# Big Data Regulatory Debates in the EU

FAYE FANGFEI WANG[*]

## Abstract

In the phenomenon of big data intertwined with cloud services, the traditional way of collecting, processing and storing data has been changing. This new phenomenon challenges traditional legal principles and increases legal uncertainty of various rights protection in the information society. Such challenges have become greater and greater in response to the emerging technologies embedded with machine intelligence. This paper seeks to remove legal obstacles in the era of big data by defining big data from a legal perspective and finding solutions to the key legal challenges of big data such as data privacy and security; intellectual property protection and jurisdictional issues. It proposes a feasible interpretation and application of the traditional concepts to the new big data phenomenon. Overall, it intends to contribute to the review and reform of the current regulatory framework of the data-driven society in the EU and promote legal certainty and a healthy economic and societal growth across the globe.

## 1.  Introduction

In the information society, computing storage becomes larger and cheaper over time. This enables datasets to be gathered together from open sources or other databases on a mass scale, which is greater than the traditional significant volume of data in banks or other individual organisations. Smarter statistical and computational analysis to these aggregated datasets from different sources may generate new value, meaning and context.[1] In addition, artificial intelligence enables machines to replicate humans making decisions. Although machines are not as clever as the people that make them, machines have the potential to continuously learn new datasets.[2] Taking the advantage of data from various sources, a more sophisticated robot can just learn like a child.[3] When such robot has access to a database of facial expressions, it may also have the ability to learn to recognise the expression of other faces and can grad-

---

[*]  Senior Lecturer in Law, Brunel University London, Email: fangfei.wang@gmail.com.

[1]  A Williamson, *Big Data and the Implications for Government* 14 Legal Information Management 253-257, p 253 (2014).

[2]  The Life Scientific, *Nigel Shadbolt Broadcasts*. (14 Apr 2015) at 09:00. BBC Radio 4; (14 Apr 2015) 21:30.

[3]  'The robot which learns like a child', BBC News, 21 October 2015, available at <http://www.bbc.co.uk/news/technology-34446447> (last accessed 6 October 2016).

ually react to the different humans engaging with him in a different way.[4] All of these generate a new phenomenon known as 'open data' and 'big data'.

In 2016, there were a number of exciting global new tech developments which benefited from databases, open data and big data. For example, Google's self-driving cars continued their evolution, racking up over 1.5 million miles in the United States. Another competitor - Tesla Motors has brought up new products including an auto-pilot function to allow its car to drive semi-autonomously, with little or no driver input on motorways. Whilst the system is undoubtedly advanced, it requires the driver to be ready to take control of the car at a moment's notice, leading to questions of liability in the event of an accident. The volume, variety and velocity of data gathered to be able to employ such technology keeps us curious as to the legal implication of big data. Another new development is a new smartphone game called Pokemon Go, which was released in July 2016, utilising GPS and augmented reality to allow the user to explore a virtual world while walking around the physical world. The game instantly brought up media reports of privacy and safety issues – firstly with the game requiring access to the users' Google account credentials, then with concerns of player's location being reported to others.[5]

Artificial intelligence also continues its advancement. If robots with access to big data become so intelligent and closely connected to human beings, they may potentially cause other social and legal consequences. For example, there is a growing concern that whether human beings would be able to live in harmony with robots. It was reported that in Japan people 'grow so attached to their robot dogs that they hold funerals for them when they "die".'[6] The phenomenon of big data in combination of artificial intelligence has also expanded into the field of legal services. It was reported that 'by combining machine learning, data science and legal expertise, some law firms are starting to offer the kind of predictive analytics services that were previously only available outside the law in fields such as financial services'.[7] With the further development of technologies making use of big data, in intelligent online dispute resolution (ODR) systems, robotic arbitrators may be able to make accurate automated decisions based on various computing technologies such as service-oriented computing and expert systems.[8] In the far future, the relationship between robots and human beings, i.e. marriages and other contractual issues may raise further legal concerns.

In the phenomenon of big data intertwined with cloud services, the traditional way of collecting, processing and storing data has been changing. This new phenomenon

---

[4] R Cellan-Jones, 'My day with a robot', BBC News, 15 September 2015, available at <http://www.bbc.co.uk/news/technology-34256655> (last accessed 6 October 2016).

[5] F Wang, 'Introductory Remark', Society of Legal Scholars Annual Conference, University of Oxford (8 September 2016).

[6] J Wakefield, 'Intelligent machines: Will we accept robot revolution?' (7 October 2015), available at <http://www.bbc.co.uk/news/technology-32334571 (last accessed 6 October 2016).

[7] 'The Relentless Advance of the Super-intelligent Attorney', Financial Times, 6 December 2016, available at <https://www.ft.com/content/af3e2a64-a069-11e6-891e-abe238dee8e2> (last accessed 20 April 2017).

[8] F Wang, Online Arbitration (Oxford: Informa, 2017).

challenges the traditional legal principles and increases legal uncertainty in the information society. Such challenges have become greater and greater in response to the emerging technologies embedded with machine intelligence. The purpose of this article is to identify key legal obstacles of big data in order to propose a regulatory strategy overall. This paper looks into the first obstacle – the concepts of 'open data' and 'big data'; and discusses the legal challenges and regulatory development for the current phenomenon in general. This paper in particular seeks to resolve the obstacle – how big data is defined as there is no consistent understanding across the globe. Thirdly, this paper discusses the legal obstacles to intertwining the concept of big data with other existing regulations such as data privacy and security; intellectual property protection and jurisdictional issues. It proposes a feasible interpretation and application of the traditional concepts to the new big data phenomenon. Overall, it intends to contribute to the review and reform of the current regulatory framework of the data-driven society in the EU and seeks for possible solutions to promote legal certainty and a healthy economic and societal growth across the globe.

In order to establish a feasible legal framework for the phenomenon of 'big data', various key legal issues should be addressed and sought for interpretation. They are:

– How should 'big data' be defined from a legal perspective?
– Under which laws should 'big data' be governed? Is it database law, privacy law, data protection law, intellectual property law or others? When disputes occur, how is jurisdiction determined?
– If there is no current legislation providing provisions on big data, in what direction should the future legislation be heading? Is it possible to interpret and apply traditional legal principles to big data issues in the digital economy? Or is it feasible to foster innovation, protect rights holders, and promote the free flow of information by reforming existing law, developing specific policies and guidance for companies and consumers, or encouraging self-regulation?

Although the European Commission has not yet fully looked into all of the main issues above, slow progression has been made since 2012. For example, in 2012 the Commission proposed a major reform of the EU legal framework on the protection of personal data in order to strengthen individual rights and tackle the challenges of globalisation and new technologies.[9]

In 2014 the Commission continued working on a reform package towards a thriving data-driven economy, which aimed to build a single, modern, strong, consistent and comprehensive data protection framework for the EU. This is in line with a growing recognition for the need of a regulatory environment that strengthens individuals' trust and confidence and enhances legal certainty for the development of innovative and sustainable data goods and services such as 'big data'. The Commission stipulated that 'the fundamental right to personal data protection applies to big data where it is

---

[9] Reform of Data Protection Legislation, available at <http://www.ec.europa.eu/justice/data-protection/> (last accessed 6 October 2016).

personal: data processing has to comply with all applicable data protection rules.'[10] In addition, there are a series of legal measures suggested in this reform package to build trust in order to exploit the full potential of the data-driven economy:

– horizontal consumer and marketing law also applies to products based on big data technology. The Commission will ensure that small and medium-sized enterprises (SMEs) and consumers, suppliers and users, are given all necessary information, are not misled, can rely on fair contracts, notably as regards the use of data collected from them; and
– the Commission will also work with Member States and stakeholders to ensure that businesses, and in particular SMEs, receive adequate guidance, notably on issues such as data anonymisation and pseudonymisation, data minimisation, personal data risk analysis, and tools and initiatives enhancing consumer awareness. The Commission will also actively support research and innovation (R&I) for related technical solutions that are privacy enhancing 'by design'.[11]

Meanwhile, the European Council also recognised the interplay between cloud computing and big data by calling for EU action to provide the right framework conditions for a single market for Big Data and Cloud Computing in its conclusions of October 2013.[12] This echoed the view of the Commission, which launched a consultation process on the concept of user-controlled cloud-based technologies for storage and use of personal data ('personal data spaces') in the phenomenon of big data.[13] On 10 January 2017 the Commission adopted the 'Building the European Data Economy' package including a Communication and a Staff Working Document,[14] which looks into key legal issues concerning:

– the rules and regulations impeding the free flow of data and present options to remove unjustified or disproportionate data location restrictions, and
– access to and transfer of data, data portability and liability of non-personal, machine-generated digital data.[15]

---

[10] *Towards a Thriving Data-Driven Economy*, COM (2014) 442 final, Brussels, 2.7.2014, p 11.

[11] *Ibid.*

[12] Commission Staff Working Document – Report on the Implementation of the Communication, 'Unleashing the Potential of Cloud Computing in Europe', Accompanying the document Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions 'Towards a thriving data-driven economy', SWD(2014) 214 final, Brussels, 2.7.2014, p 2.

[13] *Towards a Thriving Data-Driven Economy*, COM (2014) 442 final, Brussels, 2.7.2014, p 11.

[14] Commission Staff Working Document on the free flow of data and emerging issues of the European data economy: Accompanying the document Communication Building a European data economy, {COM(2017) 9 final}, Brussels, 10.1.2017, SWD(2017) 2 final.

[15] *Building a European Data Economy*, available at <https://ec.europa.eu/digital-single-market/en/building-european-data-economy> (last accessed 20 April 2017).

In order to collect feedback on the above issues to help shape the future policy agenda, the public consultation was launched between 10 January 2017 and 26 April 2017 to seek opinions on the following matters:

– whether and how local or national data localisation restrictions inhibit the free flow of data in Europe;
– whether and to what extent digital non-personal machine generated data are traded and exchanged;
– the nature and magnitude of any barriers to accessing such data;
– ways of tackling those barriers;
– emerging Internet of Things and robotics liability challenges; and
– practices and issues relating to data portability, interoperability and standards.[16]

## 2.   Solutions to the Obstacles of Defining Big Data

### 2.1.   Open Data

The relationship between open data and big data is that open data may provide a substantial amount of sources for big data. The term 'open data' refers to 'a subset of data, namely to data made freely available for re-use to everyone for both commercial and non-commercial purposes'.[17] It also includes the re-use of public sector information – 'the wide range of information that public sector bodies collect, produce, reproduce and disseminate in many areas of activity while accomplishing their Public Task'.[18]

In response to current barriers of public sector information re-usage, in 2013 the Directive on Re-use of Public Sector Information was amended in order to remove barriers to the re-use of public sector information across the European Union.[19] Member States were required to implement the revised Directive no later than 18 July 2015. In the UK, work has been progressing. For example, in November 2011, the Chancellor of the Exchequer, George Osborne, presented his Autumn Statement to Parliament, which acknowledged that making more public sector information available would

---

[16] *Public Consultation on Building the European Data Economy*, 10 January 2017 to 26 April 2017, available at <https://ec.europa.eu/eusurvey/runner/European-Economy-Consultation#> (last accessed 20 April 2017).

[17] Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions: Towards a thriving data-driven economy, COM (2014) 442 final, Brussels, 2.7.2014 (hereafter 'Towards a thriving data-driven economy'), available at <http://www.eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:52014DC0442&from=EN> (last accessed 6 October 2016), p 5.

[18] The glossary of terms, by the UK Advisory Panel on Public Sector Information, available at <http://www.data.gov.uk/glossary> (last accessed 6 October 2016).

[19] Directive 2013/37/EU of the European Parliament and of the Council of 26 June 2013 amending Directive 2003/98/EC on the re-use of public sector information Text with EEA relevance, OJ L 175, 27.6.2013, p 1-8.

help catalyse new markets and innovative products and services as well as improving standards and transparency in public services. Taking into consideration the benefits of the usage of open data, the UK Government intended to open up access to core public datasets on transport, weather and health, including giving individuals access to their online GP records. It also promised to provide up to £10 million over five years to establish an Open Data Institute to help industry exploit the opportunities created through release of this data.[20] In addition to EU legislation on open data, the Commission also promotes non-legislative measures to facilitate re-use of public sector information.[21] Main measures include engagement with Public Sector Information expert group (PSI Group); funding support to an Open Data incubator and the Legal Aspects of Public Sector Information (LAPSI); and development of an Open Data Portal.[22] The EU Open Data Portal and one-stop-shop open data digital service is to provide open data access across the EU.

Although member states, such as the UK, consider open data to be of great value to society, there is still a general concern over data privacy protection, in particular health data, in practice. This prompts the consideration of the balance of trade-off between openness and privacy as they are not mutually exclusive.[23] In order to strike the balance, data owners may need to understand their best interests to make informed decisions, while data providers may employ license restriction of using datasets. Accordingly, the European Commission has established the EU open data policy and legal framework; and prepared guidelines on recommended standard licences, datasets and charging for the re-use of documents to facilitate the implementation.[24] In addition, other legislative and technological measures are also in place to further open up data for access and re-use. For example, there are measures which promote scientific discovery and collaboration across disciplinary and geographical boundaries in the Commission's scientific information package. Moreover, a number of Commission initiatives covering sector-specific data (transport, environment, etc.) as well as

---

[20] Further Detail on Open Data Measures in the Autumn Statement 2011, 29 November 2011, available at <https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/61959/Further_detail_on_Open_Data_measures_in_the_Autumn_Statement_2011.pdf> (last accessed 6 October 2016).

[21] Non-legislative measures to facilitate re-use, *Open Data, EU Digital Single Market*, available at <https://ec.europa.eu/digital-single-market/en/non-legislative-measures-facilitate-reuse> (last accessed 20 April 2017).

[22] *Ibid*. See also *Open Data Portals*, available at <https://ec.europa.eu/digital-single-market/en/open-data-portals> (last accessed 20 April 2017). Open Data portals are web-based interfaces designed to make it easier to find re-usable information. See also *EU Open Data Portal*, available at <http://www.data.europa.eu/euodp/en/data/> (last accessed 20 April 2017). The EU Open Data Portal has been in operation since December 2012.

[23] Life Scientific, *Nigel Shadbolt Broadcasts*. (14 Apr 2015) at 09:00. BBC Radio 4; (14 Apr 2015) 21:30.

[24] *Towards a Thriving Data-Driven Economy*, COM (2014) 442 final, Brussels, 2.7.2014, available at <http://www.eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:52014DC0442&from=EN> (last accessed 6 October 2016), p 8.

Guidelines on Open Access to Scientific Publications and Research Data in Horizon 2020[25] have also been introduced to foster open data policies.[26]

## 2.2.   Extra-large Datasets: Big Data

Parallel to the regulatory development for open data, a new regulatory framework for extra-large data has also been called for. When extra-large data from a substantial amount of open data and other sources is generated, it generates a new phenomenon called 'big data'. It is suggested that 'big data refers to large amounts of data produced very quickly by a high number of diverse sources'.[27] Such data can either be created by people or generated by machines.[28] Extra-large data from various sources at a high speed does not just let us see more of the same data, but also allows us to see new, better and different for enhanced insight and decision making. For example, when sufficient data is collected in a service-oriented computing system for car parking services, drivers' credit cards will be automatically charged for the parking fee without any human interaction as the automated system will immediately identify where those drivers are and what drivers are doing. Just like the anti-theft device in cars developed in Tokyo, the car could recognise that a non-approved driver was behind the wheel, and the engine would just stop.[29] Moreover, the emerging technology of driverless cars is also supported by big data analytics.[30] In the US, big data is also used to help FBI fraud crackdown by identifying, tracking and prosecuting criminal activity in the Medicare system.[31] This prompts an urgent need for big data management. It was reported that there is 'a tenfold increase in demand for big data staff in the past five years, with vacancies rising from 1,800 in 2008 to 21,400 in 2013 – an average annual increase of 212 per cent.'[32] This also urges for regulatory updates to

---

[25] *Guidelines on Open Access to Scientific Publications and Research Data in Horizon 2020*, 11 December 2013, available at <http://www.ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-pilot-guide_en.pdf> (last accessed 6 October 2016).

[26] *Towards a Thriving Data-Driven Economy*, COM (2014) 442 final, Brussels, 2.7.2014, available at <http://www.eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:52014DC0442&from=EN> (last accessed 6 October 2016), p 8.

[27] *Big Data, EU Digital Single Market*, available at <https://ec.europa.eu/digital-single-market/en/big-data> (last accessed 20 April 2017).

[28] *Ibid.*

[29] 'To Prevent Theft: Car Seat Identifies Drivers Sitting Down', TechCrunch 21 December 2011, available at <http://www.techcrunch.com/2011/12/21/car-seat-japan/> (last accessed 6 October 2016).

[30] 'What are the policy challenges of driverless cars?', Policy Exchange 20 August 2015, available at <http://www.policyexchange.org.uk/media-centre/blogs/category/item/what-are-the-policy-challenges-of-driverless-cars> (last accessed 6 October 2016).

[31] 'Big Data help FBI fraud crackdown', 12 January 2015, Financial Times, available at <http://www.video.ft.com/3980039209001/Big-Data-help-FBI-fraud-crackdown/World> (last accessed 6 October 2016).

[32] 'Big Data Analytics Assessment of Demand for Labour and Skills 2013-2020', e-Skills UK, October 2014, available at <https://www.e-skills.com/Documents/Research/General/BigData_report_Nov14.pdf> (last accessed 6 October 2016), p 4.

be considered because the concept of 'big data' challenges the traditional legal principles. There are three main issues as follows:

– when such data comes from different sources, it raises the issue of the ownership, usage rights and copyright;
– when such data comes from different individuals, it raises the issue of privacy and data protection; and
– when such data comes from different jurisdictions, it raises the issue of restriction and harmonised standards of cross-border data transfers.

It is clear that all of the legal issues above have been covered by relevant legislation (regardless of their sufficiency) when governments, businesses and individuals are dealing with data before the name of 'big data' appears. It is debated that it is not having 'big data' that makes it different, but it is what people do with it that matters.[33] However, this does not mean that 'big data' is the same as 'open data', or should be treated the same as traditional datasets. 'Big data' might be 'closed and proprietary, not necessarily to the public at large'[34] from different sources which require further consideration for regulations.

In response to the market trend on the emerging technologies supported by big data analytics, the EU Commission recognises that 'this global trend holds enormous potential in various fields, ranging from health, food security, climate and resource efficiency to energy, intelligent transport systems and smart cities, which Europe cannot afford to miss.'[35] Subsequently, the EU has launched communications towards a thriving data-driven economy, which is to set out some operational conclusions to support and accelerate the transition to the establishment of the right framework conditions for a single market for big data and cloud computing.[36] The key tasks include 'making sure that the relevant legal framework and 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; and rapidly concluding the legislative processes on the reform of the EU data protection framework, network and information security, and support exchange and cooperation between the relevant enforcement authorities (e.g. for data protection, consumer protection and network security)'.[37]

---

[33] A Williamson, *Big Data and the Implications for Government* 14 Legal Information Management 253-257, p 253 (2014).

[34] Williamson, *supra* n 24, at 254.

[35] Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions: Towards a thriving data-driven economy, Brussels, 2.7.2014, COM(2014) 442 final, available at <https://ec.europa.eu/digital-agenda/en/towards-thriving-data-driven-economy> (last accessed 6 October 2016).

[36] *Ibid.*

[37] *Ibid.*

## 2.3.   The Definition of Big Data

In 2012 International Data Corporation (IDC) defined 'Big Data technologies' as 'a new generation of technologies and architectures designed to extract value economically from very large volumes of a wide variety of data by enabling high-velocity capture, discovery, and/or analysis.'[38] In industry, the concept of big data has been further interpreted from a technical point of view, which 'describes a holistic information management strategy that includes and integrates many new types of data and data management alongside traditional data. While many of the techniques to process and analyse these data types have existed for some time, it has been the massive proliferation of data and the lower cost computing models that have encouraged broader adoption.'[39]

As the name of 'big data' indicates, big data is something 'big' in size and it is 'data' in nature. Big data are data that are unprecedented in scale and scope.[40] Big data has been described as a phenomenon rather than a technology.[41] It is often described in terms of the 'three Vs': volume, variety and velocity.[42] An IT company, Oracle, has recently provided an architecture overview on big data, suggesting that one additional V – value – should be considered as an attribute of big data.[43] In some

---

[38] D Vesset and others, *Market Analysis: Worldwide Big Data Technology and Services 2012 – 2015 Forecast* (IDC, March 2012) (last accessed 6 October 2016).

[39] B Gaff, HE Sussman, and J Geetter, *Privacy and Big Data*, Computing and the Law (IEEE June 2014) at p 7-9, 8; and see also P Heller, D Piziak and K Knudsen, *An Enterprise Architecture White Paper – An Enterprise Architect's Guide to Big Data  –  Reference Architecture Overview* (May 2015), available at <http://www.oracle.com/technetwork/topics/entarch/articles/oea-big-data-guide-1522052.pdf> (last accessed 6 October 2016), p 4.

[40] R Schroeder and E Meyer 'Big Data: What's new?', Oxford Internet Institute, 9 October 2012, available at <http://www.ipp.oii.ox.ac.uk/sites/ipp/files/documents/Schroeder%20presentation.pdf> (last accessed 6 October 2016).

[41] Lesley Wiggins, 'If big data and analytics exist in a silo, does the outcome matter?' IBM Big Data and Analytics Hub, 25 February 2014, available at <http://www.ibmbigdatahub.com/blog/if-big-data-and-analytics-exist-silo-does-outcome-matter> (last accessed 6 October 2016).

[42] B Gaff, HE Sussman, and J Geetter, *Privacy and Big Data*, Computing and the Law (IEEE June 2014) at 8; P Heller, D Piziak and K Knudsen, *An Enterprise Architecture White Paper – An Enterprise Architect's Guide to Big Data  –  Reference Architecture Overview* (May 2015), available at <http://www.oracle.com/technetwork/topics/entarch/articles/oea-big-data-guide-1522052.pdf> (last accessed 6 October 2016); and see also 'Big Data and Data Protection', by Information Commissioner's Office (ICO), 28 July 2014, available at <http://www.pdpjournals.com/docs/88314> (last accessed 6 October 2016), p 6. It refers to the Gartner definition – 'Big data is high-volume, high-velocity and high-variety information assets that demand cost-effective, innovative forms of information processing for enhanced insight and decision making.'

[43] P Heller, D Piziak and K Knudsen, *An Enterprise Architecture White Paper – An Enterprise Architect's Guide to Big Data  –  Reference Architecture Overview* (May 2015), available at <http://www.oracle.com/technetwork/topics/entarch/articles/oea-big-data-guide-1522052.pdf> (last accessed 6 October 2016), p 4. '**Value.** Data has intrinsic value – but it must be discovered. There are a range of quantitative and investigative techniques to derive value from data – from discovering a consumer preference or sentiment, to making a relevant offer by location, or for identifying a piece of equipment that is about to fail. The technological breakthrough is that the cost of data storage and compute has

cases, dimensions of 'big data' have been further extended to 'incorporate related considerations such as Variability (the daily, seasonal and event-triggered peaks in data) and Complexity (the challenge of linking, cleaning and matching data across multiple sources)'.[44]

A recent study of Big Data Analytics concludes that 'there is no universally recognised operational definition of big data'.[45] However, the Commission has continued to try defining the term 'big data' in its recent work agenda towards a thriving data-driven economy as follows:

> 'Big data' refers to large amounts of different types of data produced with high velocity from a high number of various types of sources. Handling today's highly variable and real-time datasets requires new tools and methods, such as powerful processors, software and algorithms.[46]

To further differentiate big data from ordinary data, it is necessary to add the terms of 'new' and the 'three Cs' in addition to the 'three Vs'. The data only becomes big (in a sense that the current regulation may need to be reviewed) when it generates a new set of data when information is aggregated; and it is 'cross sectors', 'cross purpose', and 'cross border'.

## 3.    Solutions to the Obstacles of Legal And Regulatory Developments

Great debates in big data have been generated by the foremost thinkers in their fields including law academics, practitioners, computer scientists, and business. Such debates go beyond what can be applied under the existing legal frameworks in countries, prompting law makers and business to consider the key tensions and questions underlying the subject matter and setting legal and regulatory developments in an interdisciplinary context.

---

exponentially decreased, thus providing an abundance of data from which statistical sampling and other techniques become relevant, and meaning can be derived. However, finding value also requires new discovery processes involving clever and insightful analysts, business users, and executives. The real Big Data challenge is a human one, which is learning to ask the right questions, recognizing patterns, making informed assumptions, and predicting behaviour.'

[44] SAS® High-Performance Analytics, 'Transforming Big Data into Corporate Gold', SAS, September 2012, available at <http://www.sas.com/offices/europe/uk/downloads/bigdata/sas-hpa.pdf> (last accessed 6 October 2016), p 2.

[45] 'Big Data Analytics – An assessment of demand for labour and skills, 2012-2017', e-skills UK, January 2013, available at <https://www.e-skills.com/Documents/Research/General/BigDataAnalytics_Report_Jan2013.pdf> (last accessed 6 October 2016).

[46] Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions: Towards a thriving data-driven economy, COM (2014) 442 final, Brussels, 2.7.2014, available at <http://www.eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:52014DC0442&from=EN> (last accessed 6 October 2016), p 4.

### 3.1.   Data Protection and Security

*Data Privacy Protection*
In the era of big data economy, it is widely accepted that the benefits of such new phenomenon cannot simply be traded with privacy rights.[47] In the UK, the recent Information Commission Office (ICO) Big Data and Privacy Report emphasises that the key Principles of the Data Protection Act (DPA) which should be considered when using Big Data are:

– that the processing of the personal information is fair and lawful;
– that further processing purposes must not be incompatible with the original processing purpose (i.e. the 'purpose limitation' Principle); and
– that the conditions for processing must be satisfied. These are either: that the consent of the individual concerned has been obtained; that the processing is necessary for the performance of a contract that an individual has entered into; or that the processing is necessary for the purpose of legitimate interests.[48]

This implements the principle of fairness concerning the processing of data under the EC Directive on Data Protection.[49] The General Data Protection Regulation also contains relevant provisions concerning the assessment of the principle of fair and transparent processing. The General Data Protection Regulation requires that 'Any processing of personal data should be lawful and fair. It should be transparent to natural persons that personal data concerning them are collected, used, consulted or otherwise processed and to what extent the personal data are or will be processed.'[50] The assessment of principle of fair and transparent processing also requires including the users' consent, the purposes of data processing and the minimum period of data storage.[51]

Under the current EU Data Privacy Protection Framework, if an organisation is relying on users' consent as the condition for processing their personal data, that organisation must notify users of terms and conditions, and users' consent must be

---

[47] 'Big Data and Data Protection' by Information Commissioner's Office (ICO), 28 July 2014, available at www.pdpjournals.com/docs/88314 (last accessed 6 October 2016), p 2.

[48] 'Big Data and Data Protection' by Information Commissioner's Office (ICO), 28 July 2014, available at www.pdpjournals.com/docs/88314 (last accessed 6 October 2016); and *see also* Sayers, S and Davidson, B., ICO's Big Data reports – guidance for DPOs, (2014) *Privacy and Data Protection*, 14(8), 3-5, p 3.

[49] Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data (hereafter 'the EC Directive on Data Protection'), OJ L 281, 23.11.1995, p 31-50, Arts 6, 10 and 11.

[50] Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation), available at <http://www.eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32016R0679> (last accessed 6 October 2016), Recital (39). See also Recital (60).

[51] General Data Protection Regulation 2016, Recital (45).

'freely given, specific and informed'. If an organisation has collected personal data for one purpose and then decides to start analysing it for completely different purposes, it needs to make its users aware of this. Users also have the 'right to be forgotten' if their data is no longer necessary for the purposes for which they were collected or processed. The Court of Justice of the European Union (CJEU) recently interpreted and implemented the 'Right to be Forgotten' principle in the case brought by Google Spain SL ('Google Spain') and Google Inc. against the Agencia Española de Protección de Datos (Spanish Data Protection Agency; 'the AEPD') and a Spanish citizen Mr Costeja González.[52] The judgment supports that the operator of the search engine – Google Inc. – has the responsibility to remove search results concerning individuals, which is no longer necessary and relevant, upon users' requests. However, it is worth noting that this is a highly technical and demanding task for data controllers to determine the nature and sensibility of those data concerned and to deploy appropriate technical measures. While it is fully understandable that courts give such order to ensure the enforcement of the 'right to be forgotten' principle, it could become a burdensome responsibility for operators of search engines and lead to ineffective outcomes without harmonised standards.[53]

In practice, individuals always make the choice required in order to obtain desirable service or product, though they may not understand or read any terms. If in the sphere of big data, thousands of data exchanges by and about every individual occur at a high speed around the globe every day, it remains an ongoing challenge to meet the required principles of 'notice', 'consent' and 'right to be forgotten' according to the General Data Protection Regulation.

In 2017, the Commission Staff Working Document also pointed out that 'in a number of scenarios, public sector bodies could significantly improve their decision making using commercially-held information, notably for reasons of public health policy, spatial and urban planning, natural and technological risk management, managing energy supply grids or protecting the environment'.[54] It is suggested that access to public interest data shall be improved, provided that compliance with the General Data Protection Regulation is ensured in case of access to and processing of personal data.[55]

---

[52] Case C-131/12, *Google Spain SL and Google Inc. v. Agencia Española de Protección de Datos (AEPD) and Mario Costeja González*, Judgment of the Court (Grand Chamber) of 13 May 2014, para. 2.

[53] F Wang, '*The 'Right to be Forgotten' Ruling for Data Privacy Protection: Case Note for Google Spain SL and Google Inc. v. Agencia Española de Protección de Datos (AEPD) and Mario Costeja González* (*CJEU Case C-131/12, 13 May 2014)* 98 Journal of Intellectual Property Forum, 96-105, p 104 (2014).

[54] Commission Staff Working Document on the free flow of data and emerging issues of the European data economy: Accompanying the document Communication Building a European data economy, {COM(2017) 9 final}, Brussels, 10.1.2017, SWD(2017) 2 final, p 32.

[55] *Ibid.*, p 33.

*Ownership and Transfer of Data*

Nowadays, some car insurance companies offer policies for younger drivers with a discount if they allow a 'black box' to be fitted to the vehicle. Such a black box will typically measure lateral and longitudinal acceleration, possibly time stamped and with positional information obtained from GPS, in an attempt to profile the ability and safety of the driver, and thus estimating whether the driver is at high or low risk of becoming involved in an accident. Drivers determined to be at a lower risk may gain the benefit of lower insurance premiums.

Although these data are related to a specific person, they can be combined and used anonymously to help develop competitive insurance packages to other customers. Thus, these data are valuable to all insurance companies. In addition, these data could also be useful for software developers or automobile producers to generate dashboard warnings to drivers when similar unsafe driving occurs. If these data are merged with other datasets, the combined datasets may be used to help foresee other circumstances, for example, traffic jams, road works and accident blackspots. So who is the owner of these data? Is it the owner of the car or the driver? Or is it the seller, the insurance company, or the automobile manufacturer?

It is noteworthy that downloading a piece of software may be subject to a contract of sale. It is debatable whether contracts for the supply of intangible goods (such as data and software) should be considered as a contract of sale for service other than a contract of sale of goods, in particular the supply of individualised/custom-made software with datasets rather than standardised/ready-made software.[56] In the English case of *St Albans City and DC v. ICL*, Sir Iain Glidewell stated that software could constitute goods, because the software program itself is a formula, and it is of necessity contained in a physical medium.[57] A program in machine readable from must be contained on a machine readable medium, such as paper cards, magnetic cards, magnetic tapes and discs. On July 3, 2012, the European Court of Justice (CJEU) published its landmark decision in *UsedSoft GmbH v. Oracle International Corp*, the commercial distribution of software via online downloading may be based on a contract of sale of goods in addition to a licence agreement. Under the contract of sale of goods, the copyright holder of the software cannot prevent a permanent licensee from reselling his software.[58] This implies that a specific ownership may be attributed to intangible goods (i.e. downloaded software).[59]

---

[56] F Wang, *Law of Electronic Commercial Transactions: Contemporary Issues in the EU, US and China,* 356, 16 (2nd ed., Oxford: Routledge, 2014).

[57] *St Albans City and DC v. ICL* [1996] All ER 481.

[58] CJEU Case C-128/11, *UsedSoft GmbH v. Oracle International Corp*, Judgment of the Court (Grand Chamber), 3 July 2012.

[59] T Hoeren, *Big Data and the Ownership in Data: Recent Developments in Europe* 36 European Intellectual Property Review 751-754, p 753 (2014). This paper also points out that there is contradictive judgment concerning the ownership of data in the two German cases – *Nuremburg* (OLG Nürnberg 1. Strafsenat, Beschluss vom 23.01.2013 – 1 Ws 445/12) and *Saxony* (LAG Sachsen, Urteil vom 17.01.2007 – 2 Sa 808/05).

Accordingly, it raises debate over the ownership of data, and whether electronic databases should be considered as property. For example, in the US case of *Yazoo Pipeline*, it suggests that data 'could not exist apart from some physical storage medium, such as a computer, flash drive, tapes, or film' and 'could be accessed by a human user in a manner analogous to the access of traditional tangible property.'[60] In the recent English case of *Your Response Ltd v. Datateam Business Media Ltd*, it concerns whether data might be subjected to liens. It provides that:

> 'The electronic database was a type of intangible property which, unlike choices in action, was capable of possession and thus of being subject to a lien. An electronic database consists of structured information. Although information may give rise to intellectual property rights, such as database right and copyright, the law has been reluctant to treat information itself as property. When information is created and recorded there are sharp distinctions between the information itself, the physical medium on which the information is recorded and the rights to which the information gives rise. Whilst the physical medium and the rights are treated as property, the information itself has never been.'[61]

It is debatable whether the concept of intangible property over the Internet can be accepted. In the absence of legislation recognising intangible goods as property, it is suggested that copyright protection might help to protect the structure of the database if the database is based upon a highly original concept.[62]

It is also suggested that the *sui generis* rights of the owner of the website or the database may also be used to protect information.[63] According to the EC Directive on Databases, the holder of database rights may not place restrictions of the purpose to which the insubstantial parts are used (Art. 8(1)).[64] In the age of big data, Art. 8(1) of the EC Directive on Databases is unlikely to be applicable to mass data collection and analysis.

In addition, the EC Directive on Databases allows the protection of a substantial amount of time and money invested in structured data, which may also restrict the re-utilisation of database in the era of big data economy. In the CJEU case of the *British Horseracing Board Ltd and Others v. William Hill Organization Ltd*, the CJEU diminished the factor of publicity that '[the] fact that the contents of a database were made accessible to the public by its maker or with his consent does not affect the right of the maker to prevent acts of extraction and/or re-utilisation of the whole

---

[60] *Yazoo Pipeline* 459 B.R. 636. 653 (Bankr. S.D. Tex. 2011).

[61] *Your Response Ltd v. Datateam Business Media Ltd [*2014] EWCA Civ 281; [2014] 3 W.L.R. 887, para.42, available at <http://www.bailii.org/ew/cases/EWCA/Civ/2014/281.html> (last accessed 6 October 2016).

[62] Hoeren, *supra* n 48, at 752.

[63] Hoeren, *supra* n 48, at 752.

[64] Directive 96/9/EC of the European Parliament and of the Council of 11 March 1996 on the legal protection of databases, OJ L 77, 27.3.1996, p 20-28.

or a substantial part of the contents of a database'.[65] With regard to the determination of database 'extraction', in the case of *Directmedia Publishing GmbH v. Albert-Ludwigs-Universität Freiburg*, the CJEU also ruled that '[the] transfer of material from a protected database to another database following an on-screen consultation of the first database and an individual assessment of the material contained in that first database is capable of constituting an 'extraction', to the extent that – which it is for the referring court to ascertain – that operation amounts to the transfer of a substantial part, evaluated qualitatively or quantitatively, of the contents of the protected database, or to transfers of insubstantial parts which, by their repeated or systematic nature, would have resulted in the reconstruction of a substantial part of those contents'.[66] In theory, the robot search for data from billions of websites at least once a day for a substantial part of content should constitute an 'extraction'. However, in practice, robots have been commonly used for commercial purposes taking necessary precautions to avoid significant negative impacts.[67]

Cross-border data transfer in the phenomenon of big data may also experience legal barriers under the current EU framework. Transfer of data to a third country or to an international organisation may only take place if the other jurisdiction meets the adequate standard. However, one of the exceptions to this principle is that 'the transfer is necessary for the purposes of the legitimate interests pursued by the controller or the processor, which cannot be qualified as frequent or massive' (Art. 44(1)(h) of the Proposed General Data Protection Regulation). The term 'massive' hints at the phenomenon of 'big data'.[68] This may constitute a substantial impediment to the operation of big data from the EU to other parts of the world. However, this provision was abolished by the final General Data Protection Regulation.[69]

In response to the challenges to data ownership and cross-border data transfer, the Commission has recently started its study on such barriers, considering future policy actions, notably by taking into account the Trusted Cloud Europe report and recommendations by the European Cloud Partnership.[70] In addition, the Commission has also taken various other actions and measures to develop a healthy big data environment and to build the trust that is necessary to exploit the full potential of the data-driven economy.[71] They include:

---

[65] Case C-203/02, *The British Horseracing Board Ltd and Others v. William Hill Organization Ltd*, 9 November 2004.

[66] Case C-304/07, *Directmedia Publishing GmbH v. Albert-Ludwigs-Universität Freiburg*, 9 October 2008.

[67] R Brennenraedts, R te Velde, F Wang and others, *Feasibility Study on Statistical Methods on Internet as a Source of Data Gathering* (SMART 2010/030). *Final Report – A Study Prepared for the European Commission DG Communications Networks, Content & Technology* (2012), available at <http://www.ec.europa.eu/digital-agenda/en/news/feasibility-study-statistical-methods-internet-source-data-gathering-smart-2010030> (last accessed 6 October 2016), p 101.

[68] C Kuner, FH Cate, C Millard and DJB Svantesson, *The Challenge of 'Big Data for Data Protection* 2 International Data Privacy Law 47-49, 48 (2012).

[69] General Data Protection Regulation 2016, Art. 49.

[70] *Towards a Thriving Data-Driven Economy*, COM (2014) 442 final, Brussels, 2.7.2014, p 12.

[71] *Ibid.*, p 11-12.

– Launching a consultation and expert group to assess the need for guidance on specific issues of data ownership and liability of data provision, in particular for data gathered through Internet of Things (IoT) technology.[72]

– Launching a consultation process on the concept of user-controlled cloud-based technologies for storage and use of personal data ('personal data spaces'), and support Research and Innovation (R&I) on tools to assist users in selecting the data sharing policies that best match their needs. This is to enable users to better control and secure their data and to support projects aiming at reducing personal data breaches and ensuring original purposes of data collection.[73]

– Consumer and marketing law also horizontally applies to products based on big data technology. This is to ensure the availability of necessary and non-misleading information and fair contacts concerning the use of data collected from SMEs, consumers, suppliers and users.[74]

– With regard to security measures, the Commission will 'explore the landscape of security risks relating to big data and will propose risk management and mitigation measures, including guidelines, e.g. on good practices for secure data storage, to further a security culture in many sectors of society and help detect and better respond to cyber-attacks. The Commission will also support R&I to help reduce the risk of data breaches and of databases being exploited covertly for unlawful purposes.'[75] This continues and enhances the work from the Cybersecurity Strategy for the European Union and the Commission proposal for a Directive on Network and Information Security.[76] One of the most relevant legal measures concerning big data management is to establish a Public-Private Platform on Network and Information Security which identifies risk management and provides information sharing practices.[77]

In 2017, relevant policies and broad principles were suggested by the Staff Working Document to help shaping an EU framework for the free flow of data and improve sharing of commercial data and in particular machine-generated data which are either non-personal in nature or personal data that have been anonymised.[78] It was advised that the current EU legal framework of data localisation restrictions may not take into consideration technological advancement and the distributed nature of the Internet.[79]

---

[72] *Ibid.*, p 12.

[73] *Ibid.*, p 11.

[74] *Ibid.*, p 11.

[75] *Ibid.*, p 11.

[76] 'EU Cybersecurity plan to protect open internet and online freedom and opportunity – Cyber Security strategy and Proposal for a Directive', available at <http://www.ec.europa.eu/digital-agenda/en/news/eu-cybersecurity-plan-protect-open-internet-and-online-freedom-and-opportunity-cyber-security> (last accessed 6 October 2016).

[77] *Towards a Thriving Data-Driven Economy*, COM (2014) 442 final, Brussels, 2.7.2014, p 3.

[78] Commission Staff Working Document on the free flow of data and emerging issues of the European data economy: Accompanying the document Communication Building a European data economy, {COM(2017) 9 final}, Brussels, 10.1.2017, SWD(2017) 2 final, p 4.

[79] *Ibid.*, p 7.

It was stressed that in most cases the level of security of data in electronic format does not necessarily depend on its storage location, but rather on the security of the IT infrastructure and strength of the encryption techniques used.[80] Thus, it is feasible to remove the barriers of cross-border data transfers by securing data storage or processing, which may be achieved by 'removing obstacles to keep data in larger state of the art data centres, which are much less vulnerable to attacks, and enabling cross-border cooperation, i.e. one data centre being the back-up of another located in a different Member State'.[81] Liability issues for the use of big data have also been looked into, taking into consideration the complexity of data-based products and services and autonomous systematic applications.[82]

## 3.2.   Intellectual Property Protection

As shown above concerning big data management, big data involves creation, mass collection and re-utilisation of data from different sources. The creativity of big data-sets may be subject to the protection of intellectual property rights, whilst its re-utilisation from the mass collection of data (known as 'data-mining') may potentially infringe other right holders' intellectual property rights, in particular copyright. In order to protect rights holders' rights, the Commission has been investigating ways to enhance data-driven innovation based on data-mining, including text-mining in relation to copyright. The commission, thus, tries to note possible exception rules which may facilitate mass data analysis activities in Member States under the current copyright framework.[83]

As discussed earlier, it is commonly known that the cloud may be used as an enabler for big data analytics.[84] Parallel regulations on cloud computing and big data technologies are of necessity. It is noted that copyright aspects of cloud computing are considered as a very important issue for building the digital single market in Europe, and that it is recommended for the review of the private copy levies focused on trying to fix the existing system and included suggestions to improve it by the European Commission in 2013.[85] Accordingly, the issue of IP protection in the cloud-based environment and big data phenomenon was addressed by the European Parliament in 2014. The resolution of 27 February 2014 on private copying levies calls on 'the Commission to assess the impact on the private copying system of the use of cloud computing technology for the private recording and storage of protected works,

---

[80] *Ibid.*

[81] *Ibid.*

[82] *Ibid.*, p 40.

[83] *Towards a Thriving Data-Driven Economy*, COM (2014) 442 final, Brussels, 2.7.2014, p 11.

[84] 'Big Data in the Cloud', Intel IT Center, April 2015, available at <http://www.intel.co.uk/content/dam/www/public/us/en/documents/product-briefs/big-data-cloud-technologies-brief.pdf> (last accessed 6 October 2016).

[85] A Vitorino, *Recommendations resulting from the Mediation on Private Copying and Reprography Levies* (Brussels 31 January 2013), available at <http://www.ec.europa.eu/internal_market/copyright/docs/levy_reform/130131_levies-vitorino-recommendations_en.pdf> (last accessed 6 October 2016).

so as to determine whether these private copies of protected works should be taken into account by the private copying compensation mechanisms and, if so, how this should be done.'[86] In 2014 the Report on the Implementation of the Communication, 'Unleashing the Potential of Cloud Computing in Europe', further clarified that 'cloud based online content services provide unique opportunities to be remunerated on the basis of direct licensing deals, rather than indirect compensation mechanisms on devices, such as private copying levies'.[87] It further provided three key actions in respect of standards and certification; development of safe and fair contract terms and conditions; and the launch of the European Cloud Partnership with the aim to bring together the public and the private sector.[88] In addition, the Commission considered actions to follow up with the specific recommendations in the context of the on-going review of the EU copyright rules, in particular patent and trademark aspects in relation to cloud services.[89]

This reaffirms the relationship of protection between big data and intellectual property. That is, as Tepp at US Chamber of Commerce Foundation explained:

> 'Collecting and storing data involves patented hardware. Organizing and analyzing the data involves software that is probably both patentable and copyrightable. The computer processors that run those programs are likely patented. The data itself may be a propriety trade secret. Reports and interpretations that are produced are copyrightable. And trademarks will help us identify the companies that produce the best analyses and forecasts.'[90]

It is important that patent, copyright and trademark laws are efficiently employed to fight off unfair competition and unauthorised usage, and support innovation and market development. As the use of big data grows and develops, policymakers and businesses may face fundamental questions that need to further apply to the specifics of the big data marketplace as follows:

– What are the right incentives to encourage both the collection of data and a market for that data?
– How can we maximize the development and implementation of tools to analyse and interpret big data?
– Who should make those decisions: the government or the private sector?[91]

---

[86] European Parliament resolution of 27 February 2014 on private copying levies, P7_TA-PROV(2014)0179, (2013/2114(INI)), available at <http://www.europarl.europa.eu/sides/getDoc.do?pubRef=-//EP//TEXT+TA+P7-TA-2014-0179+0+DOC+XML+V0//EN> (last accessed 6 October 2016), para.30.

[87] Report on the Implementation of the Communication, 'Unleashing the Potential of Cloud Computing in Europe', SWD (2014) 214 final, Brussels, 2.7.2014, p 3.

[88] *Ibid*., p 2.

[89] *Ibid*., p 3.

[90] S Tepp, *Big Data and Intellectual Property Go Hand in Hand* (US Chamber of Commerce Foundation), available at <http://www.uschamberfoundation.org/blog/post/big-data-and-intellectual-property-go-hand-hand/34384> (last accessed 6 October 2016).

[91] *Ibid*.

Although 'cloud interoperability and data portability' increase efficiency and promote innovation, intellectual property (IP) rights may be used by the IP rights holders to prevent other cloud providers from making use of information concerning the existing products or operating in conjunction with other existing products. Such balance may be possible to be achieved through a sound legal infrastructure at community level, i.e. the interplay between the IP rights and competition rules which can be developed upon the previous experience of traditional software interoperability.[92]

Moreover, safe and fair model contract terms for cloud services in the phenomenon of big data may be of great help to enhance best practices and protect different right holders' rights. The EU Commission has been working on this. It is noteworthy that an expert group on model contract terms and conditions for cloud services for consumers and small firms and a working group with industry stakeholders on service level agreements for professional users were recently established in order to identify and disseminate best practices in respect of model contract terms for cloud services and to increase trust of prospective customers.[93]

### 3.3.   Jurisdictional Issues in Big Data and Cloud Environment

Last but not least, in the big data and cloud environment, the threat to the use of the cloud in big data mainly includes content infringement (e.g. data security, privacy and IP rights infringement) and performance infringement (e.g. non-compliance with the requirements of cloud computing services and/or product description). In such an environment, it is likely that data centres may be relocated or added at any time and as a result they may be located in various jurisdictions. This may contribute to the difficulty in identifying the location of infringement and determining the competent court and applicable law. In addition, the need of striking proper balance between cloud interoperability, data interoperability and other rights protection may also contribute to the complexity of determining jurisdiction and applicable law.[94]

As to jurisdiction concerning data privacy protection in the big data and cloud environment, under the current Commission's regulatory reform, Arts 3, 79(2) and 81 of the General Data Protection Regulation in conjunction with the Brussels I Regulation (Recast) extend and advance the existing rule in Arts 4, 17(3) and Recital (20) of the EC Directive on Data Protection.[95] For example, the General Data Protection Regulation provides a principal provision in its Art. 79(2) concerning jurisdiction for data protection, which stipulates that 'proceedings against a controller or a processor shall be brought before the courts of the Member State where the controller or processor has an establishment. Alternatively, such proceedings may be brought

---

[92]  F Wang, *Jurisdiction and Cloud Computing: Further Challenges to Internet Jurisdiction* 24 Issue European Business Law Review 589- 616, p 595 (2013).

[93]  Report on the Implementation of the Communication, 'Unleashing the Potential of Cloud Computing in Europe', SWD (2014) 214 final, Brussels, 2.7.2014, p 4; and *see also* Commission Decision of 18 June 2013 on 'setting up the Commission expert group on cloud computing contracts' (2013/C 174/04).

[94]  Wang, *supra* n 76, at 594-595.

[95]  Wang, *supra* n 76, at 610.

before the courts of Member State where the data subject has its habitual residence, unless the controller is a public authority acting in the exercise of its public powers'.

As to jurisdiction concerning IP rights infringement in the big data and cloud environment, information published online can be accessed everywhere in the world except for geo-blockings. This challenges the traditional theory of the determination of copyright infringement which is subject to the protection of national law, because once copyrighted work is distributed via the internet without the rights holder's consent, damage may occur immediately anywhere in different countries. Moreover, events giving rise to such damage may also be in multiple jurisdictions. This requires justification and interpretation of the rules of internet tort jurisdiction for online copyright infringement.[96] According to the effects approach in Art. 7(2) of the Brussels I Regulation (Recast) 2012 (originally Art. 5(3) of the Brussels I Regulation), there may be multiple locations that can qualify as the place where the harmful event occurs: (a) the place of the event giving rise to the damage; and (b) the place where the damage occurred. It is most likely that parties may need to enforce their IP rights in courts of different countries.[97] The CJEU once again confirmed in the recent case of *Pez Hejduk* that the expression of 'place where the harmful event occurred or may occur' in Art. 5(3) of the Brussels I Regulation is intended to cover both possibilities:

–   The place where the damage occurred; and
–   The place of the event giving rise to it.[98]

The principle of territoriality has also been consistently deployed to determine online copyright jurisdiction in *Pinckney* (Case C-170/12) and *Pez Hejduk* (Case C-441/13). That is, copyright rights, which were automatically granted without the need of registration, are subject to the principle of territoriality. Those rights are thus capable of being infringed in each Member State in accordance with the applicable substantive law.[99] The Court explained that the protection of copyright 'granted by the Member State of the court seised is limited to the territory of that Member State, a court seised on the basis of the place where the alleged damage occurred has jurisdiction only to rule on the damage caused within that Member State.'[100] The principle of territoriality limits the scope in which the courts of respective Member State of registration are capable of ascertaining damage caused in that Member State.

---

[96]   F Wang, *Online Copyright Jurisdiction: CJEU Ruling in Pez Hejduk* (C-441/13) *in Comparison with that in Pinckney (Case C-170/12) and Other Internet Tort Jurisdiction Cases* 100 Journal of Intellectual Property Forum 89-94, p 89(2015).

[97]   Wang, *supra* n 76, at 613.

[98]   Case C 441/13, *Pez Hejduk v. EnergieAgentur.NRW GmbH*, Judgment of the Court (Fourth Chamber), 22 January 2015, para. 18.

[99]   Case C 441/13 *Pez Hejduk*, para. 22; and Case C-170/12 *Pinckney*, para. 39.

[100]   Case C 441/13 *Pez Hejduk*, para. 36; and Case C-170/12 *Pinckney*, para. 45.

## 4.   Afterthoughts

Big data seems to transform how people live, work and think. It may also increase efficiency, productivity, safety, convenience, opportunities and profits in businesses and daily life. As pointed out in this paper, the phenomenon of big data still faces legal uncertainty as there is no single legislation which specifically tackles legal issues of big data. This paper suggests that traditional legal concepts and principles of database law, data privacy protection law, intellectual property law, contact law and private international law need to be interpreted and applied to the specifics of the big data marketplace.

In addition, best practices and guidelines may also be helpful to contribute to the well-being of citizens as well as to socio-economic progress.[101] The EU Commission has been in the process of consulting with Parliament, Council, Member States and all relevant stakeholders to draw up a more detailed, multi-layered and evidence-based action plan for advancing towards the data-driven economy of the future and addressing Europe's future societal challenges.[102] However, the question remains that whether it would be feasible to propose a balanced regulatory framework covering all areas of laws in response to the emerging technologies assisted by cloud services in the phenomenon of big data. This would then answer whether machine intelligence intertwined with big data in the cloud-based environment may foster a healthy economic and societal growth or cause chaos of social and legal order.

---

[101] *Towards a Thriving Data-Driven Economy*, COM (2014) 442 final, Brussels, 2.7.2014, p 12.
[102] *Ibid.*