in the sub-Sections below:

- the right would be non-exclusive;
- it would be an ownership-type of right as opposed to an intellectual property-type of right; and
- as such, it would be a right in individual pieces of data that will naturally extend to the entire datasets, which those individual pieces of data are part of.

Non-exclusive Character of the New Right

Data is as such non-rivalrous, non-exclusive and inexhaustible. An unlimited number of actors in the data value cycle can act with regard to the same data, without interfering with each other’s actions. Also, as soon as data becomes public, its use by others cannot be limited and the data cannot be exhausted (although often rapidly outdated). Data therefore has a potential illimitable availability, which should be reflected in the new right.

Accordingly, it would be impractical, and even senseless, to create an exclusive right in data, as this would significantly limit the possibilities of the actors in the data value cycle and thus hamper the emergence of a true EU data economy. A non-exclusive right, apart from fitting with the non-exclusive nature in essence of data itself, would allow for a shared use of data by the different actors in the data value cycle, each on their own merits.
Such shared use of the same data would stimulate competition in the data economy.

**Ownership Right vs. Intellectual Property Right**

The new right would be sculpted as a type of ownership right, with the term "ownership" having the same legal meaning as put forward in sub-Section 2.2.1 of this report, namely "the exclusive right to use, possess, and dispose of property, subject only to the rights of persons having a superior interest" and to any restrictions on the owner's rights imposed by agreement with or by act of third parties, or by operation of law, however eliminating the exclusive character of the right considering the particularities of data discussed in the previous sub-Section.

The creation of an ownership-type of right is preferred over an intellectual property right given that the latter is granted as a reward for some kind of intellectual effort, which (depending on the type of intellectual property right) is measured in terms of originality, novelty, distinctiveness or another similar criterion. Consequently, the protection of intellectual property rights is accorded to only a limited amount of "information" – as demonstrated by Figure 5.1 below.

---

267 Such as, for example, IPR holders (see however Section 5.3 below on Data Protected by Copyright)


269 Rob Kitchin, *The Data Revolution: Big Data, Open Data, Data Infrastructures and their Consequences* (SAGE 2014)
An effective functioning of the EU data economy however requires protecting other types of data that do not necessarily flow from intellectual efforts, but which merit protection as a valuable asset of the data value cycle.

Moreover, granting an intellectual property right for the protection of data would lead to an overly exclusive and protective legal framework, which would not fit with the particular characteristics of data, as explained in the previous sub-Section. Although it offers legal certainty, an intellectual property right in data would not provide the flexibility needed to perform big data analytics.

A non-exclusive ownership right would therefore, in our opinion, better suit the needs of actors in the data value cycle.

**Right in Individual Pieces of Data**

The new right would create a non-exclusive entitlement to individual pieces of data. As a result of the ever-expanding data economy we are currently witnessing, a single piece of data may already be vital to its processor\(^{270}\) and should therefore be protected.

---

\(^{270}\) Processor should be read in the broad meaning given to the term "data processing" by ISO/IEC 2382-1, revised by ISO/IEC 2382:2015 – Information technology – Vocabulary (see below).
For example, in the TOREADOR use-case on the Energy Production Data Analysis, sensors in the solar farms and smart homes generate valuable pieces of information for internal functioning and maintenance of such solar farms and smart homes.

In such example, each and every piece of information has its value in the broader bulk of data and should therefore be recognised as such by attaching to each a (non-exclusive) ownership right.

In order to determine at which point a piece of information qualifies as data and is therefore eligible for protection by the non-exclusive ownership right, we reiterate the ISO/IEC definition of data: "Reinterpretable representation of information in a formalized manner suitable for communication, interpretation, or processing."\(^{271}\)

Therefore, as soon as information is represented in a formalised manner suitable for communication, interpretation, or processing; it would qualify for protection under the new right.

In order for an actor in the data value cycle to be able to claim said non-exclusive right in data, it should in some way "process" the data. In this respect, it shall be noted that the same ISO/IEC standard provides the following definition for the concept of data processing: "systematic performance of operations upon data".\(^{272}\)

Such broad definition of processing can also be found in the GDPR, where processing is defined as "any operation or set of operations which is performed on personal data or on sets of personal data, whether or not by automated means, such as collection, recording, organisation, structuring, storage, adaptation or alteration, retrieval, consultation, use, disclosure by transmission, dissemination or otherwise making available, alignment or combination, restriction, erasure or destruction".\(^{273}\)

The non-exclusive ownership right in data could therefore be claimed by actors performing systematic operations upon it (such as collecting, recording, organising, structuring, enriching, harmonising, etc.). A natural consequence of such right would be the elasticity of the protection on the entire datasets, which those individual pieces of data are part of.

\(^{271}\) ISO/IEC 2382-1, revised by ISO/IEC 2382:2015 – Information technology – Vocabulary; please note that the same ISO/IEC standard defines 'data communication' as the "transfer of data among functional units according to sets of rules governing data transmission and the coordination of the exchange".

\(^{272}\) Ibid

\(^{273}\) GDPR, art 4(2)
5.2 Creation of a Traceability Obligation

The creation of a new non-exclusive ownership right in data may lead to having several players of the data value cycle claim ownership in all or part of the same data without being able to demonstrate that they have actually "processed" the data. Such situation would be undesirable as we do not aim for protection to be awarded to unprocessed data. Also, the situations where one can claim protection without being able to evidence the processing, should be prevented. We therefore believe that a safeguard ("garde-fou") for the non-exclusive ownership right in data should be provided in the form of a traceability obligation in order to prevent abuse and provide legal certainty.

The traceability obligation would consist in the obligation to be able to demonstrate at all times the provenance of and the processing performed on the data one is claiming ownership in. Such obligation could be accomplished by keeping traceability logs. The player concerned would have to keep logs that are updated every time the data(set) is processed. Each time (part of) the data(set) is transferred, the traceability logs file (or the relevant part thereof) should be transferred with it.

The purpose of the traceability obligation would be threefold:

- Traceability would feature as a validity requirement for the ownership. As such, the traceability logs would provide one with an ownership title in the data they relate to.
- The traceability logs file would be a key element to establish the legitimate use of data in order to show in particular that the dataset concerned has not been created in violation of rights of third parties.
- The transfer of traceability logs together with the data would enable the mapping of data flows, which would in turn allow complying with other existing legal obligations; e.g. the mapping of data flows would, for example, facilitate the data controller to respond to the data subject's request of access to his/her personal data. It would also allow for a swift and proper compliance with the obligation of notification of data breaches or security incidents under the GDPR and the NIS Directive respectively.

Unlike in the offline world, a traceability obligation in the form of log files would not be too burdensome in the information and technology environment that surrounds (big) data analytics.

Moreover, the exact modalities of the traceability obligation could be further clarified in soft law measures, such as codes of conduct or certification mechanisms. Similarly to the GDPR, a particular statute
could be awarded to such soft law measures, the establishment of which should be encouraged by the EU Member States.\textsuperscript{274}

Hence, the non-exclusive ownership right would not create a blanket freedom for its holders to do whatever they wish with the data concerned, but would be both elaborated and confined by the introduction of a traceability obligation.

5.3 Interaction with Other Rights in Data

In Chapters 2 and 3, the present report has provided an in-depth analysis of other rights that may subsist in data, be it individual pieces of data or collections of data.

Should the proposed non-exclusive ownership right come into existence, it is essential to consider its interaction with such other rights in data. To this end, Figure 5.2 below displays the possible types of data used in the framework of the TOREADOR project. The following sub-Sections aim to discuss per type of right how each would interact with the non-exclusive ownership right.

\textsuperscript{274} See Articles 40 and 42 of the GDPR
Database Protection

The use of data covered by the non-exclusive ownership right may be opposed on the basis of an exclusive *sui generis* database right. The traceability obligation, discussed in Section 5.2 would however allow resolving such issue, as it would enable the demonstration of the data's independent origin, similar to the concept of "independent creation" existing in numerous jurisdictions.

The doctrine of independent creation is used *inter alia* under US, Canadian and Japanese laws as a defence to a claim of copyright infringement. Indeed, given that copyright protection is not subject to registration, it is practically impossible to be aware of all the copyright protected works existing in the world. The principle of independent creation has been confirmed by the highest US courts as well as in the US legal doctrine.²⁷⁵ Closer to home, we observe the widespread

recognition of the concept in Belgian case law and legal doctrine. Though not (yet) unanimously accepted in the copyright laws of each Member State, we deem the defence of independent creation indispensable in a creative society where technology opens more doors than one could ever imagine were closed.

Claims of exclusive database rights conflicting with the use of data under the non-exclusive ownership right could thus be refuted by relying on a defence of "independent origin", which could be evidenced by the traceability logs file held under the traceability obligation.

**Data Protected by Copyright**

As highlighted in sub-Section 3.1.4 above, we believe that the infringement test in copyright should be rethought against a data analysis background. When an in se protected work is used for (big) data analytics, it is usually not used with the same aim as the purpose for which the work was created; i.e. the work is not exploited or, in other words, it is not used as a "work".

Such use should, in our view, not be considered as an infringement provided that it conforms to the main conditions of the three-step test laid down by Article 9(2) of the Berne Convention and Article 5(5) of the InfoSoc Directive. The vast majority of (big) data analytics does not conflict with the normal exploitation of the work nor unreasonably prejudices the legitimate interests of the rightholder.

Therefore, data used in the context of (big) data analysis and benefitting from the non-exclusive ownership right should not be exposed to copyright infringement claims, provided that the copyright-protected data is not used as a "work".

**Data Protected by Trade Secrets**

In case one would like to use data protected by trade secrets for the purpose of (big) data analytics, the protection of said data vis-à-vis the new ownership right would rely – as it does under the current legal framework – on confidentiality obligations laid down in contractual arrangements. Hence, no specific legal change would be needed in this respect.

Furthermore, the traceability logs file held under the abovementioned traceability obligation may also serve as a means to refute unwarranted

---

et al. v. NBC Universal, Inc., et al. California Court of Appeal, Second Appellate District, April 28, 2009; Gerald Morawski v. Lightstorm Entertainment, Inc. et al. No. 13-55227 (9th Cir. 2015).
allegations of misappropriation of trade secrets, as they would allow demonstrating a different provenance of the data concerned.

**Data Protected by Privacy and Data Protection Law**

Rather than recognising an exclusive ownership right of the data subject in its own personal data, we suggest the coexistence of our proposed non-exclusive ownership right and the rights of data subjects under data protection law. In our view, an exclusive ownership right of the data subject in personal data, such as – allegedly – recently adopted in France under the Digital Republic Act (please refer to sub-Section 2.3.1), would be meaningless as the GDPR already provides ample control to data subjects with respect to their personal data. Therefore, in our opinion, personal data can be owned by data controllers or processors within the limits imposed by the GDPR.

In this context, it shall be noted that the obligations under data protection law remain in any event unaffected in the framework of the new non-exclusive ownership right. Thus, all actors in the data value cycle would still need to rely on a legitimate ground for processing of personal data and data subjects would remain able to exercise their rights under the GDPR. The obligations of data security (including the strong push for pseudonymisation and encryption) would also remain intact.

Moreover, as demonstrated above, the new non-exclusive ownership right and the corresponding traceability obligation would facilitate complying with existing legal obligations under the GDPR – such as the right of access to personal data and the breach notification obligation.

### 5.4 Incidence of the New Right on Civil Law

In order to achieve legal certainty, the non-exclusive ownership right should be secured from a civil law perspective, either by adapting the civil code's existing ownership concept or by creating a new one. Indeed, given the inherent exclusive character of the traditional ownership concept, the non-exclusive ownership right over intangible assets should somehow be recognised under civil law.

Such civil law concept of non-exclusive ownership over intangible assets should clarify the rules related to the ownership as well as the title over ownership, which is to be established on the basis of traceability logs (see above). The traceability obligation should therefore also be inscribed in the civil code. The new ownership concept should also determine the evidentiary value of the traceability logs file, in order to strengthen their utility.
In addition, the new civil law should prohibit any contractual claim of exclusivity of an ownership right in data, as this would defeat the purpose of the new non-exclusive ownership right. Any provision to such effect should therefore be held to be null and void in the new civil law.

In this respect, Commissioner Oettinger published in October 2016 an article in the French newspaper Le Monde, in which he encouraged the creation of a "civil code" for data in the following words:

"[...] I call for a 'civil code' on data. Just like the civil code clearly governs, among other things, the right of ownership, possession and lease of movable and immovable property, we need a civil code for the digital age that clarifies in detail the issues that may arise with respect to rights in data" (freely translated from French).

Although it may prove burdensome to introduce a comprehensive civil code at EU level, we believe a regulation of the civil law concept of data ownership at EU level is achievable. Possible soft law measures (such as codes of conduct) could further frame and supplement the non-exclusive ownership right as well as the traceability obligation.

As already mentioned, the new non-exclusive ownership right would also be in line with EU competition law, given that access to data(sets) will be stimulated and competition will thus be enhanced.

5.5 Reflecting the New Right in Contractual Arrangements

Though the creation of a non-exclusive ownership right and the subsequent adaptations in civil law would provide the much desired legal certainty that is currently lacking, contractual arrangements between the different actors in the data value cycle would still be important to further detail the exercise of the ownership right in practice.

Thus, contractual arrangements should detail each actor's role, the exact scope of the "ownership" per actor, the permitted and prohibited processing acts, as well as the other elements enumerated in sub-Section 4.2.3. In this regard, as mentioned above, such contractual arrangement could not impose the exclusivity of an ownership right in data.

276 Günter H. Oettinger, 'Big data: Pour un 'code civil' des données numériques' (Le Monde 14 October 2016)
In order to give weight to the contractual relationship, the traceability logs file should be annexed to the contract. Given that the traceability logs would serve as the ownership title, they would be opposable to third parties (once these principles would have been implemented in the Member State laws). Although a contractual arrangement is as such not opposable to third parties, the traceability logs and thus the ownership title could be interpreted in light of the contract.

Therefore, despite the higher legal certainty created through the introduction of the ownership right, contractual arrangements still have a crucial role to play. As such, contracts would be one of the building blocks of the new system, which could be further elaborated by the creation of template or standard agreements. The latter could be developed by industry experts at EU level, similarly to what has been done in the context of cloud computing.\textsuperscript{277}

Such solution, where the new ownership principles would be set in stone but contracts would determine the further implementation, would offer legal certainty whilst providing the necessary flexibility.

\textsuperscript{277} The Expert Group on Cloud Computing Contracts was set up on 18 June 2013 to assist the EU Commission in identifying safe and fair contracts terms and conditions for cloud computing services for consumers and small firms. For more information, see <http://ec.europa.eu/justice/contract/cloud-computing/expert-group/index_en.htm>
Annex 1. Overview of the TOREADOR use-cases

Use-case #1. SAP – Application Log Analysis pilot

The Application Log Analysis pilot, which is based on a product called ‘SAP Enterprise Threat Detection' ("ETD"), has as its general purpose to detect security breaches in the systems of its users (i.e., companies). The ETD solution gathers and analyses log data, such as user logs, security audit logs or transaction logs. It cannot be excluded that part of the data collected in the context of such solution is "personal data". Such collected data is normalised and stored into a database for analysis, notably based on patterns detected by the ETD solution.

Use-case #2. Lightsource – Energy Production Data Analysis pilot

The Energy Production Data Analysis pilot is based on an asset management platform called the "Lightsource Monitoring Platform", which aims to provide information on the operation of the Lightsource solar farms and smart homes in order to improve the functioning and the maintenance of those farms/smart homes. To this end, a considerable
Ownership of Data

volume and variety of data is collected from the solar farms and smart homes and subsequently analysed.

The categories of data collected from the solar farms include both data inherent to the solar installations as such (energy-related data) and data related to the installations' surroundings (ambient data); for example:

- Energy-related data: active and reactive energy; active and reactive power; voltage, current and frequency levels; daily energy produced; total energy produced; string combiner details; etc.

- Ambient data: irradiance data; ambient temperature data; photovoltaic module temperature data; wind speed/direction data; humidity data; etc.

The categories of data collected from the smart homes include, among others: generated power/energy; consumed power/energy; export-import power/energy; frequency and voltage levels; etc.

![Data Flow Diagram]

*Figure 6.2: Schematic and simplified data flow diagram of the Energy Production Data Analysis pilot*
Use-case #3. JOT – Clickstream Analysis pilot

The ultimate goal of the Clickstream Analysis pilot is to generate tools and models needed to tackle and fight online fraud. It therefore aims to create a model to detect fraud which would make it possible to determine in real-time to what extent a user is fraudulent. In order to do so, the pilot is collecting and processing numerous data. More particularly, web traffic data is matched and cross-checked with marketing campaign performance data provided by the pilot Owner. This data encompasses the whole spectrum of information that can be collected by means of a pixel on an Internet webpage (including for instance: IP addresses, search terms, real URL, cookies, etc.) and on a Search Engine Marketing Platform (including for instance: campaign category, number of impressions and clicks, performance metrics, etc.).

Figure 6.3: Schematic and simplified data flow diagram of the Clickstream Analysis pilot

Use-case #4. DTA / Avio Aero – Aerospace Products Manufacturing Analysis pilot

The Aerospace Products Manufacturing Analysis pilot aims to develop a "smart factory" concept by subjecting its production and quality data to
the TOREADOR Platform's big data analytics. The analytics will then be used to optimise, in real-time, the performance of the production process. To this end, an on-going analysis is required of variables liable to affect the quality of the manufacturing process or of the resulting products. The practical implementation of the Aerospace Products Manufacturing Analysis pilot will be carried out by Avio Aero, an Italian manufacturer in the aerospace sector.

The following internal and external variables may affect the quality of the manufacturing process or of the products and may thus be analysed in the framework of the pilot:

- Internal variables: properties of materials, design data, machine output, quality data, warehouse data, etc.
- External variables: temperature, pressure, humidity, etc.

Figure 6.4: Schematic and simplified data flow diagram of the Aerospace Products Manufacturing Analysis pilot
[14] Osborne Clarke, Legal study on Owernship and Access to Data (European Union bookshop 2016)
[22] Commission, 'Towards a Thriving Data-Driven Economy' (Communication) COM(2014) 442 final
[38] Hoeren T, 'Big Data and the Ownership in Data: Recent Developments in Europe' (2014) 36(12) EIPR 751
[40] Hugenholtz B, 'Abuse of Database Right. Sole-source Information Banks under the EU Database Directive', in François Lévêque and Howard Shelanski (eds), *Antitrust, Patents and
Ownership of Data
[58] Oettinger G H, 'Big data: Pour un 'code civil' des données numériques' (Le Monde 14 October 2016)
[61] Purtova N, 'Property Rights in Personal Data' (BOXPress 2011)
[62] Purtova N, 'The Illusion of Personal Data as No One’s Property' (2015) 7(1) Law, Innovation and Technology 83
[65] Sappa C, 'Public Sector Databases - the Contentions between _sui generis_ Protection and Re-use' (2011) 17(8) CTLR 217
## Glossary

<table>
<thead>
<tr>
<th>Term</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Article 29 Working Party</strong></td>
<td>The Article 29 Working Party is an independent advisory body on privacy and data protection composed of representatives from the national data protection authorities, the European Data Protection Supervisor and the EU Commission</td>
</tr>
<tr>
<td><strong>BDA</strong></td>
<td>Big Data Analytics</td>
</tr>
<tr>
<td><strong>Charter</strong></td>
<td>EU Charter of Fundamental Rights</td>
</tr>
<tr>
<td><strong>CJEU</strong></td>
<td>Court of Justice of the European Union</td>
</tr>
<tr>
<td><strong>CSIRT</strong></td>
<td>Computer Security Incident Response Team</td>
</tr>
<tr>
<td><strong>DG Connect</strong></td>
<td>Directorate General for Communications Networks, Content and Technology of the European Commission</td>
</tr>
<tr>
<td><strong>DSA</strong></td>
<td>Data Sharing Agreement</td>
</tr>
<tr>
<td>Acronym</td>
<td>Definition</td>
</tr>
<tr>
<td>---------</td>
<td>------------</td>
</tr>
<tr>
<td>DSM</td>
<td>Digital Single Market</td>
</tr>
<tr>
<td>ECHA</td>
<td>European Chemicals Agency</td>
</tr>
<tr>
<td>ERP</td>
<td>Enterprise Resource Planning</td>
</tr>
<tr>
<td>ETD</td>
<td>Enterprise Threat Detection</td>
</tr>
<tr>
<td>EU</td>
<td>European Union</td>
</tr>
<tr>
<td>ECHR</td>
<td>European Court of Human Rights</td>
</tr>
<tr>
<td>F/OSS</td>
<td>Free and Open Source Software</td>
</tr>
<tr>
<td>GDPR</td>
<td>General Data Protection Regulation</td>
</tr>
<tr>
<td>IEC</td>
<td>International Electrotechnical Commission</td>
</tr>
<tr>
<td>IPR</td>
<td>Intellectual Property Right</td>
</tr>
<tr>
<td>ISO</td>
<td>International Standards Organisation</td>
</tr>
<tr>
<td>ISP</td>
<td>Internet Service Provider</td>
</tr>
<tr>
<td>MPL 2.0</td>
<td>Mozilla Public Licence V2.0</td>
</tr>
<tr>
<td>NCA</td>
<td>National Competent Authority</td>
</tr>
<tr>
<td>NDA</td>
<td>Non-Disclosure Agreements</td>
</tr>
</tbody>
</table>
### NIS Directive

### PECS providers
Public Electronic Communication Service Providers

### PIA
Privacy Impact Assessment

### PSD
Public Sector Database

### PSI Directive

### Radio Equipment Directive

### RFID
Radio-Frequency Identification

### ROI
Return On Investment

### SLA
Service Level Agreement

### SME
Small and Medium-sized Enterprise

### sui generis
The term "sui generis" is a generic one and means “of its own kind”

### TFEU
Treaty on the Functioning of the European Union

### Trade Secrets Directive

### TRIPS
Trade-Related Aspects of Intellectual Property Rights