UNIVERSITÀ DEGLI STUDI DI PARMA

Dottorato di ricerca in Diritto Commerciale:

Proprietà Intellettuale e Concorrenza

Ciclo XXVIII

BIG DATA

Balancing the Web User's and the Service Provider's Rights in the Big Data Era

Coordinatore:

Chiar.mo Prof. Pietro Vagliasindi

Tutor:

Chiar.mo Prof. Giulio Enrico Sironi

Dottorando: Andrea Tuninetti Ferrari

INDEX

CHAPTER I

THE DIGITAL REVOLUTION: BIG DATA

CHAPTER II

BIG DATA: SOURCING DATA FROM THE WEB

CHAPTER III

BIG DATA: EXPLOITATION AND ENFORCEMENT

CHAPTER I

THE DIGITAL REVOLUTION: BIG DATA

1. The «digital revolution» and Big Data: scope of work and preliminary remarks

Over the past ten years the change from mechanical and electronic technology to the digital electronics – the so-called «digital revolution» – accomplished as a result of the simplification of the technical means and the diffusion of the Internet web: nowadays, two billion people use smartphones (which have become as powerful as supercomputers), two billion (active) social media accounts populate the social networks, two billion users have turned to cloud-based services, and an unimaginable amount of hardware is connected to the Internet (the so-called «*Internet of Things*»), while corporations tune their business decisions on the analysis of huge amounts of data («*Big Data*») that they source from the abovementioned phones, computers and devices connected to the Internet.[1]

As inferable from the above, the digital revolution wholly relies on the interplay between billions of users and a high number of corporations running data-based businesses (*e.g.* Internet service providers, software houses, and, more generally, multinational groups with the necessary skills and resources to manage and exploit Big Data). These corporations, from now on, will be jointly referred to as the «*service providers.*»[2]

---

[1] For a brilliant overview of the digital innovations, see I. SIDHU, T.C. DOYLE, *The Digital Revolution: How Connected Digital Innovations Are Transforming Your Industry, Company & Career*, London, 2016: "*Once* everything *is connected to the Internet, we will have at our fingertips data on every activity, interaction, and condition known to man. Translating this data into information, of course, will require an immense effort. But thanks to infinitely scalable resources now available to everyone via the Internet and cloud, we now have the power required to collect, store, and process this information. With better analytical tools now being developed, we increasingly have the capability to translate this information into actionable knowledge and insights required for solving our problems and addressing our aspirations*" (p. 2 *et seq.*).

[2] For the sake of clarity, the word «*service*» (in the wider expression "service provider") here refers to both (i) the services that corporations provide to users in exchange for the possibility to store and exploit the user's data and user-generated content, and (ii) the

Users come into play mostly at a stage where they (more or less knowingly) supply data to the service providers, prevalently by actively communicating and uploading content on the social network platforms, but also by enjoying online services (*e.g.* video on demand, e-commerce services, and other Web-based services), and by using the above-referred Internet-connected sensors and machines (*e.g.* home automation systems, black box-equipped vehicles, healthcare analysis tools, *etc.*).

Service providers, on the contrary, are not only concerned with the collection of input data among the users, but also engage in the subsequent data processing, mining, and exploitation activities.

This factual situation has important legal implications, because it requires jurists to answer at least two questions:

1. *what are the legal schemes (and the boundaries) governing the allocation of data from the user to the service provider*; and
2. *what are the legal schemes that enable the service provider to protect its investment in data collection and data analytics.*

This paper is about Big Data and focuses on the legal implications of the collection of input data (chapter II), and of the enforcement of rights arising in connection with output data resulting from data analytics processes (chapter III).

One last preliminary remark. As the Internet is accessible from anywhere in the World, and as the national boundaries that separate a Country from another are easily crossed while navigating the Web, I do not mean to

---

business services (and goods) that multinational companies are able to enhance and improve thanks to their investments in data analytics. The category mentioned in (i) corresponds to the «*information society service providers*,»[2] usually referred to as the Internet service providers («*ISPs*»), which can be subdivided into (a) *access providers*, granting the user access to the Web, (b) *service providers*, providing other Internet-related services (*e.g.* e-mail boxes, chat-rooms, search engines, *etc.*), (c) *content providers*, which make available proprietary or licensed content (*e.g.* news, audio/video recordings, pictures, *etc.*), (d) *host providers*, leasing memory space on their servers to service and content providers who do not have their own, and (e) *maintainers*, acting on behalf of the other providers and supplying administrative and technical maintenance services to websites.

anchor the forthcoming legal analysis to the principles of Italian law on an exclusive basis. Italian law – which is the one I practice as an attorney-at-law and on which I focussed my J.D. and Ph.D studies – remains the main benchmark for the legal analysis of the intellectual property, civil, privacy and competition issues that will arise throughout this paper. However, I may often need to refer to the solutions adopted in other jurisdictions to provide a legal basis for developing an argument. This will lead to looking at the laws and case precedents of the United States of America on a prevalent basis for a practical reason: the vast majority of the service providers are US-based corporations, so that data flows (from the user to the service provider) are in fact governed by the laws of the US, and the biggest datasets are arguably retained in the US.

With this scope of work in mind, in this introductory chapter I mean to set the necessary background for the analysis that will be at the core of chapters II and III.

The forthcoming paragraphs in this chapter, therefore, aim at presenting:

(a) from the user's perspective, the reasons underlying the data flows that generate Big Data, by briefly recapping how the right to Internet access developed into a fundamental individual right (§2), which I believe will clarify why users are nowadays so eager to actively contributing to the creation of the Web;

(b) from the service providers' perspective, the reasons why data is the new "*oil of the Internet*"[3] and, above all, why investment in data deserves a high level of protection (§3); and

(c) the legal issues that arise in connection with Big Data, and which will be at the core of this paper (§4).

[3] Big Data has been described with many metaphors. One of the most famous is Hal Varian's (Chief Economist at Google) analogy: "*data is to information as sand is to silicon chips.*"

2. From the user's perspective: the right to Internet access

Looking at the digital revolution from the user's perspective, it has taken quite a long way for users to get to a point where they have ultimately become an active player of the Internet web.

The mechanic and analogue technologies, on which many communications *media* used to rely, lay at the basis of the traditional way of *passively* enjoying third party-generated content (*e.g.* films, music, radio, books, *etc.*), and is also consistent with the principle that *media* can only develop and deliver *one* communication pattern at a time (*e.g.* telephones permit vocal communication, television permits one-way delivery of audiovisual sequences, *etc.*). The same limitations applied in the first phase of the Internet era (so-called «*Internet 1.0*»), where users were bound to passively enjoying third party-generated content.[4]

Against this background, the recent digital revolution has had a material impact on the way we relate and interact with the others, the environment and, ultimately, life itself. People not only communicate and interact in the real world, but, by connecting to the Internet, they equally communicate and interact in a fictitious environment, the so-called «*cyberspace*.» [5] The

---

[4] According to B. KRISHNAMURTHY, G. CORMODE, *Key differences between Web 1.0 and Web 2.0*, in *First Monday*, 2008, "*content creators were few in Web 1.0 with the vast majority of users simply acting as consumers of content*" (the article is available at http://firstmonday.org/article/view/2125/1972). In the Internet 1.0 era personal web pages were common, consisting mainly of static pages hosted on ISP-run web servers, or on free web hosting services such as GeoCities (source: Wikipedia).

[5] F. DI CIOMMO, *Il diritto di accesso all'informazione in Internet*, in AA.VV., *Internet e diritto civile, Atti del convegno svoltosi a Camerino il 26 e 27 settembre 2014* a cura di C. PERLINGIERI e L. RUGGERI, Napoli, 2015, pp. 77-78. The Author highlights that "*the [digital] revolution does not have its roots in any cultural, philosophical or political movements (even though, as it was easy to foresee, it gave birth to such movements), because it was simply determined from the widespread use of the new means of communication (the* medium*). It is probably the first time in recent human history that a process' innovation influences so directly the human behaviours in such a way as to determine such important cultural and social changes. The daily use by millions of people worldwide of computers connected to local webs sharing the protocols used in the Internet has created the conditions for the birth of the so-called global community or cybernetic community*."

communication on the Internet differs from any other communication transmitted through the other means, because the cyberspace is characterised by «*multimedia*,» *i.e.* the possibility to transfer a combination of different content forms – such as text, audio, images – and by «*interactivity*,» which refers to the user's possibility to move and act freely in the cyberspace, also proactively contributing by uploading, publishing materials, or enjoying «*on demand*» content.[6]

The services provided, the transactions perfectioned, and, more generally, any activities taking (virtual) place in the cyberspace normally do also affect the real world (this is the case, for instance, of a purchase made on an online store, which originates the seller's obligation to supply the good, and the buyer's obligation to pay the purchase price).[7]

---

[6] J. HUNTLEY, N. MCKERREL, S. ASHGAR, *Universal Service, the Internet and the Access Deficit*, SCRIPTed Vol 1(2), 2004, p. 301, available at https://script-ed.org/wp-content/uploads/2016/07/1-2-Huntley.pdf. The unlimited communication possibilities offered by the Internet were firstly acknowledged in *Reno v. American Civil Liberties*, whereby the Supreme Court of the United States of America, in assessing the constitutionality of certain provisions of the Communication Decency Act of 1996, stated that "*the Internet is a unique and wholly new medium of worldwide human communication. Anyone with access to the Internet may take advantage of a wide variety of communication and information retrieval methods. ... All of these methods can be used to transmit text; most can transmit sound, pictures, and moving video images. Taken together, these tools constitute a unique medium- known to its users as "cyberspace"- located in no particular geographical location but available to anyone, anywhere in the world, with access to the Internet*" (see *Reno v. American Civil Liberties Union*, available at https://supreme.justia.com/cases/federal/us/-521/844/case.html). The U.S. case law went on to qualify the broadband as an "*information service*," rather than a "*communication service*," on accounts that "*Internet access is a capability for manipulating and storing information*. " See the U.S. Federal Communications Commission in its Declaratory Ruling and Notice of Proposed Rulemaking dated 14 march 2002, available at https://apps.fcc.gov/edocs_public/attachmatch/FCC-02-77A1.pdf, which emphasises that "*Internet access services ... alter the format of information through computer processing applications such as protocol conversion and interaction with stored data*," and the U.S. Supreme Court's decision of 27 June 2005 on *National Cable & Telecommunications Association et al. v. Brand X Internet Services et al. (04-277) 545 U.S. 967 (2005)* (available at https://www.law.cornell.edu/supct/html/04-277.ZS.html).

[7] In other words, what occurs in the cyberspace has, among others, legal implications, as the user's and the ISPs' conduct in the cyberspace can be the source for civil or even criminal liability. An ISP may be held liable on grounds that, for example, it suspended or interrupted the provision of services, or that it failed to identify the user, or the IP (*Internet Protocol*) associated with the user, who committed an offence on the ISP's

From a legal standpoint, the evolution of the Internet to an interactive web (the so-called «*Internet 2.0*»)[8] and, subsequently, a semantic web («*Internet 3.0*»),[9] is the result of a progressive evolution of the «*right to Internet access*» – *i.e.* the right for an individual to access the Web and exercise into the Web his fundamental rights to information, expression and communication – towards a fundamental/constitutional rank. Considering the importance acquired nowadays by the Internet web and, more generally, the digital means, the possibility for anyone to be a part to the Web is considered to be instrumental to an accomplishment of the human personality. Hence, the right to access shall not be interpreted to be a mere right to "*connect to the Web*," but as an incentive for lawmakers to grant the availability of, and the access to, the Web to the maximum possible extent.[10]

In 2010 a proposal was made to amend the Italian Constitution, by including the right to Internet access within the scope of the constitutional

---

website, or else that the ISP, in its capacity as content provider, made unlawfully available certain content on the web.

8 T. O'REILLY, *What Is Web 2.0 - Design Patterns and Business Models for the Next Generation of Software*, available at http://www.oreilly.com/pub/a/web2/archive/what-is-web-20.html. Web 2.0 describes the World Wide Web websites characterised by user-generated content, usability (ease of use, even by non-experts), and interoperability (meaning that a website can work well with other products, systems and devices) for end users.

9 T. BERNERS-LEE, J. HENDLER, O. LASSILA, *The Semantic Web - A new form of Web content that is meaningful to computers will unleash a revolution of new possibilities*, available at https://www.scientificamerican.com/article/the-semantic-web/, according to which Internet 3.0 is the web of data that can be processed by machines.

10 F. DI CIOMMO, *op. cit.*, p. 85, P. PASSAGLIA, *Diritto di accesso ad Internet e giustizia costituzionale comparata. Una (preliminare) indagine comparata*, available at http://www.giurcost.org/studi/passaglia.htm, and A.R. POPOLI, *Social network e concreta protezione dei dati sensibili: luci ed ombre di una difficile convivenza*, in *Dir. inf.*, 2014, p. 989. The latter Author underscores that the Internet communication qualifies as one of the possible ways to exercise the freedom to express and the freedom of thought pursuant to Article 21 of the Italian Constitution. Considering the current lack of regulation, there is an actual risk that the exploitation of the web could be *de facto* devolved to the exclusive management of the big economic and institutional players, thereby limiting the abovementioned freedoms. As a consequence, as mentioned below, Stefano Rodotà proposed at the 2010 Internet Governance Forum to amend the Italian Constitution to add Article 21-*bis* governing the right to Internet access.

rights, as follows: "*Everyone has an equal right to access the Internet web, in equal conditions, with technically adequate modalities and which remove any economic and social obstacle.*"[11]

At a national level, the recognition of the fundamental/constitutional rank of the right to Internet access has moved forward since the first decade of the 21st century: by now, some jurisdictions already contain provisions aimed at erasing the «*digital divide.*»[12]

---

[11] The proposal for an amendment of the Italian Constitution by introducing the abovementioned draft Art. 21-*bis* is available at http://www.senato.it/japp/-bgt/showdoc/frame.jsp?tipodoc=Ddlpres&leg=16&id=00519114&part=doc_dc&parse=si. According to G. AZZARITI, *Internet e Costituzione*, 2011 (available at http://www.costituzionalismo.it/articoli/392/), the proposal for the constitutional amendment shall be read to impose a duty on the public institutions – first of all the Italian lawmaker – to guarantee the availability of the and the access to the Web to the maximum number of individuals, to minimise the risk that, lacking more detailed rules, the use of the Internet may be *de facto* exclusively governed by the ISPs

[12] For an overview of the national legislations and case law precedents regarding the recognition of the right to Internet access, see F. DI CIOMMO, *op. cit.*, p. 85 *et seq.*.

Since 2010, any provider operating on the territory of **Finland**, in its capacity as a "*provider of a universal service*," has a legal duty to provide any Finnish citizen (who so applies) with broadband Internet connection, meaning a connection of at least 1 Mbps. The target pursued by the Finnish government is to be able to provide a connection of at least 100 Mbps to the Finnish in the short term, reaching the remotest corners of the Country, in such a way as to erase the so-called *digital divide* across the Country (see http://www.bbc.com/news/10461048).

Consistently, **Spanish** law of 4 March 2011, no. 4 mandates that the providers shall put any individual in a position to avail of a broadband connection (*i.e.* 1 Mbps or more), irrespective of the technological device, that is connected to the web, thereby qualifying the broadband connection among the universal services. Pursuant to Article 52 of Spanish law 4/2011 the provision broadband connection falls within the definition of universal service, which shall be ensured irrespective of the deployed technology and from the availability of a fixed infrastructure.

Estonia had taken a similar view since 2000, when its Telecommunications Act expressly qualified the right to Internet as a universal service, subsequently bolstering such recognition in 2010, by prescribing that Internet access must be made available to all citizens, irrespective of the latter's geographical position, and at a fairly accessible price (see Article 5, para. 1, of Estonian law no. 151/2010).

Overseas, the **Costa Rica**'s Constitutional Court concluded that "*the Government's delay in the creation of an open telecommunications marketplace equals a violation of a fundamental right, which seriously mines the consumers' freedom of choice and jeopardises the fight to the digital divide.*" In so ruling, the Constitutional Court expressly recognised the Internet as a fundamental communication tool, which plays an essential role in overcoming the technical barriers that the other communication means are not able to supersede (see *Sala Constitucional De La Corte Suprema De Justicia*,

At an international level,[13] in 2013 the European Parliament kicked-off discussions regarding a private citizen's petition for the inclusion of a draft Art. 3(a) (*Access to the Internet and the information society*) within the Treaty on European Union,[14] while, in the same year, the United Nations published a report, whereby they expressly qualified the Internet as a fundamental human right falling within the scope of Art. 19 of the Universal Declaration of Human Rights, which protects the right to freedom of opinion and expression.[15]

Judgement of 30 July 2010, *sentencia*: 12790, *expediente*: 09-013141-0007-CO, the English unofficial translation of which is available at http://www.technollama-.co.uk/costa-rican-court-declares-the-internet-as-a-fundamental-right).

Even more so, the **Ecuador**'s and **Greece**'s Constitutions expressly contemplate the right to Internet access: the Greek Constitution states that "*everyone has a right to take part in the information society,*" and "*the State has a duty to facilitate the access to the information circulating in an electronic form, as well as the production, exchange and diffusion of said information*" (see Art. 5A(2) of the Greek Constitution); the Ecuador's Constitution grants the citizen a right to the access to the communication and information technologies and tasks the Government with the duty to make said right effective and enforceable (see Art. 16 and Art. 17 of the Ecuador's Constitution).

[13] The aforementioned provisions of law and decisions set the route to be pursued for the crystallisation of the right to Internet access as a fundamental right. The next step shall necessarily be a more clear-cut definition of the perimeter and the targets of the right to Internet access. The recognition of the right to Internet access is in line with the objectives set forth in the first Conference of Ministers responsible for Media and New Communication Services held in Reykjavik, Iceland, between 28 and 29 May 2009, which investigated the role of the *media* in the modern society, keeping an eye on the rapid technologic changes. The speeches of the Conference are available at http://www.coe.int/t/dc/files/ministerial_conferences/2009_media_communication/default_EN.asp

[14] The Parliament's notice to the Member States is available at http://www.europarl.europa.eu/meetdocs/2014_2019/plmrep/COMMITTEES/PETI/CM/2016/1010/1028726EN.pdf, whereby the Parliament acknowledges that Article 3(a) "*should formalise the importance of internet access as a vital tool for encouraging the development of the European digital economy based on online content and applications, in order to promote innovation and economic growth and to improve services for citizens and businesses. The aim of the recognition of the right to Internet access as one of the fundamental principles of the European Union should be to promote a digital single market and genuine interoperability among technology services.*"

[15] See United Nations' resolution titled "*The promotion, protection and enjoyment of human rights on the Internet*" (A/HCR/20/L.13), available at https://documents-dds-ny.un.org/doc/UNDOC/LTD/G12/147/10/PDF/G1214710.pdf?OpenElement. See also the United Nations' report titled "*Report of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression*" (A/HRC/17/27), available

Following the recognition of a priority rank of the right to Internet access, the national constitutional and civil courts across jurisdictions have been facing the uneasy task to find a balance between Internet access and other individual rights, such as exclusivity afforded by copyright laws.[16] Overall, the court's unanimous response was that the right to Internet access shall prevail, in principle, on copyright;[17] however, Internet access shall not

---

at http://www2.ohchr.org/english/bodies/hrcouncil/docs/17session/A.HRC.17.27_en.pdf, whereby the Special Rapporteur, Mr. Frank La Rue, concludes that "*Given that the Internet has become an indispensable tool for realizing a range of human rights, combating inequality, and accelerating development and human progress, ensuring universal access to the Internet should be a priority for all States. Each State should thus develop a concrete and effective policy, in consultation with individuals from all sections of society, including the private sector and relevant Government ministries, to make the Internet widely available, accessible and affordable to all segments of population. At the international level, the Special Rapporteur reiterates his call on States, in particular developed States, to honour their commitment, expressed* inter alia *in the Millennium Development Goals, to facilitate technology transfer to developing States, and to integrate effective programmes to facilitate universal Internet access in their development and assistance policies*" (§§85-86).

[16] For an overview of the case precedents discussing the conflicts between right to Internet access and intellectual property, see F. DI CIOMMO, *op. cit.*, p. 92.

[17] The Federal Court of **Australia** took a stand in *Universal Music Australia Pty Ltd v. Sharman License Holdings Ltd.:* Kazaa, which was available worldwide and free of charge, enabled users to share copyright-protected musical works. Kazaa was found by the Court to infringe, and a provisional order was made to restrain future infringements, without intruding on freedom of speech and communication. The continuation of the Kazaa system was assessed not be a contravention of the order if the system is modified, in a manner agreed by the applicants or approved by the Court. See the Federal Court of Australia's decision of 5 September 2005 in *Universal Music Australia Pty Ltd v. Sharman License Holdings Ltd [2005] FCA 1242*, available at *http://www.austlii.edu.au/au/cases/cth/federal_ct/2005/1242.html*.

Later in 2006 the Federal Court of Australia's 18 December 2006 decision in *Cooper v. Universal Music Australia Pty Ltd [2006] FCAFC 187* rejected an appeal by a website owner and an ISP against findings of infringement of copyright of sound recordings by the operation of a website (MP3s4FREE) which linked to mp3 recordings (appeal decision is available at http://www.austlii.edu.au/au/cases/cth/FCAFC/2006/187.html).

In 2011 the Federal Court of Australia further elaborated on the topic, and ultimately came to the conclusion that the violation of the copyright laws committed through an unauthorised download does not justify the suspension of the access to the Internet to be imposed on the infringers, on grounds that it is not the access itself the means of the violation, but the actual use – which is made possible through the access to the Internet, but does not coincide with the same – of a given software, which makes the copyright's violation possible. See the Federal Court of Australia's decision dated 4 February 2010 in *Roadshow Films Pty Ltd v. iiNet Limited* (No. 3) [2010] FCA 24, available at http://www.austlii.edu.au/au/cases/cth/FCA/2010/24.html. The Court found that: "*it is*

obvious that the respondent's provision of the Internet was a necessary precondition for the infringements which occurred. However, that does not mean that the provision of the internet was the 'means' of infringement. The provision of the internet was just as necessary a precondition to the infringements which occurred in the Kazaa proceedings, but no ISP was joined as a respondent. The focus in that proceeding was correctly upon the more immediate means by which the infringements occurred, namely the Kazaa system. … The provision of the Internet was also a necessary precondition to the infringements that occurred by the people who accessed Mr Cooper's website, but, again, the focus in those proceedings was rightly upon the narrower and more specific 'means' of infringement, namely the website and the ISP that hosted it. As with cases like Kazaa and Cooper, in the present circumstances there are also other necessary preconditions to bring about infringement, such as the computers upon which the infringements occurred or the operating systems on those computers, for example, Microsoft Windows. The use of the BitTorrent system as a whole was not just a precondition to infringement; it was, in a very real sense, the 'means' by which the applicants' copyright has been infringed. This is the inevitable conclusion one must reach when there is not a scintilla of evidence of infringement occurring other than by the use of the BitTorrent system. Such conclusion is reinforced by the critical fact that there does not appear to be any way to infringe the applicants' copyright from mere use of the internet. There will always have to be an additional tool employed, whether that be a website linking to copyright infringing content like Mr Cooper's website in Cooper, or a p2p system like the Kazaa system in Kazaa and the BitTorrent system in the current proceedings. Absent the BitTorrent system, the infringements could not have occurred."

The **French** *Conseil Constitutionnel* expressly considers the right to internet access to enjoy protection under the *Déclaration des droits de l'homme et du citoyen de 1789*, and the right to Internet access shall therefore prevail, in principle, on copyright. Consistently with the findings of the Federal Court of Australia, the *Conseil Constitutionnel* – in denying the compliance with the French Constitution of the statutory provisions, pursuant to which the *Commission de protection des droits*, in its capacity as the French Authority, was empowered to suspend the individual user's access to the Internet in cases where the user's account was found to have breached third parties' copyright – stated that, having regard to the freedom to communicate and the freedom of thought and opinion set forth in Article 11 of the *Déclaration des droits de l'homme et du citoyen de 1789*, a public Authority cannot be vested with the power to limit the individual's right to access the Internet, in cases where the protection of the right at stake, such as the copyright, would not *per se* prevail on the invidual's right to access the Internet. See *Conseil constitionnel*, 10 June 2009 no. 2009-580 DC, English version available at http://www.conseilconstitutionnel.fr/conseilconstitutionnel/root/-bank/download/2009580DC2009_580dc.pdf, according to which "*The powers to impose penalties created by the challenged provisions vest the Committee for the protection of copyright, which is not a court of law, with the power to restrict or deny access to the internet by access holders and those persons whom the latter allow to access the internet. The powers vested in this administrative authority are not limited to a specific category of persons but extend to the entire population. The powers of this Committee may thus lead to restricting the right of any person to exercise his right to express himself and communicate freely, in particular from his own home. In these conditions, in view of the freedom guaranteed by Article 11 of the Declaration of 1789, Parliament was not at liberty, irrespective of the guarantees accompanying the imposition of penalties, to vest an administrative authority with such powers for the purpose of protecting holders of copyright and related rights.*"

constitute legal basis for infringing third party's proprietary rights: according to the Irish High Court, a court shall in fact have the power to *restrict* the right to access, by prohibiting the user the use of a particular channel/means he deployed to access the Web and infringe others' proprietary rights, without necessarily *denying* the right to access in its entirety.[18]

The route towards the recognition of a constitutional rank to the right to Internet access also stimulated discussions regarding the scope of such right: provided that the user has a right to navigate the Web, does he have also an unrestricted right to access content, data and information available into the Web (in cases where, of course, the information is not protected by copyright, or access thereto is restricted on a subscribers only basis)?[19]

From a mere technologic standpoint, there is no doubt that this is feasible, considering that the search engines (*e.g.* Google, Bing, Yahoo! Search) make web pages they have indexed available from their «*cache*.»[20] However, in

---

[18] In *Emi Records (Ireland) Ltd et al. v. Eircom Ltd.* the **Irish** High Court analysed the issue of copyright infringement, and concluded that taking "*the subscriber .. off service except for phone or television internet access ... is a serious sanction. Some would argue that it is an imposition on human freedom. There is no freedom, however, to break the law. Further, while it is convenient to have internet access at home, most people in Ireland have only to walk down to their local town centre to gain access for around €1.50 an hour.*" According to the High Court, the key to reaching a balance between enforcement of copyright and right to Internet access is the concept of *accessibility*: there is a substantial difference between the right to Internet access, in its widest meaning, and a supposed "*right to access the Internet from the user's home.*" The High Court's decision relies on the reasoning that, considering the current *status* of technology, a limitation of the right to access the Internet from home (or from another particular place/channel) does not automatically prevent the individual from accessing the Web *tout court*, because there remain other (technical/factual) solutions that the user may implement to access the Internet at an affordable price, which still ensures the possibility for everyone to exercise his freedom to communicate and express ideas, in compliance with the third parties' proprietary rights. See *Emi Records (Ireland) Ltd et al. v. Eircom Ltd.*, [2010] IEHC 106. The High Court's decision of 16 April 2010 is available at https://www.scribd.com/document/39179082/EMI-Records-v-Eircom-Ltd.

[19] F. DI CIOMMO, *op. cit.*, p. 100.

[20] A «*cache*» is an information for the temporary storage of web documents; search engines cache websites to provide a way of retrieving information from websites that are temporarily not accessible or a way of retrieving data more quickly than by clicking the direct link. A relevant side effect of caching information is that nothing that is ever uploaded on the Internet is lost or *forgotten*.

landmark case *Gonzàlez v. Google* the Court of Justice of the European Union denied the subsistence of any right to access content and information. The Court of Justice considered that, in cases where the retention of a link is found to be "*inadequate, irrelevant or no longer relevant or excessive in the light of the time that had elapsed*," a search engine must consider requests from individuals to remove links to freely accessible web pages resulting from a search on their name. In case of the engine's denial, the individual may seek the competent authorities (*e.g.* data protection authorities) to intervene and order the search engine to remove the links from search results.[21]

The Court of Justice, in a nutshell, recognised that any individual, at certain conditions, has a *right to be forgotten*. The Court concluded that the individual's right that the Internet service provider remove the indexing, thereby making the information no more available on the Internet, shall in principle prevail on the service provider's economic rights and, above all, on the interest of the public (the other web users) to gather such information, unless, of course, serious reasons (*e.g.* the data subject's public role) legitimate that the interest of the public override the data subject's privacy.[22]

The legal issues dealt with in the case law precedents presented above make it clear that the statute governing the rights granted to a user navigating

---

[21] Court of Justice of the European Union, Judgment of 13 May 2014, *Google Spain SL and Google Inc. v. Agencia Española de Protección de Datos (AEPD) and Mario Costeja González*, available at http://curia.europa.eu/juris/celex.jsf?celex=62012CJ013-1&lang1=en&type=TXT&ancre.

[22] Para. 81 of the Court of Justice's decision reads as follows: "*it is clear that it cannot be justified by merely the economic interest which the operator of such an engine has in that processing. However, inasmuch as the removal of links from the list of results could, depending on the information at issue, have effects upon the legitimate interest of internet users potentially interested in having access to that information, in situations such as that at issue in the main proceedings a fair balance should be sought in particular between that interest and the data subject's fundamental rights under Articles 7 and 8 of the Charter. Whilst it is true that the data subject's rights protected by those articles also override, as a general rule, that interest of internet users, that balance may however depend, in specific cases, on the nature of the information in question and its sensitivity for the data subject's private life and on the interest of the public in having that information, an interest which may vary, in particular, according to the role played by the data subject in public life.*"

the Internet is still in the making. However, the advent of the digital era continuously raises issues, the solution of which requires jurists to take a new legal approach and to provide responses that are in line with the digital progress.[23]

The Italian regulator has somehow openly recognised that the current framework is not suitable to answer all the new legal questions posed by the digital revolution, especially in terms of the web user's rights, so that in 2014 the Chamber of Deputies elaborated a *Declaration of Rights in the Internet*,[24] which relies on the principles of equality, freedom and dignity, and which aims at establishing, in view of future legislative reforms, the rights to Internet access, net neutrality, anonymity and oblivion.[25]

3. From the service provider's perspective: Big Data and the «*Internet of things*»

Looking at the digital revolution from the service providers' perspective, the widespread exercise by individuals of their right to Internet access,[26]

---

[23] C. GALLI, *I diritti IP nel mercato globale e nella nuova economia digitale: le ragioni di un Convegno*, in *Dir. ind.*, 2015, p. 105. The Author explains that the evolution of the Internet "*while representing one of the most evident and clamorous demonstration of the new digital economy, with which the companies in any sector must cope, on the other hand poses a series of brand new problems, which imply the interaction of different disciplines and the necessity to compare and integrate the expertise of the legal experts and of the IP attorneys with that of the web's technicians and experts.*"

[24] The latest version of the draft Constitution of the Internet (2015) is available at http://www.camera.it/application/xmanager/projects/leg17/commissione_internet/dichiarazione_dei_diritti_internet_pubblicata.pdf.

[25] C. GALLI, *ibidem*; P. PAGANINI, *Verso l'Internet delle cose*, in *Dir. ind.*, 2015, p. 111.

[26] The switch from the Internet 1.0 (where users were limited to the passive viewing of content) to the Internet 2.0 (where users are in a position to interact and collaborate with each other in a social media dialogue as creators of user-generated content in a virtual community - see above) brought to the current scenario, where users spend a constantly increasing amount of time on social platforms. Teenagers on an average spend nine hours a day on social media platforms, according to the report by Common Sense Media (*The Common Sense Census: Media Use by Tweens and Teens*, available at https://www.commonsensemedia.org/research/the-common-sense-census-media-use-by-tweens-and-teens or http://edition.cnn.com/2015/11/03/he-alth/teens-tweens-media-screen-use-report/), while 30% of all time spent online by all users is now allocated to social media interaction (see E. ASANO, *How Much Time Do People Spend on Social Media?*, available at http://www.socialmediatoday.com/-marketing/how-much-time-do-

combined with other enabling factors – such as the miniaturisation of computer power, wireless interconnectivity, the decrease in the cost of storage[27] and, overall, the expansion of the Internet[28] – led to the production of large-scale quantities of data.

Service providers have subsequently realised that, by capturing, structuring, and analysing these huge volumes of data, and understanding the relationships between data, the data holder could gain valuable, non obvious, information, facts, relationships, indicators that, where read at a macro level, may reflect event probabilities or consumer preference trends that may guide business decision. [29] Service providers have therefore made major

---

people-spend-social-media-infographic: "*Broken down, time spent on social media differs across each platform. YouTube comes in first, consuming over 40 minutes of a person's day (i.e. 1 year and 10 months in a lifetime). Facebook users will spend an average of 35 minutes a day, totalling 1 year and 7 months in a lifetime (some statistics include Facebook, Facebook-owned Instagram, and Facebook Messenger for total time spent on Facebook). Snapchat and Instagram come in next with 25 minutes and 15 minutes spent per day, respectively. Finally, users will spend 1 minute on Twitter, spanning 18 days of usage in a lifetime*").

[27] Aaron Levie (Co-Founder at Box, a online file sharing and content management service for businesses), in his 2014 speech at the Stanford University ("*Building for the Enterprise*"), remarked that in 2004 it was very difficult both for users and businesses to share files. Web-mails, for example, would grant user only a 15MB storage space – because of the cost of storing data – which was next to nothing. His idea to start a file hosting company followed three key factors: "*cost of storage dropping dramatically; more powerful browsers and networks; more locations and people to share with*." The speech is available on iTunes University (Stanford course "*How to build a Start-up*").

[28] See D.A. PRANGE, *Navigating the protection of big data*, in *Intellectual Property Magazine*, 2017, p. 54.

[29] A widely used example of how the Big Data works is the joint research carried out by Stanford, Columbia and Microsoft Corporation in 2010: researchers developed a new way to predict harmful interactions between pharmaceuticals based on the Internet rather than on chemical interactions or trials. In cooperation with Microsoft, the university researchers analysed logs of millions of online searches made by consenting users of the Google, Bing, and Yahoo! search engines. Using statistical techniques, researchers observed that users who searched for the names of two drugs – Paxil and Pravastatin – were likely to also enter search terms related to hypoglycaemia. This correlation led the researchers to hypothesise, and later to experimentally confirm, that Paxil and Pravastatin can cause adverse effects when taken together. See W. WHITE, N.P. TATONETTI, N.H. SHAH, R.B ALTMAN, E. HORVITZ, *Web-Scale Pharmacovigilance: Listening to Signals from the Crowd*, 2013, available at http://jamia.bmj.com/content/20/3/404.full.pdf;

investments in data, to understand market trends, improve business processes and products, and ultimately gain competitive advantage.

The peculiarity of this process is that it starts with (input) data and it also ends up with (output) data. Data is collected, then processed, so that the data holder might ultimately tune its business decisions based on data analytics.

Input data has a wide variety of sources and is different in nature.

Data can be generated from e-mail, video, click streams, social media platforms, entries on search engines, Internet-based transactions, personal or shared computers, mobile phones, tablets, *etc.*.

Data is also sourced from interconnected instruments and sensors, *i.e.* physical devices embedded with software and network connectivity, that enable devices to collect and exchange data. These devices – *e.g.* health monitoring implants, home automation systems, automobiles with built-in sensors, *etc.* – collect data the same way computers or smart phones do, and then autonomously flow the data to a data receiver. This is known as the «*Internet of Things*,» an expression which aims at underscoring the devices' ability to create and communicate data, and, ultimately, to even take decisions based on said data.[30]

---

[30] The concept of the «*Internet of Things*» is ascribable to a speech by Peter T. Lewis to the Congressional Black Caucus Foundation 15th Annual Legislative Weekend. See, for example, http://www.chetansharma.com/IoT_History.htm. Internet of Things is considered to be an evolution of the "*machine-to-machine*" whereby direct communication occurs between devices using any communications channel. To make an example of how the Internet of Things works, I. SIDHU, T.C. DOYLE, *The Digital Revolution cit.*, p. 8 describe how the Tesla Model S "*unquestionably [became] the best electric car on the planet.*" This is "*not only because it is* electric, *but because it is* digital. … *Virtually everything in the car that can be measured has an active sensor on it that is connected to the car's digital network. You can tell your Tesla Model S to park itself neatly into your garage, so you don't have to wedge your body out when finished. And with its mobile app, you can remotely check the cabin temperature on a hot day and tell the vehicle to power up the AC, so it will be at a desired temperature when you get to the car. The car has dozens of other cool features that leverage digital technology. But there's one feature that sets it apart from virtually any other vehicle on the road. Aside from a handful of parts that need routine replacement—think tires and wiper blades—the bulk of the vehicle's components and functions were designed to be upgraded, not by mechanics wielding wrenches, but by software engineers working in Tesla's Silicon Valley research and development labs. Like an iPhone, the Tesla S gets*

Data may consist in (either self-generated, or derived) market data, personal data, including sensitive data, confidential information, government data, employee data, *etc.*.[31]

This huge amount of data is usually referred to as «*Big Data.*»

Although there is not a universal definition of Big Data,[32] Big Data is mostly described through reference to the five «*V's,*» *i.e.* its key factors:

(a) ***Volume****:* the sheer amount of data generated and data intensity that must be ingested, analysed, and managed to take decisions based on complete data analysis;

(b) ***Velocity****:* how fast data is being produced and changed and the speed with which data must be received, understood, and processed; and

---

better every time the company releases a new software update over the Internet. They can make the car safer, more reliable, and even more pleasurable."

[31] See the 2014 White House report titled *Big Data: Seizing Opportunities, Preserving Value*, available at https://www.whitehouse.gov/issues/technology/big-data-review.

[32] G. PRESS, *op. cit.;* the first documented use of the term «*big data*» appeared in a 1997 paper by two NASA scientists, Michel Cox and David Ellsworth, in an attempt to describe the problems they were facing with computer graphics, due to the fact that "*data sets are generally quite large, taxing the capacities of main memory, local disk, and even remote disk. We call this the problem of big data. When data sets do not fit in main memory (in core), or when they do not fit even on local disk, the most common solution is to acquire more resources.*". The term was added to the Oxford English Dictionary in 2013 (see http://mashable.com/2013/06/13/dictionary-new-words-2013/#wlzvptDvuuqE) and appeared in Merriam-Webster's Collegiate Dictionary only in 2014 (see http://www.computerworld.com/article/2489571/it-management/selfie--big-data-and---make-the-2014-dictionary.html). The Oxford English Dictionary's definition is largely linked to the issues that gave birth to the phenomenon itself, *i.e.* the difficulties arising in connection with the management and storage of huge amounts of data that accrue in the business IT systems: "*data of a very large size, typically to the extent that its manipulation and management present significant logistical challenges*" (See G. PRESS, *Big Data News: A Revolution Indeed*, available at http://www.forbes.com/sites/gilpress/2013/06/18/big-data-news-a-revolution-indeed/#5496d3b7b9fb). The Tech American Foundation's definition goes beyond, to underscore that the analysis of data is suitable to balance the management issues with the provision of data, at a granular level, which add value to the business: "*large volumes of high velocity, complex, and variable data that require advanced techniques and technologies to enable the capture, storage, distribution, management, and analysis of the information*" (See the TechAmerican Foundation's report titled *Demystifying Big Data*, available at http://www1.unece.org/stat/platform/pages/viewpage.action?-pageId=80053387).

(c) ***Variety**:* the rise of information coming from new sources both inside and outside the walls of the enterprise or organisation creates integration, management, governance, and architectural pressures on the information technology systems;

(d) ***Veracity**:* the reliability and truthfulness of data for the purposes of taking a business decision based on data analytics; and

(e) ***Value**:* the ability to easily access data and deliver quality analytics that enables informed decisions.[33]

The above description, however, is probably not satisfactory in underscoring how important Big Data actually is for businesses.[34] As a phenomenon, Big Data is, above all, "*an opportunity to gain a more complex understanding of the relationships between different factors and to uncover previously undetected patterns in data by leveraging advances in the technical aspects of collecting, storing, and retrieving data along with innovative ideas and techniques for manipulating and analysing data.*"[35]

This is the reason why companies and start-ups have created *ad hoc* tools and trained specialised personnel («*data scientists*») to enable the capture, storage, search, sharing, and, above all, analysis of data in a way that is valuable to the organisations. The opportunities presented by Big Data have substantially required corporations to accept a cultural shift, whereby more

---

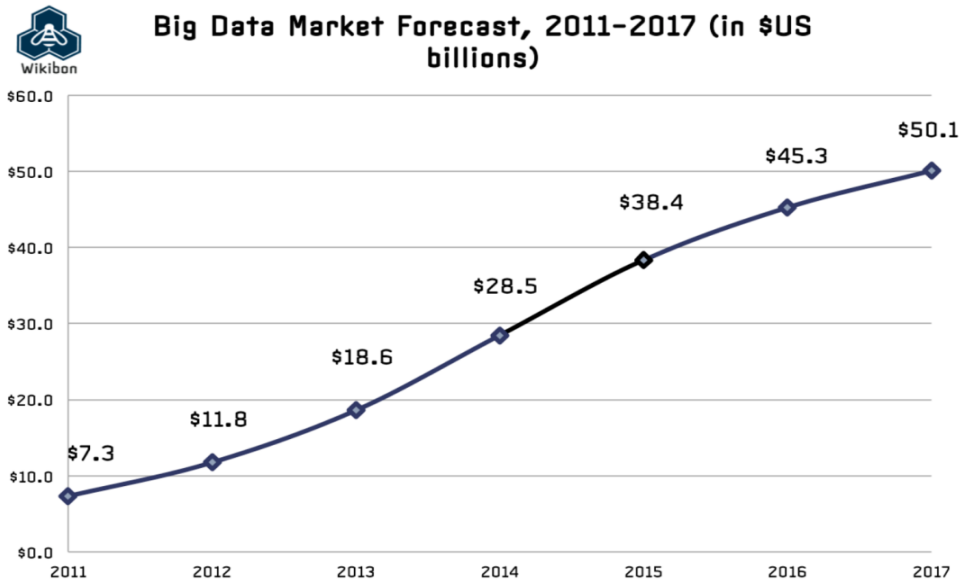
[33] D. NAVETTA, *Legal Implications of Big Data: A Primer*, *Issa Journal*, 2013, available at https://c.ymcdn.com/sites/www.issa.org/resource/resmgr/journalpdfs/feature0313.pdf. See also http://antitrust.freshfields.com/big-data.

[34] In 2014 the UC Berkeley's Master of Information and Data Science Program launched a survey among forty data science *thought leaders*, asking the question "*What is Big Data?*" The answers consistently argued that that it's not the size of data that matters, but the tools being used or the insights that can be drawn from a dataset (see https://datascience.berkeley.edu/what-is-big-data/).

[35] Among the different definitions presented in the Berkeley's survey, I chose the definition given by Amy Escobar, data scientist at 2U ("*an education technology company that partners with top colleges & universities to bring their degree programs and credit-bearing courses onlines,*" see www.2u.com), because I deem that her definition best pinpoints the rationale of investing in Big Data, *i.e.* the pursuit of business-relevant information that is otherwise hidden below the unstructured data that input operations provide to companies.

and more decisions are made by algorithms with transparent logic, operating on documented immutable evidence, so that 'Big' shall refer more to the pervasive nature of this change than to any particular amount of data.[36]

No surprise, then, if Big Data was a USD 28.5 billion market in 2014, and it has been estimated that the Big Data market will be nearly twice as double, USD 50.1 billion, by the end of 2017 (see the chart below).[37]



Source: Wikibon

---

[36] *Ibid.*, see, in particular, the definition of Big Data according to Daniel Gillick, Senior Research Scientist at Google. The White House's Executive Office of the President, in the 2014 report mentioned *supra* (*Big Data: Seizing Opportunities, Preserving Value*), comments Big Data's definitions as follows: "*Most definitions reflect the growing technological ability to capture, aggregate and process an ever-greater volume, velocity, and variety of data . In other words, «data is now available faster, has greater coverage and scope, and includes new types of observations and measurements that previously were not available.» More precisely, big datasets are «large, diverse, complex, longitudinal, and/or distributed datasets generated from instruments, sensors, Internet transactions, e-mail, video, click streams, and/or all other digital sources available today and in the future.»*" The notion of Big Data validated by the White House is therefore shorthand for the collection, processing, analysis and use of vast exploitable datasets of unstructured and structured digital information.

[37] See J. KELLY, *Big Data Vendor Revenue and Market Forecast 2013-2017*, available at http://wikibon.org/wiki/v/Big_Data_Vendor_Revenue_and_Market_Forecast_2013-2017. Prospectively, McKinsey's report on the Internet of Things estimates the potential economic impact between USD 4 trillion and USD 11 trillion a year by 2025 () Between 2013 and 2022, Cisco estimates that digital transformation will generate $19 trillion in economic activity, nearly half of which will be from the replacement of activities or things that will simply fade away like the local travel agent, printed encyclopaedia, and classified newspaper advertisement ().

Input data is, however, usually *per se* not conclusive, and may not provide added value to corporations, as there is a large gap to be bridged before Big Data can be harnessed effectively. Hence, following data collection comes data processing: data is processed on internal/external platforms by using a wide variety of algorithms and criteria – *e.g.* business intelligence and analytics applications, data visualisation, machine learning, pan-enterprise search – to produce output data.[38]

Output data is then communicated to addressees and places within the business organisation, where output data can be most effectively used for a variety of purposes, *e.g.* business analysis, product development, sales and marketing, risk management, capital management or finance.[39] Output data may also be transmitted outside the business organisation, *e.g.* data may be sold (or licensed) to a third party purchaser (or licensee).[40]

Output data is then exploited in accordance with the data holder's business vision. Here are just a few examples of how Big Data affects some key sectors:[41]

---

[38] R. KEMP, *Legal aspects of managing Big Data*, in *Computer Law & Security Review*, 2014, pp. 489-490.

[39] *Ibid.*: "*In insurance, for example, vehicle on board telematics and location based services can inform the insurer of a driver's general skill and care and where he or she was when the accident occurred. This data can be used by underwriters to assess risk and premium costs, claims assessors and to evaluate fault, the finance department to allocate capital based on risk and hence pay-out profile, the compliance team for reporting to the regulator, by product development to consider new product offerings and for marketing purposes.*"

[40] Data brokers collect detailed employment information concerning approximately 190 million individuals (including salary information) and sell it to debt collectors, financial institutions, and other entities (see, for example, http://www.nbcnews.com/technology-/exclusive-your-employer-may-share-your-salary-equifax-might-sell-1B8173066);

[41] See *The Big Data Market: 2016 - 2030 - Opportunities, Challenges, Strategies, Industry Verticals and Forecasts*, which the website Research and Markets summarises as follows: "*-In 2016, Big Data vendors will pocket over $46 Billion from hardware, software and professional services revenues. - Big Data investments are further expected to grow at a CAGR of 12% over the next four years, eventually accounting for over $72 Billion by the end of 2020. - The market is ripe for acquisitions of pure-play Big Data startups, as competition heats up between IT incumbents. - Nearly every large scale IT vendor maintains a Big Data portfolio. - At present, the market is largely dominated by hardware sales and professional services in terms of revenue. - Going*

(a) *The banking sector:* the banking sector very much relies on Big Data, with trading platforms generating and facilitating the exchange of high volumes of data (regarding trading in securities, derivatives and other financial instruments) between index providers, data re-vendors, asset managers, banks and brokers, based on market practice rules and contractual regulations (above all, license agreements).[42]

(b) *The insurance sector:* Big Data enables risk to be assessed much more precisely by reference to specific data about the insured, and hence enables the premium to be calculated more accurately. Insurers can source data by a number of telematics, location-based services, home sensors and wearables that fall within the definition of Internet of Things in the paragraph above.

(c) *The air transport industry:* airlines have generated and hold vast amounts of data about customers' preference during all stages of their journey (from the ticket purchase to arrival at the airport destination), so that the players that are capable of predicting the air travel preferences will obtain a particular competitive advantage.[43] For

---

forward, software vendors, particularly those in the Big Data analytics segment, are expected to significantly increase their stake in the Big Data market. - By the end of 2020, the author expects Big Data software revenue to exceed hardware investments by over $7 Billion" (http://www.researchandmarkets.com/research/9zv8f6/the_big_data).

[42] *Ibid*., p. 484. The Author comments that "*as an alphabet spaghetti of new rulebooks finally emerges from the 2008 financial crisis, the financial instrument trading regime that has applied to equities across the EU since 2007 will shortly be extended to most other asset classes by MiFID II. MiFID II effectively takes MiFID I's regulatory template for public price transparency for equities and extends it to the secondary market for bonds, OTC derivatives and most structured finance products. It makes its contribution to the dawning era of Big Data by requiring pre- and post- contract price data to be disclosed and reported to the market for trades an all the securities that it regulates. As was the case for MiFID I and equities after 2007, MiFID II is likely to lead to hefty growth in the market data world. The degree of transformation that the new rulebooks are imposing, no just on IT platforms and data but across the whole spectrum of financial instrument trading, sets the scene for widespread adoption of Big Data techniques in the banking sector as trading operations and procedures that have developed incrementally since the onset of computerised trading in the 1970s are re-written to comply with the more prescriptive.*"

[43] See http://www.sita.aero/content/big-data-big-insights.

example, Farecast's crunchbase took Big Data files of airline ticket prices relative to days before the flight to be able to calculate the optimum time for flight purchase: it analyses 200 billion flight price records to make its predictions, saving passenger an average of USD 50.00 a flight;[44]

(d) *The entertainment industry:* the entertainment industry is in the midst of a digitalisation process, whereby online consumption and streaming are replacing the traditional way of enjoying audiovisual copyrighted works. With supply and demand increasingly operating online on a global scale, Big Data will enable existing structured datasets to be combined with unstructured data from sources like social media and mobile to predict the "*next best thing.*" For example, Netflix uses Big Data to construct the perfect television series based on its customers' specific preferences.[45]

---

[44] See https://www.crunchbase.com/organization/farecast#/entity.

[45] See http://www.salon.com/2013/02/01/how_netflix_is_turning_viewers_into_puppets/: "*Every single day, Netflix, by far the largest provider of commercial streaming video programming in the United States, registers hundreds of millions of such events. As a consequence, the company knows more about our viewing habits than many of us realize. Netflix doesn't know merely what we're watching, but when, where and with what kind of device we're watching. It keeps a record of every time we pause the action — or rewind, or fast-forward — and how many of us abandon a show entirely after watching for a few minutes. … For at least a year, Netflix has been explicit about its plans to exploit its Big Data capabilities to influence its programming choices. «House of Cards» is one of the first major test cases of this Big Data-driven creative strategy. For almost a year, Netflix executives have told us that their detailed knowledge of Netflix subscriber viewing preferences clinched their decision to license a remake of the popular and critically well regarded 1990 BBC miniseries. Netflix's data indicated that the same subscribers who loved the original BBC production also gobbled down movies starring Kevin Spacey or directed by David Fincher. Therefore, concluded Netflix executives, a remake of the BBC drama with Spacey and Fincher attached was a no-brainer, to the point that the company committed $100 million for two 13-episode seasons.*" Besides, Netflix released an anonymised dataset containing the movie rental histories of approximately 480,000 of its customers: researchers established that they could re-identify some of the Netflix customers at issue by accessing and analysing publicly available information concerning movie ratings performed by such customers. The Netflix contest eventually led to a lawsuit against the company and regulatory scrutiny from the Federal Trade Commission; see A. NARAYANAN AND V. SHMATIKOV, *Robust De-anonymization of Large Datasets (How to Break Anonymity of*

(e) *The healthcare sector:* also the healthcare sector has undergone massive data-driven innovations over the past few years,[46] with the goal to simplify the collection and analysis of information from multiple sources, based on the principle that each patient may enjoy tailor-made treatments if the stakeholders are in a position to structure data provided by hospitals, laboratories, and physician offices. Big Data may eventually help find a cure to cancer and prevent other diseases.[47]

(f) *The public sector:* Government departments across the World are growing digital databases which contain the information that States know about citizens. This so-called data estate is a valuable asset to Governments, but also raises concerns from various standpoints (*e.g.* protection, growth, maintenance, monetisation of databases by a State; compliance with privacy laws; respect of fundamental freedoms) and requires to find a balance between a State commercial interests and the maximisation of benefits of technological progress for citizens.

(g) *Marketing:* horizontal markets are also widely affected by Big Data: providers like Amazon and Target can forecast with a reliable degree of certainty when their customers will order a given product, *before*

---

the Netflix Prize Dataset, the University of Texas, 2008 available at https://arxiv.org/PS_cache/cs/pdf/0610/0610105v2.pdf.

[46] The 2013 report issued by McKinsey & Co on *The Big Data Revolution in the US Healtcare* (available at http://www.mckinsey.com/industries/healthcare-systems-and-services/our-insights/the-big-data-revolution-in-us-health-care) acknowledges that pharmaceutical companies have been aggregating years of research and development data into medical databases, providers have digitised their patient records, public stakeholders have shared their health-care knowledge and clinical trial results on patients covered under public insurance programs.

[47] A well-known attempt to use Big Data in the healthcare sector is Google Flu Trends, a web service operated by Google, which provided estimates of flu activity: by aggregating more than 50 million keywords entered in the form of Google search queries, it attempted to make accurate predictions about flu activity. D. BUTLER, *When Google Got Flu Wrong*, 494 Nature 155, 155-56 (2013), available at http://www.nature.com/news/when-google-got-flu-wrong-1.12413 describes how Google Flu Trends ultimately provided misleading information due to undetected biases in their practices.

the customer places the order, and send them dedicated marketing materials earlier on in order to win the business.[48]

(h) *Legal services:* Big data analytics may prove to be an effective tool to analyse the business of providing legal services. The provision of certain legal services is about using data (*e.g.* case law precedents, decisions, as well as law firms' templates and know-how) to make predictions about outcomes of trials. On this basis, the market already offers certain tools that rely on Big Data analytics or large datasets for use by lawyers: Lex Machina (which analyses large datasets to try to predict the likely outcome of intellectual property cases), KMS Technology (which undertakes an analysis of the structure and language contained in agreements with the objective of drafting new agreements and auditing and reviewing agreements more efficiently), and Judicta (which converts unstructured case law into highly structured data for predictive purposes) are recent examples.[49]

4. Big Data: infinite potential *versus* intellectual property, privacy, competition issues

The dichotomy – which is going to be quite recurring in this paper – between the user's and the service provider's rights on data poses questions that trigger a variety of legal disciplines.

At the opposites of the above-referred dichotomy are the user's exercise of his right to Internet access and the service provider's expectation to protect its investment in Big Data: lawmakers and jurists are therefore required to provide sensible, interdisciplinary, responses to the brand new issues arising in the Big Data era. A flexible approach is very much needed, because, while

---

[48] See http://www.nytimes.com/2012/02/19/magazine/shopping-habits.html.

[49] See M. REBEIRO, M. EVANS, *Big Data: protecting rights and extracting value*, in *Practical Law*¸ 2015, p. 4, also available at http://www.nortonrosefulbright.com/files/-big-data-protecting-rights-and-extracting-value-125453.pdf.

establishing a (difficult) balance between the user's and the service provider's rights, there remains a risk that (privacy and competition) regulation may be improperly invoked to justify the infringement of intellectual property rights, thereby frustrating the service provider's investment in data and innovation.

The following §§4.1, 4.2 and 4.3 therefore aim at presenting certain main issues that may arise in connection with Big Data and which will be analysed – with a particular focus on the intellectual property aspects – throughout this paper.

4.1 Big Data and (Intellectual) Property

The retention and dissemination of data on the Internet pose a number of questions: from the user's perspective, questions concern the user's right to access the Web while maintaining control over his personal, digital data; turning to service providers, which process and retain data that is generated and/or communicated into the Web, issues may arise when seeking legal protection for the datasets they compile and exploit, with a view to maximising their investment in data mining[50] and data analytics.

The difficulty lays with the status of national and international regulations, and the lawmakers' ability to keep the pace of the constant growth (in scale and complexity) of this «*web market of information*,» as I would define the Internet digital marketplace, where information is communicated at an incredibly fast rate, and where large datasets are collected, analysed and, more generally, exploited by those who possess the necessary skills and resources to extract value from those large amounts of data, so that it becomes critical for lawmakers to provide modern solutions aimed at finding a satisfactory balance between the conflicting interests pursued by the users, who *inject* data

[50] According to Wikipedia, «*Data mining*» is the "*computing process of discovering patterns in large data sets involving methods at the intersection of artificial intelligence, machine learning, statistics, and database systems. It is an interdisciplinary subfield of computer science.*"

into the Web, and by the service providers, which *extrapolate* data from the Web.

Commentators have observed that the draft *Declaration of Rights in the Internet* constitutes an improvable effort to find that balance, by proposing "*a precise cultural model, which aims at protecting all rights, to information, privacy, etc., except one, the property right. [The draft] Constitution ... ignores the property rights, as if property were not to be protected, or it did not even exist, on the Web.*"[51]

As most of the decisions mentioned in §2 above suggest, the key to balancing the user's and the service providers' rights is to establish the cases where the user's right to access and privacy shall not unjustifiably deny the enforcement of proprietary rights, including those rights acquired by the service provider on the data it collected (or by the subsequent assignee of said data). From this standpoint, the draft *Declaration of Rights* of the Internet would appear to give priority to privacy to the highest possible extent, by establishing that the exceptions to privacy protection shall apply restrictively.[52] On the one hand, this declaration does not appear to be very much in line with what normally happens on the Web, where the user's data circulates almost freely, and, on the other hand, there is a risk that subjecting the power for courts to interfere with the freedom of electronic communications to very narrow conditions may result in indiscriminately denying the protection and enforcement of intellectual property rights in the name of privacy.

---

51 P. PAGANINI, *Verso l'Internet delle cose*, in *Dir. ind.*, 2015, p. 111.

52 P. PAGANINI, *id.*, p. 112 comments that Article 7 of the draft Constitution sets out that "*any individual's informatics systems and devices and the freedom and secrecy of his information and electronic communications are inviolable. Derogations are possible under the circumstances and modalities established by a provision of law and with the motivated authorisation of the judicial authorities.*" While this protection is of course of essence for those who exercise the right/duty to information, it is a material boundary to the right owner who wishes to enforce his property rights, so that there is a risk, once again, that "*data protection [may be] an argument often abused by those who are responsible for illegal conduct to the detriment of third parties' proprietary rights.*"

Big Data-oriented business organisations have a lot to lose in this scenario. These businesses in fact exploit a number of assets which are instrumental to the analysis and storage of Big Data, *i.e.* (i) the «*platform infrastructure*» – consisting of hardware (servers, storage, user devices, routers, *etc.*), software and networks (local network, Internet connectivity) – and (ii) the «*information architecture*» – *e.g.* data model, structure, design, schemes, format (all as distinct from any data content, such as Big Data) – which governs data flows, and therefore constitutes the intermediate step between the platform infrastructure and data itself.

All the abovementioned assets are suitable to enjoy protection in the form of intellectual property rights: the infrastructure triggers traditional software copyright issues ("*rights in computer languages, software «look and feel», etc.*") and the interrelationships between copyright and database right in relation to database software and accessing and extracting the data held in the software, while the documentation describing the information architecture will normally enjoy the copyright protection, with the structure being protectable by copyright in the European Union under Art. 3 of directive 96/9/EC on the legal protection of databases.[53]

When it comes to dealing with data itself, the main intellectual property rights that afford a level of protection to data remain, in general terms, copyright, database rights and confidentiality, while, as said, patents and rights to inventions can apply to software and business processes that manipulate and process data, but generally not to data itself. Trademarks can apply to data products, but, again, generally not in relation to data.

The enforcement of rights on data, however, emerges to be quite a complex task for data holders.

The very source of Big Data's huge potential is somehow also the cause for its weakness: Big Data exists as long as there is disclosure (see further

[53] R. KEMP, *Legal aspects of managing Big Data*, in *Computer Law & Security Review*, 2014, pp. 486-487.

chapter III, §2.1), communication and sharing of data, which is the result of considerable investment in a multiple-layer process, consisting in the mining, structuring, and analysis of huge datasets. Data sharing, however, relies on the assumption that, where necessary, the data holder is in a position to enforce its proprietary rights in data.

The matter is to establish who is the legitimate data holder, and what are the boundaries to which the assignment and use of data must be subject.[54] But how can someone claim ownership on data?

As further discussed *infra* (chapter III), the difficulty in answering this very practical issue lays with the nature of data itself, which cannot be *per se* "*owned.*"

A «*dataright holder*»[55] may in fact face complex legal hurdles when trying to enforce their rights in data it has accumulated, the reason being that, as a general principle, **there is no property *in* data**: from a purely civil law standpoint, no enforcement can be legitimately/effectively pursued, unless the enforcing party proves *ownership* (or other title) on the asset, the rights on

---

[54] C. GALLI, *I diritti IP nel mercato globale e nella nuova economia digitale: le ragioni di un Convegno*, in *Dir. ind.*, 2015, pp. 105-106. The Author considers that these questions are still unanswered at least by the Italian Courts: the Administrative Court of the Lazio Region in 2014 filed a petition with Constitutional Court seeking the latter's assessment on the compliance with the Constitution of the Communications Authority's (AGCOM) regulation on web piracy: while the Administrative Court alleged that the intellectual property rights may be threatened and put at risk by the intervention of an independent Authority, the Administrative Court apparently failed to raise any concerns regarding the ISPs and the Social Networks' technical possibility to intervene at any stage of the data creation, dissemination and collection (see T.A.R. Lazio, sez. I, ruling dated 26 September 2016, no. 10016, available at http://www.foroitaliano.it/wp-content/uploads/2014/09/Tar-Lazio-10016-14.pdf). The Italian Constitutional Court ultimately rejected the Administrative Court's petition on procedural grounds, thereby abstaining from analysing the merits (see *Corte Costituzionale*, decision no. 247 of 21 October 2015, available at www.cortecostituzionale.it/stampaPronuncia-Servlet?anno=2015&numero=247...PDF).

[55] From now on I will use «*dataright holders*» to identify those companies who gather and process Big Data. This wording, which is used in M. MATTIOLI, *Disclosing Big Data*, in *Minnesota Law Review*, 2014, pp. 535-583, intentionally avoids employing the verb "*own*," to reflect the principle that information cannot be *owned*. See *infra* in this paragraph.

which he is trying to enforce. Put it very simply, the service provider which collects data may not ultimately claim that such data is proprietary to itself.

At the outset of the analysis of a possible legal framework governing the allocation of rights on Big Data, it is probably worth questioning whether the traditional legal schemes of civil law constitute an effective parameter for the analysis of the current market of information. The relevance of this question lays with the necessity to prevent the risk that civil law may result in an unjustified obstacle to the enforcement of intellectual property rights, [56] the aim of intellectual property law being, among others, to support and protect the investments of those who have the skills and resources to seek innovation.[57]

The answer is that the traditional legal schemes are probably not entirely reconcilable with the transactions that take place into the digital world, and that the constant interconnection between society, technology and human relationships has now made it increasingly necessary for jurists to re-think the basic notions of civil law.

The circumstance that the devices in the Internet of Things era are interconnected for a reason (*i.e.* to retrieve and analyse data) implies that intellectual property law shall now aim at protecting not only hardware equipment and software platforms, but also the input and output thereof, *i.e.*

---

[56] C. GALLI, *Il diritto d'autore e la tutela della proprietà industriale sulla rete di Internet*, in AA.VV., *Internet e diritto civile cit.*, p. 189, comments that "*the law and the web appear to move at different speeds, because the width and pervasivity of the rights which can be exercised are jeopardised by the extreme difficulty to enforce.*" There emerges a "*necessity to re-think the IP exclusivity rights by virtue of a synergic, and not conflictive, relationship between exclusivity, competition and contract, and ultimately between IP law and private law … avoiding to artificially create subjective positions that are unrelated to actual liabilities.*"

[57] For an analysis of how intellectual property law can reward investment in innovation (especially in the field of computer-implemented inventions and databases) and provide legal tools to reduce the business risks deriving from investments see C. GALLI, *Diritti di proprietà intellettuale e remunerazione degli investimenti*, (Relazione al Convegno "IP e costituzioni", Università di Pavia, 23-24 settembre 2005), in *AIDA*, 2005, pp. 68-79; R. ROMANO, *Innovazione, rischio e "giusto equilibrio" nel divenire della proprietà intellettuale*, in *Riv. dir. civ.*, 2015, pp. 532-553.

data and algorithms.[58] So, as mentioned above, while there can no be property *in* data – because data falls outside the trivial notion of *asset* that can be *owned*, so that the data holder may not enforce rights in data based on an alleged ownership right – the data holder (so long as it has collected data lawfully) shall however be granted *access* to those datasets it created and analysed *on an exclusive basis*, in return for the investment that it made on data mining and data analytics.[59]

Therefore, any conclusion that the holder of data cannot simply enforce rights *in* data may be just one side to this story, because **rights *arise in connection with* data**, such latter rights being suitable to enjoy the protection afforded by intellectual property law.[60]

4.2 Big Data and Privacy

Also regulation plays an increasingly important role towards the definition of a legal framework for Big Data.

---

[58] P. PAGANINI, *Verso l'Internet delle cose cit.*, p. 111.

[59] According to P. PERLINGIERI, *op. cit.*, pp. 338-342 a distinction shall be made between data which vests the right owner with exclusivity rights (*e.g.* patentable information, or author's works), and data the economic value of which lays entirely with the data's suitability to be traded in the context of a transaction, in consideration of the competitive advantage the data can give to the data holder. Under this perspective, any assessment on who is entitled to use the latter data (*e.g.* collected and decompiled in a data base), it is of essence to determine who is granted access to such data, and not just who is the owner thereof. P. PAGANINI, *op. cit.*, p. 112 poses the following practical question: "[imagine] *a fridge which communicates to the insurance company the fats that we consume without any need. Who is the owner of the data throughout the entire process: the user, the fridges' manufacturer or the manufacturer of the device we bring with ourselves, or else the insurance company? Putting ourselves in the shoes of the industry looking ahead, these will be the matters to be dealt with. It is not a purely legal matter, on the contrary, it is the product itself that we are studying and marketing.*"

[60] With this in mind, contracting for data likely becomes one of the most efficient means to exploit Big Data, as contract law confers strong, enforceable rights and imposes strong, enforceable obligations, although contracts are of course *in personam* (*inter partes*) – unlike rights, which are *in rem* (*erga omnes*). Contracting for data is one of the possible rewards sought by businesses investing in Big Data, as a data holder "*is entitled in principle to impose a charge for use of its data by users whether or not it has IP rights in respect of that data*" according to the UK High Court's decision in *Attheraces Ltd & Another v. The British Horse Racing Board* [2005] WEHC 3015 (Ch).

Big Data is the "Holy Grail" of marketing, because it enables service providers to profile users.[61] With this in mind, one of the most significant – and probably the most evident – legal challenges associated with Big Data, especially on the consumer marketing side, is privacy.

To satisfy the principle of notice and awareness,[62] the data subject must be made aware of the uses to which his or her personal information will be put, and to whom such personal information will be disclosed. The notice is intended to allow the data subject to make an informed choice as to the collection and use of the subject's personal information, and to consent (or not) to that collection and use.

In a Big Data world, there is a risk that the provisions of notice and consent may be circumvented: a data subject who agrees to the provisions set forth in a privacy policy that his personal data may be collected, used, and disclosed for "*marketing purposes*" may not understand that such data may end up being stored in the third parties' databases and combined and disclosed in ways that do not manifestly appear to fall within the scope of the possible data dissemination represented under the privacy policy.[63]

The problem here is twofold.

---

[61] D. NAVETTA, *op. cit.*: "*Big Data can allow marketers to target customers precisely and efficiently by providing advertising and product and services offers that are specifically tailored to a particular individual, based on his or her attributes. Big Data combined with the use of mobile devices can result in offers to individuals that are highly relevant, delivered at the right time, and (with mobile and geo-location tracking) at the right place.*"

[62] See Article 7 of Italian legislative decree no. 196 of 30 June 2003 (the *Italian Data Protection Code*).

[63] D. NAVETTA, *op. cit.*, makes the following example: "*if an ecommerce vendor disclosed to a marketer that an individual customer purchased a deep fryer, such information could be combined into a profile about the individual in a database owned by a data broker. If the data broker later sells access to the database to a health insurance company, whose algorithms put people who purchase deep fryers into a high risk category, in the world of Big Data the initial, relatively innocuous data disclosure (that was consented to) could suddenly serve as the basis to deny a person health care (or result in higher health care rates).*"

The first issue concerns data dissemination: as said, the consumer may not prevent his data from being combined with other existing profile data in a manner that reveals more about the person than contemplated at the time of disclosure (*e.g.* data provided to an insurer for the purposes of calculating the price of the policy may ultimately disclose to third parties the data subject's data about genetic pre-disposition to illnesses). The data subject may ultimately lack an understanding of the interpretations, inferences, and/or deductions that may be draw from his combined data using Big Data mining techniques and analytics.[64]

Other privacy-related concerns relate to the data subject's right to access his or her data to ascertain whether it is accurate and complete and seek for their update or deletion.[65] Except for the established and highly visible data controllers, the general public does not know what entities may be collecting information about them and creating profiles. While data subjects may be able to identify companies to whom they have provided personal information, and may have a direct relationship with such companies, the same is not necessarily true in the case of data aggregation companies, also known as «*data brokers.*» In most cases data subjects do not have a direct relationship with them and brokers typically do not receive information directly from the data subjects. Even if a consumer can identify a data broker that holds his or her profile, without a contract the consumer may have a hard time seeking access to his or her personal information.[66]

---

[64] According to D. NAVETTA, *op. cit.*, a related issue may be the manipulation issue: Big Data (especially those inferred from digital data) are frequently infused with the data subject's opinions, which opinions subsequently mix with the subjective judgments of those who collect, organise and analyse it. One technique for mitigating privacy-related risks associated with Big Data may be anonymisation, aimed at preventing data controllers from individually identifying the persons to whom the dataset relates. This technique may allow organisations to work with Big Data while mitigating privacy concerns, and has been used in many sectors, including healthcare, banking and finance and online advertising.

[65] See, again, Art. 7 of the Italian Data Protection Code.

[66] Besides, many data brokers are able to pinpoint a user's identity and specific preferences without having any information traditionally considered personally identifiable

The EU General Data Protection Regulation 2016/679 ("GDP Regulation") – which entered into force on 24 May 2016, repealing directive 95/46/EC, and which will be directly enforceable in all the EU Member States starting from 25 May 2018 – provides a proposed solution to this issue, by imposing a duty over data controllers to make data available in a structured, commonly used, machine-readable and interoperable format that allows the data subject to transfer data to another controller, if the data subject so wishes (so-called «*data portability*»). [67] As former European Commission vice-president Joaquín Almunia emphasised during an event organised by the Privacy Platform in 2012, "*data portability is at the heart of competition policy*," as lack of data portability may amount to an abuse of dominance under Art. 102 of the Treaty on the Functioning of the European Union.

4.3 Big Data and Competition

Given the increasing interplay between privacy and competition, the European Data Protection Supervisor issued a report underscoring that competition law should intervene to address privacy and data-related concerns and perceived regulatory gaps.[68]

---

information: the information provided by the mere access, via mobile or personal computer, to a given website, when combined with specific behavioral and other data, can supply enough information to identify a person individually, so that the collection and aggregation of seemingly harmless data about a person may be used to reveal sensitive information (*e.g.* health status, sexual orientation and financial status) without the prior data subject's consent. In this regard, see B. SEGALIS, N. SHAH, *FTC Looks to Link Do-Not-Track, Big Data Privacy Concerns; Seeks Solutions*, available at http://www.infolawgroup.com/2012/03/articles/data-privacy-law-or-regulation/ftc-looks-to-link-donottrack-big-data-privacy-concerns-seeks-solutions/.

[67] See Art. 20 DGPR. See also the *Guidelines on the right to data portability* issued by Article 29 Data Protection Working Party on 13 December 2016 and I. GRAEF, *Data portability at the crossroads of data protection and competition policy*, available at http://www.agcm.it/component/joomdoc/eventi/convegni/20161109_07.pdf/download.html.

[68] See the European Data Protection Supervisor's March 2014 *Preliminary Opinion on Privacy and competitiveness in the age of big data: The interplay between data protection, competition law and consumer protection in the Digital Economy* (https://edps.europa.eu/sites/edp/files/edpsweb_press_releases/edps-2014-06-big-data_en.pdf). With reference to the interplay between privacy and competition laws A. LAMADRID, S. VILLIERS, *Big Data, privacy and competition law: do competition*

The above-referred European Data Protection Supervisor's Opinion follows the path traced by EU case law precedents, pursuant to which competition law should address (also) privacy issues: in *Google/DoubleClick*, the Commission assessed whether the combination of the parties' databases would impede competition; privacy issues were raised but did not play any decisive role in the Commission's decision. In *Facebook*/*WhatsApp*, the Commission analysed whether Facebook could use WhatsApp as a potential source of user data to improve its advertising, but concluded this was not the case to a point that it could hamper competition. Outside merger control, the Commission considered that codes and structure of databases could be an essential facility to which access should be given in cases such as *IMS Health* and *Reuters*.[69]

This resulted in an increasing concern of competition agencies that, if companies acquire unique datasets that others cannot replicate, an harm to competition may occur, because data may raise barriers to entry through the creation or strengthening of market power; hence, if a dataset is considered unique and 'essential' for competing businesses, a dominant company may be required to grant access to rivals to avoid potential abuse of dominance cases. Some others argue that data is infinite in quantity and easily replicable, and therefore exclude competition concerns.[70]

---

authorities know how to do it?, in *CPI Antitrust Chronicle*, 2017, p. 7 underscore that "*it is relatively common for agencies in Europe and plaintiffs in the U.S. to think of competition law as a hammer suitable for all sorts of nails.*"

[69] See Commission Decision of 11 March 11 2008 in case No. COMP/M.4731 – *Google/ DoubleClick*; Commission Decision of 3 October 3 2014 in case No. COMP/M.7217 – *Facebook/Whatsapp*; Commission decision of 3 July 3 2001 in case No. COMP D3/38.044 – *IMS Health*; Commission decision of 20 December 2012 in case No. case AT.39654 – *Reuters Instrument Codes*; Commission Decision of 6 December 2016 in case No. COMP/M.8124 – *Microsoft/LinkedIn*.

[70] See A. MANTELERO, *Big data: i rischi della concentrazione del potere informativo digitale e gli strumenti di controllo*, in *Dir. inf.*, 2012, pp. 135-144; ID., *Observatory on ICT law: the control of over digital information in the Big Data Era*, in *Contratto e impresa. Europa*, 2012, pp. 961-966; G. MUSCOLO, *Innovazione nella rete e diritti non titolati: il ruolo di know-how, copyright, banche dati e pratiche commerciali sleali*, in *Dir. ind.*, 2015, pp. 114-116; Freshfields Brukahus Deringer 2017 Antitrust Newsletter,

In March 2016, the *Bundeskartellamt*, the German Competition Agency, opened proceedings against Facebook on suspicion of it having abused its market power by infringing data protection rules.[71] The *Bundeskartellamt* raised concerns that Facebook might abuse its dominant position in social networks through its privacy terms and conditions. In particular, the German Competition Agency flagged that Facebook's privacy terms could violate German privacy laws. Although the mere violation of privacy law by a dominant company would not be actionable under antitrust law, the *Bundeskartellamt* will assess whether Facebook's position allows it to impose contractual terms that would otherwise not be accepted by its users.

Then, in May 2016 France's *Autorité de la concurrence*, together with the same German *Bundeskartellamt*, published a joint paper on data and its implications for competition law, whereby the agencies outlined that Big Data may actually benefit consumers by providing better services, but potential harm to competition may derive from monopolisation of data-related markets.[72]

Big Data – Be prepared for closer scrutiny ahead of commercial practices and deals (http://antitrust.freshfields.com/big-data); Jones Day's, *European Antitrust Enforcers Move on Holders of Big Data*, 2016, available at http://kluwer-competitionlawblog.com/2016/05/26/european-antitrust-enforcers-move-on-holders-of-big-data/.

[71] *Bundeskartellamt*, Press Release: *Bundeskartellamt initiates proceeding against Facebook on suspicion of having abused its market power by infringing data protection rules*, 2 March 2016, available at: https://www.bundeskartellamt.de/-SharedDocs/Meldung/EN/-Pressemitteilungen/2016/02_03_2016_Facebook.html.

[72] *Autorité de la concurrence* and *Bundeskartellamt*, *Competition Law and Data*, 10 May 2016, available at: http://www.autoritedelaconcurrence.fr/doc/reportcompetitionlawand-datafinal.pdf. The joint *Autorité de la concurrence* - *Bundeskartellamt* report (a) makes clear that the competitive impact of data depends on many factors that need to be considered on a case-by-case basis, and (b) when discussing possible theories of harm it makes general arguments about how standard theories of harm often used in antitrust could apply to data, applying the same logic that would apply to any other asset.

To an extent, there were follow-ups to the French and German agencies' report also in other EU Countries.

In the **United Kingdom**, the Competition and Markets Authority came to a different conclusion when it investigated the use of data in the insurance sector. In finding that Big Data was producing a range of benefits for consumers and the use of such data was

The concerns of the European competition agencies are no doubt very sensible, but agencies may probably be mindful that privacy can be a parameter of competition subject to the narrow condition that personal data is actually relevant to the peculiarities of the case. On the contrary, and as further discussed in chapter II below, Big Data is not always personal data and, when it is not, privacy regulation cannot be used beyond its scope of application.

Similarly, when assessing the value or indispensability of a given set of data – which is at the basis of the antitrust judgment – a case-by-case approach may still be the safest way to distinguish data that is rivalrous from data that is not,[73] and that the relevance of data lies not in volume itself, but in the analytics, how it is processed and used. In a nutshell, «*it depends*» may be a frustrating answer, but it might, in this case, be the right one when discussing the opportunity for competition agencies to assess Big Data.[74]

---

working well, it decided that there was no need to launch an in-depth market study. In November 2015, the UK's Financial Conduct Authority (FCA) issued a call for input in relation to Big Data in retail general insurance. The FCA's inquiry in part seeks to understand how Big Data could affect competition in retail insurance products, in particular private motor and home insurance, and how this could affect consumers. It is due to publish a feedback statement with its initial findings later this year. Those findings could include a more in-depth competition investigation.

In **Italy** and **Spain** there has been no case to date focusing on Big Data. However, in 2009 the Spanish Competition Authority in 2009 issued decisions (upheld by the Supreme Court) against several monopolist electricity providers, finding that they had abused their dominant positions by delaying access to their customer databases to new entrants; and, remarkably, on 17 June 2016 the Civil Court of Milan ascertained Google's contractual liability towards its software supplier Attrakt for having Google abused of its economic dominance position and having attempted to eliminate Attrakt from the market (*Trib. Milano*, 17 June 2016, available at http://www.mmlex.it/wp-content/uploads/2016/10/Attrakt-sentenza-ITALIANO.pdf).

[73] R. MAHNKE, *Big Data as a Barrier to Entry*, in *CPI Antitrust Chronicle*, 2015, pp. 3-4: "*There are data, and then there are data. Some businesses are successful and valuable because they have access to data that others cannot easily obtain. For example, Craigslist. When other companies have attempted to scrape its listings, Craigslist has tried to stop them, with success. Is it impossible for anyone to compete with Craigslist in its core business? Maybe not, but anyone trying faces a substantial barrier to entry. Does every online or digital market work like this? Of course not, but some do.*"

[74] A. LAMADRID, S. VILLIERS, *Big Data, privacy and competition law cit.*, p. 9.

CHAPTER II

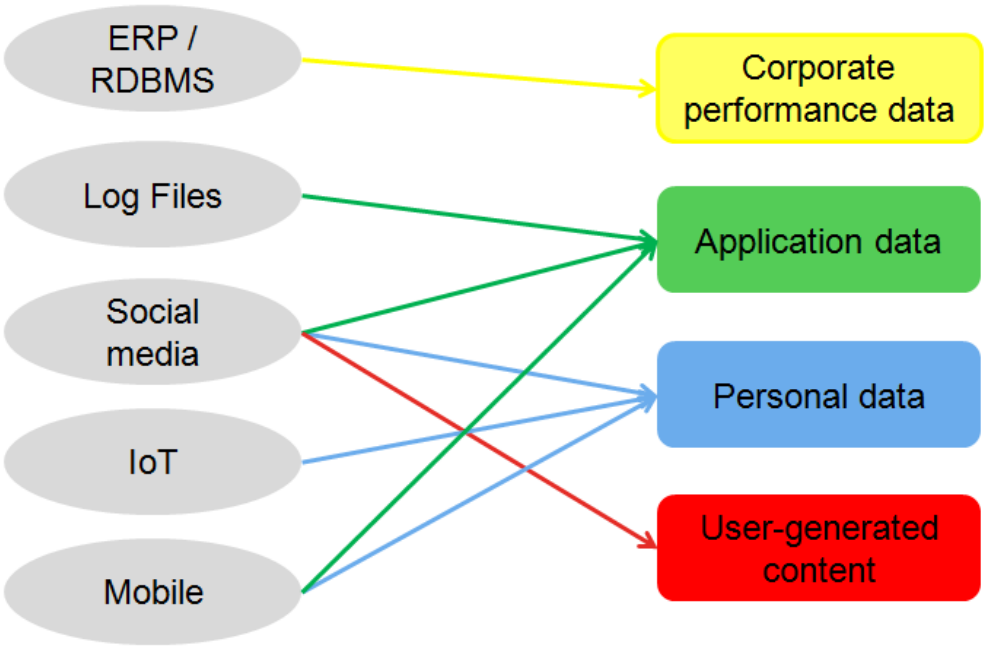
BIG DATA: SOURCING DATA FROM THE WEB

1. Input data: overview

This chapter is concerned with input data – *i.e.* Big Data that the service provider collects in view of subsequent analysis, structuring and use – and the legal aspects that shall be considered when the service providers injects data into its business systems.[1]

As anticipated in chapter I, there are a number of sources of input data: Internet of things (IoT) sensors/machines/applications, corporate management and planning software, social *media*, mobile devices and all sorts of Internet-connected computers.

Fig. 1 – Sources to input data



These sources generate a variety of data, which I would subdivide in four main categories (as depicted in fig. 1 above), namely:

[1] The distinction I drew in chapter I, §3, between input and output aims at explaining Big Data as a two-phase process: (i) an «*upstream phase*,» where Big Data is collected and injected into the business organisation, and (ii) a «*downstream phase*,» which is the result of the analysis of Big Data, and the output of which is structured, meaningful data ready for business use.

(a) *Corporate performance data*, which is probably the most "traditional" category of data – *e.g.* stocks, sales, prices, customers, suppliers, accounting, *etc.* – that companies source from software like ERP (*enterprise resource planning*) and RDBMS (*relational database management system*) to learn how to improve their business strategy;

(b) *Application data*, as I would describe any possible records of a particular activity that was performed by human and/or machine and can be tracked in a log file by an application of IoT sensors, *e.g.* logins and access records (for the activities performed on websites or software platforms); "play, fast forward, rewind, stop" records (which, as mentioned in chapter I, §3, is what Netflix does while we are watching a film); gas and oil consumption, tyres wear, brakes and accelerator usages (that are recorded by black box-equipped cars of the IoT generation); geo-localisation and other tracking records (which now any smart phone or other IoT device can reveal), *etc.*;

(c) *Personal data*, *i.e.*, according to Art. 4 of the GDP Regulation, "*any information relating to an identified or identifiable natural person («data subject»).*" Personal data, therefore, is not just an individual's name and surname, but any kind of information which may lead to identify, directly or indirectly, such individual ("*an identification number, location data, an online identifier or … one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person*"); and

(d) *User-generated content*, *i.e.* any form of content created by users and made available publicly on the system where the content was generated or uploaded, *e.g.* posts, blogs, tweets, videos, wikis or comments.

As far as this chapter is concerned with the legal issues that may arise in connection with the collection and storage of Big Data, my analysis will prevalently focus on the collection of personal data for the following reasons.

Corporate performance data is normally generated in the context of the very same business organisation that will reuse it, so I do not see major issues affecting the dataright holder's right to gather and retain such data. Issues may indeed concern the dataright holder's ability to keep data confidential, a matter that will be dealt with in chapter III.

As for application data, I would make the following distinction: if application data may be exploited as inferences that ultimately identify the user, then application data falls within the definition of personal data, with the consequences that are discussed *infra*. Conversely, the scenario that appears to be more frequent is that machine-generated data is both automatically generated and anonymised: then the collection of data should not in principle trigger any risk of unauthorised collection of data, as such data is at its outset proprietary to the data collector.[2]

User-generated content is not *per se* data. From an intellectual property standpoint, user-generated content falls within the scope of copyrighted work, so the right for the host provider to qualify as the licensee (or assignee) of the rights of economic exploitation on such work is very much governed by contract law. However, as this paper is concerned with Big Data, user generated content will be considered prevalently to the extent that it contains data and, in particular, personal data.

Put it the other way round, I have decided to focus this chapter on input personal data, because the nature of personal data triggers a number of legal issues that make collection and retention of personal data more complex than other data. When the data subject is navigating the Web or using a social network platform, the user is also likely to enjoy consumer and privacy protection in his capacity as both consumer and data subject, so that there is a variety of statutes that intervene when data is collected at a user's level.

---

[2] This approach would be consistent with the copyright principle governing computer-generated works. Pursuant to Art. 9(3) of the **UK** Copyright, Designs and Patent Act 1988, the author of computer-generated works shall be taken to be the person by whom the arrangements necessary for the creation of the works are undertaken.

2. The relationship between the web user and the Internet service provider: market for services paid for by personal information

From a privacy law perspective, a vast portion of Big Data are personal data.

The 2014 Preliminary Opinion of the European Data Protection Supervisor highlighted that personal data the data subject injects into the Web – which data the great social media and other Web players subsequently retain and process – qualifies as the actual «*consideration*» for the services (search engines, e-mail, social networks, *etc.*) that the Internet Service providers (hereinafter also referred to as the «*ISP*») grant on an apparently free of charge basis: the EU report describes this as the "*market for services paid for by personal information*."[3]

Offers of free services abound on the Internet, but the focus on the (absence of a) *price* rather than on the (actual) *cost* of «*free*» services has led consumers into a position of vulnerability.[4]

By way of example, online so-called «*freemium*» games are available free of charge if users register by disclosing personal details. These games monitor online activity to learn how to convert «*free riders,*» *i.e.* non-paying players, into paying customers, or to deliver targeted, more lucrative forms of advertising. A small minority of users pay for additional features, but still a relevant portion of revenue is generated by free riders.[5] This information

---

[3] See *Preliminary Opinion of the European Data Protection Supervisor*, p. 27, available at https://secure.edps.europa.eu/EDPSWEB/webdav/shared/Documents/Consultation/-Opinions/2014/14-03-26_competitition_law_big_data_EN.pdf.

[4] In 2010, the *Wall Street Journal* focused a series of articles on this monitoring, finding that the USA's 50 top websites on average installed 64 pieces of tracking technology onto the computers of visitors, usually with no warning. See J. ANGWIN, *The Web's New Gold Mine: Your Secrets*, the *Wall Street Journal*, 30 July 2010, available at http://online.wsj.com/article/SB10001424052748703940904575395073512989404.html.

[5] See J. HOOFNAGLE, J. WHITTINGTON, *Free: Accounting for the Costs of the Internet's Most Popular Price*, *UCLA Law Review*, 2014, p. 606 *et seq.*, available at http://www.uclalawreview.org/pdf/61-3-2.pdf. The Authors report Disney as an example: like others in the industry, Disney "*places financial value on the number of consumers it identifies, the personal information they provide, and the extent to which Disney can*

allows game producers to then "*fine-tune content in order to lower the cost of loss leaders, raise rates of conversion, reduce the time it takes to convert free riders into paying customers*," and ultimately maximize the lifetime value of each customer, thereby raising average revenue per paying user.[6]

Free offers are so widespread, that consumers tend to try them without serious awareness, to subsequently find themselves bound to contractually agreed services, with constantly changing policies, which generate unforeseen circumstances. The EU report mentions that in the European Union 46% of users of social networking or sharing sites felt insufficiently informed about the possible consequences of disclosing personal information.[7]

The following paragraphs aim at analysing how certain Internet service providers – *i.e.* the social networks, the search engines and the video-sharing websites – amass the users' personal data, with a view to assigning data to third parties and/or studying, elaborating, and using personal information for business-related purposes:[8] §§2.1, 2.2 and 2.3 provide an overview of the

---

track consumer activity in order to modify the game and thus increase the rate of conversion of consumers from free players to paying customers." In the Club Penguin's page, one of the most popular Disney's games, Disney "*explains that «if personal information is not provided to us, then Club Penguin may not agree to Club Penguin membership.» When consumers want to try Disney's Club Penguin game, they must provide a name and email address, which is then authenticated when the consumer responds to the invitation to register sent to that email address. This is a game intended for small children, so the email invitation is addressed to the parents of the child and comes complete with the privacy policy. In the process of registration and play, the consumer's Internet Protocol (IP) addresses are made available to the firm, along with any data the firm chooses to collect while monitoring the player's choices and chats. Moreover, cookies are placed on the player's computers. The policy explains that the firm «operate[s] globally and may transfer your personal information to individual companies of The Walt Disney Family of Companies or third parties in locations around the world for the purposes described in this privacy policy.*»"

6 See J. HOOFNAGLE, J. WHITTINGTON, *op. cit.*, p. 627.

7 See J. HOOFNAGLE, J. WHITTINGTON, *op.cit.*, p. 606 *et seq.*, available at http://www.uclalawreview.org/pdf/61-3-2.pdf. The Authors refer that the UK's Office for Fair Trading has even been investigating in-game app payments and has identified possible consumer law breaches, not least in the use of what may amount to the use of "emotional blackmail."

8 R. CATERINA, *Cyberspazio, social network e teoria generale del contratto*, in *AIDA*, 2011, p. 93, and M. GRANIERI, *Le clausole ricorrenti nei contratti dei social network dal punto di vista della disciplina consumeristica dell'Unione europea*, in *AIDA*, 2011, p.

aforementioned service providers' terms of service, which are subsequently analysed as to the applicable law (§3), the rights granted by the user on his data to the service provider (in exchange for the right to use the provider's software) (§4), and the service provider's right to retain data subsequently to the termination of the agreement with the user (§5).

2.1 Introduction to the social networks' terms of service

The notion of social network, at least for the purposes of this paper, includes web-based services which act as online sites to connect people with similar interests, activities, backgrounds or real-life connections, with a view to building social networks or social relations with other people.

Although each social network has its own peculiarities, the most widespread platforms – such as Facebook, Twitter, Instagram, LinkedIn – share almost the same common features:

(a) "*Social media services are (currently) Web 2.0 Internet-based applications;*

(b) *User-generated content is the lifeblood of social media;*

(c) *Individuals and groups create user-specific profiles for a site or app designed and maintained by a social media service;*

(d) *Social media services facilitate the development of social networks online by connecting a profile with those of other individuals and/or groups.*"[9]

---

127 highlight that, form a procedural standpoint, a case-by-case analysis of the terms of service of each and every social network (and the same applies to the other web-based services) is not viable, due to the increasing number of social networks currently on the market, and on the peculiarities each of them features, so that any legal analysis shall be cantered around the features that are common to the majority of the social networks, and it shall then be checked whether the outcome of the general analysis can be suitably applied to the social network at stake.

9 J.A. OBAR, S. WILDMAN, *Social media definition and the governance challenge: An introduction to the special issue. Telecommunications policy, Quello Center Working Paper no. 2647377*, p. 745; D.M. BOYD, N.B. ELLISON, *Social Network Sites: Definition, History, and Scolarship*, in *Journal of Computer-Mediated Communication*, 2007, p. 211.

The social networks are digital environments, where countless communications are conveyed in a brand new fashion of contents and technical schemes, which pose legal questions not only concerning the unprecedented, legally relevant, situations social networks generate online, but also as to the consequences the social networks give life to offline, *i.e.* in the real world, due to the way the social networks have changed the communication and relationships between individuals.

AGCOM (*Autorità Garante della Comunicazione*), the Italian regulator in the field of communications, has recently acknowledged that the digital world has in fact become a preferential environment for individuals to communicate and build relationships and, against this background, AGCOM concludes that is not only the so-called «*Voice over IP* services» (*i.e.* the delivery of voice communications and multimedia over Internet Protocol networks, normally the Internet) that have entered the telecommunications market, but also the social networks, which "*not only influence personal and private relationships, but also interact with the traditional* media, *thereby creating brand new, hybrid, services (social tv, social news, social advertising).*"[10]

The user is, however, allowed to provide his contributions within the technical boundaries imposed by the Internet service provider's platform[11] and

---

[10] By way of example, the traditional voice broadcasting system has changed as a result of the innovations brought by providers like Skype, Viber, Google Voice with the *Over-The-Top* (OTT) services, *i.e.* audio, video, and other media transmitted via the Internet without an operator of multiple cable or direct-broadcast satellite systems distributing the content. Similarly, the *Short Message Service* (SMS) has been progressively replaced by the freeware, instant messaging applications (WhatsApp, WeChat, Facebook Messenger). See AGCOM's 2014 report, p. 32.

[11] L. LESSIG, *Code Version 2.0*, available at http://codev2.cc/download+remix/Lessig-Codev2.pdf P. SAMMARCO, *Le clausole contrattuali di esonero e trasferimento della responsabilità inserite nei termini d'uso dei servizi del web 2.0*, in *Dir. inf.*, 2010, p. 634. There emerges a situation whereby the relationship between the user and the ISP is from its outset characterised by a disparity in principle: while the user has a limited freedom to operate, and very limited visibility on what happens to his data/content once he has disclosed them to the ISP, the ISP is in a position to regulate and manage the IT infrastructure.

in accordance with the terms of service set forth therein, which govern the relationship between the Web user and the social network.

From a wider viewpoint, this means that – in the lack of an intrusive intervention by courts and lawmakers – the rules governing the Web are being increasingly dictated by the Internet service providers, which are in a privileged position to both *regulate* and *manage* the web-based communication flows. Most of the services that are offered online, in fact, feature brand new technological applications, which – from a purely civil law perspective – are unprecedented and likely to fall outside the statutory legal schemes.[12] The service providers' terms of service therefore aim at providing some sort of «*full coverage*» regulation for the services they offer, in such a way as to prevent court litigation to the maximum possible extent.[13] Consistently, terms of service also exclude that either party shall be considered to be contractually bound to provide the core services, so that no liability may arise in connection with the social network's failure to provide the social networking services.[14]

The vast majority of the social networks' terms of service provide for the grant by the user of a non-exclusive, transferable, royalty-free, worldwide

---

[12] P. PERLINGIERI, *Nuovi profili del contratto*, in *Riv. crit. dir. priv.*, 2001, pp. 229-230.

[13] G. DE NOVA, *op. cit.*, p. 3 underscores that nowadays a relevant number of contracts are intended to be "auto-sufficient," meaning that the parties' goal is to regulate all possible circumstances arising out of, or in connection with, the agreement, despite the fact that they have identified Italian law as the applicable law.

[14] F. ASTONE, *Il rapporto tra gestore e singolo utente: questioni generali*, in AIDA, 2011, pp. 122-123; R. CATERINA, *Cyberspazio, social network e teoria generale del contratto*, in *AIDA*, 2011, p. 100, according to whom "*The contractual conditions imposed by the main social network websites mainly aim at preventing that the relationship with the user may result in judicial litigation.*" S. SCALZINI, *I servizi di online social network tra privacy, regole di utilizzo e violazione dei diritti dei terzi cit.*, p. 2574 comments that the social network agreements differ from any other agreements, as confirmed, for example, by the fact that the amendment of the Facebook's terms of service regarding the possibility for Facebook to share the users' information with Facebook's affiliates, such as Instagram, was the subject matter of a *referendum* amongst the Facebooks users, whereby Facebook undertook that "*if more than 30% of all active registered users vote, the results will be binding*" (see https://newsroom.fb.com/news/2012/12/responding-to-your-feedback/). Only 2% of the Facebook users ultimately voted, resulting in Facebook autonomously taking the decision to share information with its affiliates.

license to the service provider to exploit the data and copyrighted content the user generates while enjoying the social networking services. The user also consents to the further communication of his data to third parties.[15] The license agreement underlying the social network agreement shall terminate following the de-activation (for whatever reason) of the user's account, provided that the service provider is normally entitled to retain a backup copy of the user-generated data.

The service provider, in turn, grants the user a non-exclusive, royalty-free, worldwide license to use the software (which can either be proprietary to the host or licensed thereto by a third party content provider). The software is provided «*as is*,» *i.e.* without any express or implied warranty by the service provider, which accepts no liability should the user be prevented from enjoying the social networking services, or should the account host contents that breach third parties' rights.[16]

As regards the user's conduct, the terms of service provide for the user's undertaking not to take any action that is prohibited under the applicable law (*e.g.* violating third party's proprietary rights, uploading malicious files, such as malware or viruses, or harassing, intimidating or disparaging the other users, uploading pornographic, violent or other prohibited contents, or else uploading third party's sensitive data without the required consents).

The breach of the prohibitions above may result in the service provider deleting the unauthorised content, or in the de-activation of the account.

---

[15] The terms of service usually grant the user the possibility to select the distribution list of content he generates, but in certain other cases the potential addressees are the entire web (see Twitter's terms of service, version dated 30 September 2016, available at https://twitter.com/tos?lang=en: "*This license authorizes us to make your Content available to the rest of the world and to let others do the same*").

[16] S. SICA, G. GIANNONE CODIGLIONE, *Social network sites e il «labirinto» delle responsabilità*, in *Giur. mer.*, 2012, pp. 2715-2733

2.2 Introduction to the search engines' terms of service

The most widely used search engines are Google, Bing (Microsoft), Baidu and Yahoo! Search.[17] Search engines are software systems that enable the user to search for information on the web.

All the most widespread search engines also provide ancillary services, such as data and file storage (*e.g.* Google Cloud Platform, Microsoft OneDrive, Baidu Pan), translation services (*e.g.* Google Translate, Bing Translator, Baidu Translate), e-mail accounts (*e.g.* Gmail, Microsoft MSN, Yahoo! Mail), maps (*e.g.* Google Maps, Bing Maps, Baidu Maps), and news (*e.g.* Google News, Bing News).

While the user of a social network platform is generally not allowed to enjoy the services and interact if he does not register (and login) – so that the acceptance of the social network's terms of service is express and usually considered to be in writing[18] – the user of a search engine is usually led to

---

[17] The abovementioned four search engines total 98.73% of the 2017 market share. Google alone secures an 80.52% share. Source: https://www.netmarketshare.com/search-engine-market-share.aspx?qprid=4&qpcustomd=0.

[18] See M. BASSINI, *Point and click: la tutela del consumatore nel commercio elettronico*, ILSU Working Paper no. 2008-13/IT, available at *web* *www.diritto.it* and A.R. POPOLI, *op cit.*, p. 1001. The execution of the agreement takes place in compliance with the 'point and click' mechanism, whereby a user activates and icon by means of a single or double mouse click. The effectiveness of the acceptance of the terms and conditions set forth under the terms of service of a website is questioned by certain Scholars, according to whom, lacking any form of negotiation between the user and the ISP, the statutory requirement under art. 1321 of the Italian Civil Code – pursuant to which there must be a mutual undertaking between the parties for the agreement to come into force – would not be complied with, so that any provision of services by the ISP to the user would occur in the absence of any agreement (N. IRTI, *Scambi senza accordo*, in *Riv. trim. dir. proc. civ.*, 1998, p. 360); on the contrary, other Scholars underscore that the requirement of Art. 1321 is complied in all cases where the parties' will converge to a unitary design, also in the absence of a previous, express, negotiation (G. OPPO, *Disumanizzazione del contratto?*, in *Riv. trim. dir. proc. civ.*, 1998, p. 538 *et seq.*). This latter opinion would appear to be consistent with the concept of conclusive conduct (*comportamento concludente*) pursuant to Art. 1326 of the Italian Civil Code. Also, the Italian law maker has partially addressed the issue, with the following provisions: (i) legislative decree of 7 March 2005 (*Codice dell'amministrazione digitale*) explicitly rules in favour of the electronic execution of agreements; and (ii) Art. 51, para. 2, of legislative decree of 21 February 2014 no. 21 implementing directive 2011/83/EU on consumer rights explicitly contemplates the possibility to enter into agreement by selecting an icon on the website, which has been interpreted to be an implicit reference to the point and click mechanism.

accept the engine's terms of service by merely using the engine, without the need to take any further action.[19] To enjoy some of the ancillary services (*e.g.* e-mail and data storage) the user must create an account.

Similarly to the social networks' case, the parties to a search engine agreement enter into a cross-license agreement, pursuant to which the search engine provider grants the user a non-exclusive, royalty-free, worldwide license to use the software, while the user grants the right to the ISP – and "*those [the ISP] work[s] with*" – a worldwide license to "*use, host, store, reproduce, modify, create derivative works (such as those resulting from translations, adaptations or other changes [the ISP] makes so that [the user's] content works better with [the ISP's] Services), communicate, publish, publicly perform, publicly display and distribute such content*."[20]

Most notably, the search engine records a number of data, among which:

(a) *Search entries:* the pre-installed settings of every search engine provide that the engine tracks and retains every search, unless the user opts out by toggling the «*tracking off*» option in the activity controls page;

(b) *Voice recordings:* anyone who uses Google's voice search or the voice-activated assistant, Google Now, has his searches stored;

(c) *Location:* the use of the service provider's maps usually result in the provider tracking the location history, matching the places with the dates where the searches were done.

The service provider also reserves the right to use automated systems to "*analyze [the user's] content (including emails) to provide [the user] personally relevant product features, such as customized search results,*

---

[19] See Google's Terms of Service (*Your Content in our Services*) (version dated 14 April 2014), available at https://www.google.com/intl/ALL/policies/terms/: "*By using our Services, you are agreeing to these terms;*" see also Yahoo! Search's Terms of Service (version dated 20 January 2014), available at https://policies.yahoo.com/ie/en/yahoo-/terms/utos/index.htm: "*By using the Yahoo Services, you agree to follow these Terms*."

[20] See Google's Terms of Service.

tailored advertising, and spam and malware detection. This analysis occurs as the content is sent, received, and when it is stored." [21]

Also the search engines' terms of service provide that the license continues even if the user stops using the services.

2.3 Introduction to the video hosting providers' terms of service

Similar provisions govern the cross-license between the user and the video-hosting services,[22] *i.e.* websites which allow users to distribute their video clips (*e.g.* YouTube, MySpace, Bilibili – the most popular video hosting website in China – Vimeo, Metacafe, *etc.*).

YouTube's platform allows users to post comments to the shared videos. The terms of service make it clear that, while the license granted by the user in relation to the video clips terminates upon removal of the video clip from the Website, or deletion of the user's account, the license granted by the user

---

[21] See Google's Terms of Service.

[22] See, for example, the YouTube's Terms of Service (version dated 9 June 2010), available at https://www.youtube.com/static?template=terms&gl=GB:

- Acceptance of the terms of service: "*You can accept the Terms by simply using the Service. You understand and agree that YouTube will treat your use of the Service as acceptance of the Terms from that point onwards*"

- Cross-license agreement: "*When you upload or post Content to YouTube, you grant:*

a. to YouTube, a worldwide, non-exclusive, royalty-free, transferable licence (with right to sub-licence) to use, reproduce, distribute, prepare derivative works of, display, and perform that Content in connection with the provision of the Service and otherwise in connection with the provision of the Service and YouTube's business, including without limitation for promoting and redistributing part or all of the Service (and derivative works thereof) in any media formats and through any media channels;

b. to each user of the Service, a worldwide, non-exclusive, royalty-free licence to access your Content through the Service, and to use, reproduce, distribute, prepare derivative works of, display and perform such Content to the extent permitted by the functionality of the Service and under these Terms.

The above licenses granted by you in Content terminate when you remove or delete your Content from the Website."

in the textual comments he submits is perpetual and irrevocable, without prejudice to the user's ownerships rights, which are retained by the user.[23]

The user further accepts that he may be exposed to other users' comments that are "*factually inaccurate, offensive, indecent, or otherwise objectionable*" to the user, and agrees to waive any legal or equitable rights or remedies he has or may have against YouTube with respect to any such content.[24]

3. The relationship between the web user and the Internet service provider: applicable law

The vast majority of the social networks and the other service providers mentioned in the paragraphs above belong to California-based corporations (*e.g.* Facebook, which recently purchased Whatsapp,[25] LinkedIn, Twitter, Google, Pinterest, MySpace, Yahoo!, *etc.*). This circumstance does have materially legal implications, because it leads to a factual situation whereby the terms of service drafted pursuant to the US law end up circulating worldwide, following a mere translation into the local languages. No material amendments are made based on the local laws of the jurisdictions where terms

---

[23] See YouTube's Terms of Service: "*The above licenses granted by you in Content terminate when you remove or delete your Content from the Website. The above licenses granted by you in textual comments you submit as Content are perpetual and irrevocable, but are otherwise without prejudice to your ownerships rights, which are retained by you.*"

[24] See YouTube's Terms of Service: "*You further understand and acknowledge that in using the Service, you may be exposed to Content that is factually inaccurate, offensive, indecent, or otherwise objectionable to you. You agree to waive, and hereby do waive, any legal or equitable rights or remedies you have or may have against YouTube with respect to any such Content.*"

[25] Facebook purchased Whatsapp in 2014. On its terms of service Whatsapp (version dated 25 August 2016, available at https://www.whatsapp.com/legal/#terms-of-service) represents the following in relation to the flows of data between Facebook and Whatsapp: "*Nothing you share on WhatsApp, including your messages, photos, and account information, will be shared onto Facebook or any of our other family of apps for others to see, and nothing you post on those apps will be shared on WhatsApp for others to see.*" For this segregation to be effective, however, the user must opt-out upon acceptance of the Whatsapp terms of service, otherwise Whatsapp will share the data with Facebook.

of service are going to be applied and eventually enforced.[26] Besides, the terms of service provide that, in case of discrepancies between the translation into local language and the original English version, the latter will prevail.[27]

Hence, the governing law the user is brought to accept is, in the vast majority of cases, the US, California law,[28] and the choice of law contributes to the diffusion of what has become a global - not national - legislation,[29]

---

[26] F. GALGANO, *Diritto ed economia alle soglie del nuovo millennio*, in *Contr. impr.*, 2000, p. 199: the Author considers that the above-referred contractual standards are "*merely translated, without any adaptation – not even from at a conceptual level – in view of the national laws of the various Countries; [any such amendments] may jeopardise the [standard contracts'] international uniformity;*" see also G. DE NOVA, *Il contratto alieno*, Torino, 2008, pp. 2-3; and S. PATTI, *La globalizzazione del diritto e il contratto*, in *Obbl. contr.*, 2009, p. 498.

[27] See, for example, Facebook's Terms of Service (version dated 30 January 2015) available at https://www.facebook.com/terms: "*This agreement was written in English (US). To the extent any translated version of this agreement conflicts with the English version, the English version controls.*"

[28] See the Facebook's Terms of Service, which not only qualify the laws of the State of California as the governing law, but also defer any disputes to the exclusive jurisdiction of the U.S. District Court for the Northern District of California or a state court located in San Mateo County ("*You [meaning «the user»] will resolve any claim, cause of action or dispute (claim) you have with us arising out of or relating to this Statement or Facebook exclusively in the U.S. District Court for the Northern District of California or a state court located in San Mateo County, and you agree to submit to the personal jurisdiction of such courts for the purpose of litigating all such claims. The laws of the State of California will govern this Statement, as well as any claim that might arise between you and us, without regard to conflict of law provisions*"). Other social networks defer any disputes to the Courts of New York (MySpace), or San Francisco (Twitter). In fewer cases – as long as EU users are concerned – the governing law is English law, and the Courts of England shall resolve any dispute arising from the terms: see, for example, YouTube's terms of use (https://www.youtube.com/-static?template=terms&gl=GB).

[29] C. PERLINGIERI, *Gli accordi tra i siti di social networks e gli utenti*, in AA.VV., *Internet e diritto civile cit.*, p. 206 qualifies the self-regulations set forth by multinational corporations as a "*global, non national, legal statute*" (*diritto globale non statale*), thereby recalling F. GALGANO, *id.*, pp. 199-202, who recognises that "*the global society has now its own legal statute, which is named the New* lex mercatoria *... A legal statute created by businesses, without the intervention of the Countries' law makers, and built up on around rules that are aimed at governing in a consistent manner ... the commercial relationships that are entered into within the unitary economic markets.*" The Author goes on to consider that this phenomenon does affect the Court's judgments on the lawfulness of the abovementioned regulations: considering that all these regulations fall outside the statutory legal schemes (*contratti tipici*), the Courts may be required to assess whether the regulations meet the legal test (to be qualified as fully valid and enforceable): in making such an assessment, the Court may not only keep in

which is the outcome of the creation of private regulations and terms of service. National lawmakers and authorities have little control over these globalised rules, unless Countries successfully put pressure on the service providers to have local laws applied, as it has happened in very few cases, *e.g.* in Germany Facebook was led to elect German law as the law governing the relationship with German users.[30]

Against this background, the possibility for (at least the European Union)[31] users of a social network to enjoy the protection afforded by local laws will very much depend on whether the users are deemed to fall within the definition of «*consumer*,» the rights of whom cannot be waived by virtue of the choice of law set forth in the social networks or other Web-based agreements.[32] This matter becomes of essence, on account that the service

---

mind the rules governing the validity and enforceability of contracts pursuant to the *lex fori*, but also the equivalent rules of the international society, so that - even if the contract does not entirely meet the *lex fori* test – the Court may still conclude in favour of the recognition of the private regulation, as the global community treats such regulation as fully valid and enforceable. P. PERLINGIERI, *Mercato, solidarietà e diritti umani*, in *Rass. dir. civ.*, 1995, pp. 95-96 expresses criticism in this regard, on account that the "*democratic connotation*" of the market rules is at stake, so that the Countries and the Courts are expected to keep control of the purely market-driven regulations.

30 See Facebook's special terms of use applying to German users (version dated 2 February 2016), available at https://www.facebook.com/terms/provisions/-german/index.php: "*Ziffer 15.1 wird ersetzt durch: Diese Erklärung unterliegt deutschem Recht,*" pursuant to which any disputes arising between the ISP and the German user shall be governed by the German law.

31 The issue regarding the enforceability of the choice of law and venue has arisen, however, also in the US, where the United States District Court for the Eastern District of Pennsylvania found that the arbitration provision set forth in the general terms of a multiplayer role-playing game set in the virtual world (named "*Second Life*") is unenforceable as an "*unconscionable agreement,*" meaning that the arbitration clause is the result of "*an oppression through the existence of unequal bargaining positions*" (see *Bragg v. Linden Research, Inc.* 487 F. Supp. 2d 593 (E.D.Pa. 2007).

32 See Art. 5 of the Rome Convention (80/934/EEC Convention on the law applicable to contractual obligations opened for signature in Rome on 19 June 1980), pursuant to which "*... a choice of law made by the parties shall not have the result of depriving the consumer of the protection afforded to him by the mandatory rules of the law of the country in which he has his habitual residence: if in that country the conclusion of the contract was preceded by a specific invitation addressed to him or by advertising, and he had taken in that country all the steps necessary on his part for the conclusion of the contract, or if the other party or his agent received the consumer's order in that country, or if the contract is for the sale of goods and the consumer travelled from that country to*

providers – as discussed *infra* – process a number of the user's data pertaining to the user's personality.

It may in principle be questionable whether the consumer regulation may be invoked in cases where the service is granted on a free-of-charge basis. The *rationale* underlying the consumer's protection rules is that the consumer shall be protected in cases where the agreement with a business party is not entered into at an arm's length, and the disparity between the parties' rights during the performance of a bilateral agreement (such as a sale and purchase agreement) may ultimately lead the business to obtain better contractual conditions, to the consumer's detriment. And there may be arguments to conclude that, whenever the relationship between the consumer and the business does not qualify as the provision of goods or services *for consideration*, the consumer regulation does not apply.

Whether the provisions of services by the Internet service provider is remunerated by consideration will be discussed in §4 below.

For the time being, the scope of application of directive 2011/83/EU on consumer rights (the "Consumer Directive") would appear to embrace any transaction that occurs into the Web, so long as the transaction is entered into between a consumer and a trader, *i.e.* a professional who is acting "*for purposes relating to his trade, business, craft or profession in relation to contracts covered by th[e] Directive*" (Art. 2(2)).

The circumstance that agreements between Internet service providers and users in principle fall within the scope of application of the Consumer Directive may be inferred from the circumstances that these agreements are not listed among the agreements to which the Consumer Directive does not apply (Art. 3(3)), and that the Consumer Directive expressly contemplates contracts governing the delivery of digital content (Recital 19; Art. 2(11)), irrespective of whether such contracts are for consideration of free-of-charge.

---

another country and there gave his order, provided that the consumer's journey was arranged by the seller for the purpose of inducing the consumer to buy."

Consistently, As far as the Italian consumers are concerned, Art. 3 of Italian legislative decree dated 6 September 2005 no. 206 (the "Italian Consumer Code") also supports the conclusion that the user is a consumer. The definition of «*product*» includes "*any product destined to a consumer, also in the context of a provision of services … [product] supplied or made available for consideration or free of charge in the context of a commercial activity.*"

The criterion for the application of consumer laws is therefore not the consideration, but the capacity in which the business is supplying the goods or services. From this angle, the Internet service providers undoubtedly qualify as traders for the purposes of the Consumer Directive, as they pursue an economically relevant target (again, see §4 below).[33]

The governing law clause set forth in the service provider' terms of service may therefore qualify as null and unenforceable over the European Union territory: Art. 6 of EC Regulation no. 593/2008 on the law applicable to contractual obligations (Rome I) sets out that a choice may not have the result of depriving the consumer of the protection afforded to him by provisions that cannot be derogated from by agreement by virtue of the law of the Country where the consumer has his habitual residence. Consistently, Art. 36, para. 5, of the Italian Consumer Code qualifies as null and void any clauses opting for the governing law of a non-EU Member State, which does not grant the consumer the minimum protection the consumer would have benefited pursuant to Italian law, or another EU Member State, as long as the agreement has a close connection with the territory of one Member State.[34]

---

[33] P. SAMMARCO, *op. cit.*, pp. 639-640.

[34] Also, R. CATERINA, *Cyberspazio, social network e teoria generale del* contratto, in *AIDA*, 2011, p. 97 concludes that a joint reading of Art. 3, para. 2, of law no. 218 of 31 May 1995 on the conflicts of laws and of Art. 15 *et seq.* of EC Regulation of 22 December 2000 on jurisdiction and the recognition and enforcement of judgments in civil and commercial matters shall import the possibility for an Italian user to litigate with the ISP before the Italian Courts, irrespective of the choice of venue set forth in the social network agreement.

In practical terms, these provisions mean that – whenever a connection exists between the agreement and the EU Member State where the user has his place of residence (*e.g.* the service is provided to users that are located in the Italian territory) – the minimum level of protection afforded by the laws of the place of residence cannot be waived by any unfair provisions (*i.e.* provisions that are unilaterally in favour of the professional and more restrictive for the consumer).[35] Notwithstanding the choice of law agreed between the parties, unfair terms (a) must be expressly understood and accepted in writing by the consumer in order to be valid and enforceable, and (b) are null and void if they have the effect to limit the possibility for the consumer to bring a lawsuit against the service provider in case of total or partial breach of contract.

Similar arguments may apply to the service provider's excluded liability, so long as such discharge encompasses, among others, the provider's liability arising in connection with any harm caused to the user as a consequence of the provider's conduct or omission or breach of contract.[36]

In cases where the user avails of the services in his capacity as a business, of course, no consumer regulation will apply, but the clauses that limit or exclude the service provider's liability shall be specifically agreed in writing

---

35 Unfair terms are usually referred to as *clausole vessatorie* under Italian law.

36 See, for example, Facebook's terms of service: "*Facebook is not responsible for the actions, content, information, or data of third parties, and you release us, our directors, officers, employees, and agents from any claims and damages, known and unknown, arising out of or in any way connected with any claim you have against any such third parties. if you are a California resident, you waive California civil code §1542, which says: a general release does not extend to claims which the creditor does not know or suspect to exist in his or her favor at the time of executing the release, which if known by him or her must have materially affected his or her settlement with the debtor. we will not be liable to you for any lost profits or other consequential, special, indirect, or incidental damages arising out of or in connection with this statement or Facebook, even if we have been advised of the possibility of such damages. our aggregate liability arising out of this statement or Facebook will not exceed the greater of one hundred dollars ($100) or the amount you have paid us in the past twelve months. applicable law may not allow the limitation or exclusion of liability or incidental or consequential damages, so the above limitation or exclusion may not apply to you. in such cases, Facebook's liability will be limited to the fullest extent permitted by applicable law*. "

in most EU jurisdictions (in Italy pursuant to Art. 1341 of the Italian Civil Code, and provided that no liability for fraud or gross negligence could be validly waived pursuant to Art. 1229 of the Italian Civil Code).

4. The relationship between the web user and the Internet service provider: the cross-license agreement

By accepting the terms of service, the user enters into an agreement that is expressly qualified to be «*free of charge*,» in that the Internet service provider undertakes neither to provide the services – not even for a limited period of time – nor to ensure that the software will perform in accordance with the expected use. Consistently, based on the terms, no provider's liability shall in any case arise in connection with the provision of services; on the other side, the user does not pay for any of the social networking services and, like the service provider, is entitled to terminate the agreement at any time.

The lack of a duty to perform the contractually agreed services may *prima facie* exclude any contractual liabilities on the parties to the agreement as well as the possibility for any further circulation of the user profile, meaning that, for example, the user's heirs may not take on the profile of a user who passed away, as further discussed in §5.2 below.

The circumstance that the use of the provider's service will never result in any dispute – on accounts that there is no duty to perform any act on either party – may be construed as an argument to deny that the relationship between the service provider and the user even qualifies as contract. The exchange of mutual, reciprocal promises between the parties being the substantive element to qualify undertakings to perform an act (or to refrain from performing an act) as a contract in most jurisdictions.[37]

And, assuming that the relationship between the user and the service provider is a contract, one may still doubt that such contract is a bilateral

[37] Art. 1321 of the Italian Civil Code provides that "*A contract is the agreement of two or more parties to establish, regulate or extinguish a patrimonial legal relationship among themselves.*"

contract, pursuant to which one party (the service provider) undertakes to provide goods or services for consideration.[38]

However, deepening into the analysis of the contractual mechanism set forth under the agreement between the user and the service provider, results may vary, in that the provision of social networking, video hosting or search engine services is, on one hand, *free of charge*, but, on the other, is by no means *cost-free*, for the reasons anticipated in §2 above.

At the outset of the analysis of the contractual relationship between the service provider and the user, the user's acceptance shall be investigated with specific reference to the user's opt-in for the purposes of the data processing. In fact, by entering into the agreement, the user consents that his personal data will be processed. Even more so, the consent to the data processing qualifies as a condition precedent for the use of the service provider's services and software. For example, the Facebook's Terms of Service provide as follows:

"*Our goal is to deliver advertising and other commercial or sponsored content that is valuable to our users and advertisers. In order to help us do that, you agree to the following:*

1. *You give us permission to use your name, profile picture, content, and information in connection with commercial, sponsored, or related content (such as a brand you like) served or enhanced by us. This*

[38] Italian commentators have questioned the apparent nature of the social networks agreement from an Italian law perspective, and have come to different conclusions, all of which highlight the peculiarity of the lack of an *express* consideration: «*free of charge contract*» (*contratto gratuito atipico*) or «*network contract*» (*contratto di rete*) (F. ASTONE, *Il rapporto tra gestore e singolo utente: questioni generali*, in *Ann. it. dir. ind.*, 2011, p. 107); «*courtesy relationship*» (*rapporto di cortesia*) (P. SAMMARCO, *op. cit.*, p. 634); «*access to Internet contract*» (*contratto di accesso ad Internet*) (M. GRANIERI, *op. cit.*, p. 128); «*bilateral undertakings and associative schemes*» (*accordi bilaterali e schemi associativi*) (S.A. CERRATO, *I rapporti contrattuali (anche associativi) tra i soggetti del social network*, in *AIDA*, 2011, pp. 168-218; W. VIRGA, *Inadempimento di contratto e sanzioni private nei social network*, in *Ann. it. dir. ind.*, 2011, p. 222, who argues that the social network agreement is somewhere in "*between an agreement and a community*," due to the fact that the vertical relationship between the user and the ISP is strictly related to the horizontal connections that the user establishes with the other users).

means, for example, that you permit a business or other entity to pay us to display your name and/or profile picture with your content or information, without any compensation to you. If you have selected a specific audience for your content or information, we will respect your choice when we use it.

2. *We do not give your content or information to advertisers without your consent.*
3. *You understand that we may not always identify paid services and communications as such.*"

From a purely economic standpoint, it is crystal clear that, while the use of the software and the social network services are granted on a free of charge basis, it is the profile-connected activity that generates the service provider's revenues. Consideration for the advertisement displayed on the user's pages is paid by the advertiser to the provider; in addition, data relating to the preferences shown by the social network community is assigned for consideration to subsequent purchasers.[39]

The meaning of the user's consent to data processing for advertising purposes does not only have privacy implications: consent qualifies as a condition precedent to both (a) enter into the social network agreement, and (b) enjoy the social networking services.[40] Put it otherwise, by entering into the social network agreement, the user undertakes to provide his personal data as consideration for the social networking services, so that the very personal features belonging to the user, in his capacity as an individual, are at stake: his rights to privacy, image, identity.[41]

---

[39] E. ROSATI e G. SARTOR, *Social networks e responsabilità dei provider*, in *European University Institute - EUI Working Papers*, available at www.cadmus.eui.eu.

[40] S. THOBANI, *Il consenso al trattamento dei dati come condizione per la fruizione dei servizi online*, in AA.VV., *Internet e Diritto civile*, *cit.*, p. 459 *et seq.*.

[41] For the purposes of the Italian Data Protection Code «*personal data*» shall mean "*any information relating to natural persons that are or can be identified, even indirectly, by reference to any other information including a personal identification number*" (Art. 4, para. 1, let. b).

Overall, the analysis of the interests at stake and the economic (in the sense of «*economically valuable*») transaction the user and the service provider put in place, as discussed in §§4.1 to 4.3 below, can be construed as a purely contractual relationship, whereby the user is granted access to the services, by consenting to the processing of his personal data and granting the social network provider the right to (directly or indirectly) profile the user for the purposes of targeted advertising: in spite of the apparent qualification of gratuitous contract, the social network agreement meets – both in fact and from a legal standpoint – the necessary requirements to qualify as a bilateral agreement. Based on the most widespread social networks' terms of use, the user grants the service provider a non-exclusive, transferable, sub-licensable, royalty-free, worldwide license to exploit the user's personal data and user-generated content in exchange for the personal, royalty-free, non assignable and, of course, non exclusive license to use the service provider's software.

The transaction entered into between the user and the service provider, however, poses a number of questions regarding the possibility for the user to market his very personal features (§4.1); the scope, limits and revocability of the consent to the processing of the user's personal data (§4.2); and the enforceability of the user's right to exploit the service provider's software pursuant to the abovementioned cross-license agreement (§4.3).

4.1 The marketability of the user's digital identity

The agreement between the user and the service provider relies on a scheme, pursuant to which information is a «*tradable commodity.*»[42]

---

[42] See A. DE FRANCESCHI, M. LEHMANN, *ibidem;* P. PERLINGIERI, *L'informazione come bene giuridico*, in *Rass. dir. civ.*, 1990, p. 326; F. AGNINO, *Fino a che punto è possibile disporre contrattualmente dei propri diritti? (Vedi contratto FB)*, in *Giur. mer.*, 2012, p. 2560. The question as to whether information can be treated as a *res* is still debated in the Italian legal culture: some commentators – while openly recognising that information can be the subject of contractual obligations – had denied that information can qualify as a 'good,' on grounds that information can only be indirectly protected (*e.g.* through the statutory protection afforded by trade secrets, confidential information and privacy), see D. MESSINETTI, *Oggettività giuridica delle cose incorporali*, Milano, 1970, p. 36. However, other commentators have concluded that an express legal provision is

It is questionable whether trading the individual's identity complies with the Charter of Fundamental Rights of the European Union, Art. 3, para. 2, let. c) of which sets out the "*prohibition on making the human body and its parts as such a source of financial gain.*" The provision has been interpreted to establish a market-inalienability principle, pursuant to which the human body is not *tout court* indisposable («*inalienability*»), but the rights on human body cannot be traded or waived for consideration («*consideration*» having the widest possible meaning in the circumstances of the case).

However, no such provision would appear to govern the disposability of the «*identity*» (which is of course something different from the body). The individual's personal features other than the body are protected under the different perspective of Art. 8, according to which "*…data must be processed fairly for specified purposes and on the basis of the consent of the person concerned*").

Hence, while any acts of disposition concerning the human body shall be governed by the market-inalienability principle, consent qualifies as the necessary requirement for any commercial exploitation of the data subject's personal data.[43] From this perspective, the consent to the processing of the user's personal data for the purposes of using the services made available into the Web does not appear to generate the risk that any prospect of a "*financial gain*" may lead the user to irremediably waive his personal rights. A similar situation appears to occur, *mutatis mutandis*, whenever an actor or sportsman undertakes to be a brand's testimonial, thereby allowing the contracting party

---

not necessary to support the categorisation of information as a 'good' (see V. ZENO ZENCOVICH, *'Cosa', Digesto delle discipline privatistiche. Sezione civile*, Torino, 1989, p. 453; V. ZENO ZENCOVICH, *Sull'informazione come «bene» (e sul metodo del dibattito giuridico)*, in *Rass. dir. civ.*, 1999, p. 485; A. DE FRANCESCHI, M. LEHMANN, *id.*, p. 54), on the basis that the broad wording used Art. 810 of the Italian Civil Code – pursuant to which a 'good' is something that is capable of being the 'subject of rights' – allows information to be included within the scope of the definition of 'goods.'

[43] G. RESTA, *La disponibilità dei diritti fondamentali e i limiti della dignità (note a margine della Carta dei Diritti)*, in *Riv. dir. civ.*, 2002, p. 806 *et seq.*.

to use his image for commercial purposes.[44] Besides, the trend launched by ISPs and other *media* consisting in the collection of personal data and images other than those of famous people, *i.e.* the so-called *common people*, testifies the increasing market request for the latter's personal data, which is exploited (among others) as a marketing tool to promote the consumerism.[45]

This conclusion is without prejudice to the individual/user's right to dignity – which right the user may at any time invoke *ex post* to cease any processing that may put dignity at risk, it being understood that the right to dignity cannot be used to preventively limiting the same user's personal freedoms.[46]

4.2 The user's consent to the processing of his personal data by the Internet service provider

The marketability of personal features and personal data relies on a revised, amended concept of privacy, resulting in a shift from privacy as an individual's inalienable and fundamental right[47] "*to be let alone*," to an

---

[44] P. CRUGNOLA, *Problemi giuridici relativi all'uso di fotografie per la pubblicità commerciale*, in Dir. aut., 1973, p. 423, M. RICOLFI, *Questioni in tema di regime giuridico dello sfruttamento commerciale dell'immagine*, in *Nuova. giur. civ. comm.*, 1992, p. 51, C. SCOGNAMIGLIO, *Il diritto all'utilizzazione del nome e dell'immagine delle persone celebri*, in Dir. informatica, 1988, p. 1 *et seq.*.

[45] V. ZENO ZENCOVICH, *Profili negoziali degli attributi della personalità*, in *Dir informatica*, 1993, p. 545.

[46] P. FEMIA, *Interessi e conflitti culturali nell'autonomia privata e nella responsabilità civile*, Napoli, 1996, p. 559.

[47] See Article 8 of the European Union Charter of Fundamental Rights ("*1. Everyone has the right to the protection of personal data concerning him or her. 2. Such data must be processed fairly for specified purposes and on the basis of the consent of the person concerned or some other legitimate basis laid down by law. Everyone has the right of access to data which has been collected concerning him or her, and the right to have it rectified. 3. Compliance with these rules shall be subject to control by an independent authority.*") and Art. 16 of the Treaty on the Functioning of the European Union ("*1. Everyone has the right to the protection of personal data concerning them. 2. The European Parliament and the Council, acting in accordance with the ordinary legislative procedure, shall lay down the rules relating to the protection of individuals with regard to the processing of personal data by Union institutions, bodies, offices and agencies, and by the Member States when carrying out activities which fall within the scope of Union law, and the rules relating to the free movement of such data. Compliance with these rules shall be subject to the control of independent authorities.*

individual's "*right to keep control over his personal data,*"[48] *i.e.* to govern the circulation and diffusion of his own data, just like it happens for any other (tangible or intangible) good that is susceptible to be subject to a property right.[49]

This evolution of privacy is somewhat consistent with a concept of information as tradable commodities, so that also the individual's right to privacy on certain data can now constitute (either implicitly or explicitly) the actual subject matter of a transaction,[50] *i.e.* data can be traded for a

---

The rules adopted on the basis of this Article shall be without prejudice to the specific rules laid down in Article 39 of the Treaty on European Union.").

[48] S. RODOTÀ, *Tecnologie e diritti*, Bologna, 1995, p. 50 *et seq.*. On the evolution of the concept of privacy see *I servizi di online social network tra privacy, regole di utilizzo e violazione dei diritti dei terzi*, in *Giur. mer.*, 2012, p. 2572.

[49] On the relationship between right to privacy and property right see C. PRINS, *Property and Privacy: European prospectives and the commodification of our identity*, in *Information law series*, available at https://papers.ssrn.com/sol3/papers.-cfm?abstract_id=929668. The Author observes that "*Many of the arguments that have been forwarded in favor of ... a proprietary perspective derive from American sources. There has been relatively little discussion in Europe of whether such an approach could resolve the pressing problems of personal data protection - a fact that is not entirely surprising, given the European human rights-oriented approach to privacy protection. ... although it is all too often argued that the creation of a property right is not in line with the human rights-based approach to privacy, the European system appears to offer considerable leeway and even openings for a property rights model. But ... the property argument fails to recognize the data protection challenges that arise with present-day developments in the area of context-aware computing. In a society in which our behavior and identities (i.e. not individual data as such), become the object of commodification, the debate on data protection mechanisms must be structured along lines of control and visibility, rather than ownership. This then will require a debate on the role of the public domain in providing the necessary instruments that will allow us to know and to control how our behavior, interests and social and cultural identities are «created»*" (added emphasis). On the debate in the U.S. regarding the possibility to trade personal data that are protected under privacy laws just as the other goods/properties see (i) in favour of the tradability M.J. RADIN, *Incomplete Commodification in the Computerized World*, in N. ELKIN-KOREN, N. WEINSTOCK NETANEL (eds.), in *The Commodification of Information*, New York, NY, 2002, pp. 3-21; (ii) against the tradability see P. SAMUELSON, *Privacy as Intellectual Property?*, 52 *Stanford Law Review*, 2000, pp. 1140-1142.

[50] See S.F. BONETTI, *La tutela dei consumatori nei contratti di accesso ad internet: i contratti dei consumatori e la privacy tra fattispecie giuridiche e modelli contrattuali italiani e statunitensi*, in *Dir. inf.*, 2002, p. 1093 *et seq.*.

consideration (so long as the entire transaction is susceptible of being economically valuable, in the sense presented in §4.3 below).

But, to what extent is personal data marketable within the scope of an agreement?

Directive 95/46/EC on data protection expressly contemplates the free flow of personal data – together with the free movement of goods, persons, services and capital – as a means to facilitate the exchange of goods and develop the internal market, provided that movement/flows do not prejudice the fundamental rights of the data subject.[51]

The above-referred GDP Regulation on privacy then acknowledges that the scale of collection and sharing of personal data has increased significantly since the adoption of directive 95/46/EC, so that it has become necessary to ensure a higher level of protection to personal data,[52] thereby vesting the

---

[51] See recitals 3 and 4 of 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: "*(3) Whereas the establishment and functioning of an internal market in which, in accordance with Article 7a of the Treaty, the free movement of goods, persons, services and capital is ensured require not only that personal data should be able to flow freely from one Member State to another, but also that the fundamental rights of individuals should be safeguarded; (4) Whereas increasingly frequent recourse is being had in the Community to the processing of personal data in the various spheres of economic and social activity; whereas the progress made in information technology is making the processing and exchange of such data considerably easier*" (added emphasis). The free flow of personal data was previously dealt with in the Convention for the Protection of Individuals with regard to Automatic Processing of Personal Data (European Treaty Series – No. 108) executed in Strasbourg on 28 January 1981, which set the general principle that no party to the Treaty shall, for the sole purpose of the protection of privacy, prohibit or subject to special authorisation transborder flows of personal data going to the territory of another party. Exceptions to the principle obviously apply, but it is worth noting that the general principle is in the sense of guaranteeing the freedom of data flows. At the 30th International Conference of Data Protection and Privacy Commissioners, Strasbourg, 17 October 2008 the European data protection authorities discussed a draft resolution on the urgent need for protecting privacy in a borderless world, and for reaching a "Joint Proposal for setting International Standards on Privacy and Personal Data Protection" (see http://www.garanteprivacy.it/web/guest/home/docweb/-/docweb-display/docweb/1760379).

[52] See recital 6 of 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): "*Rapid technological*

national supervisory authorities with the task to monitoring the application of this Regulation, in an attempt to find a balance between the protection of the fundamental rights and freedoms of natural persons in relation to processing and the free flow of personal data within the European Union.[53]

Against this background, the definition of personal data is wide enough to include not only the user's name, surname and username, but also the place of residence, domicile (if any), the e-mail address, the telephone number (including mobile), all of which data become possible addresses for the data controller to reach and exploit for commercial (advertising, or other) purposes. Such data are processed also by the Internet service provider in the performance of the agreement with the web user.

The Italian Authority (*Garante per il trattamento dei dati personali*) has issued a number of resolutions setting forth a framework for the commercial use of the abovementioned personal data by the data controllers.[54] The user's

---

developments and globalisation have brought new challenges for the protection of personal data. The scale of the collection and sharing of personal data has increased significantly. Technology allows both private companies and public authorities to make use of personal data on an unprecedented scale in order to pursue their activities. Natural persons increasingly make personal information available publicly and globally. Technology has transformed both the economy and social life, and should further facilitate the free flow of personal data within the Union and the transfer to third countries and international organisations, while ensuring a high level of the protection of personal data."

[53] See Art. 51 of EU Regulation 2016/679.

[54] The Garante sanctioned the broadcasters which had sent advertising text messages to the user's mobile phones despite the fact that the users had previously denied their consent for such a processing (see Garante's resolution of 12 March 2003, docweb no. 29844, available at http://www.garanteprivacy.it/web/guest/home/docweb/-/docweb-display/-docweb/29844). Also the Italian Courts sanctioned data controllers on similar grounds, see *Trib. Latina*, 19 June 2006, in *Foro it.*, 2007, c. 324. The Garante also prohibited the unsolicited telephone calls from commercial operators: the 2008 newsletter highlights that the Privacy Dept. of the *Guardia di Finanza* had discovered a number of cases where the data controllers had transferred the users' personal data to third acquirers, which had approached the data subjects with direct marketing activities, without checking whether the data subject had granted their consent (see Garante's newsletter no. 316 of 2 December 208, available at http://www.garanteprivacy.it/-web/guest/home/docweb/-/docwebisplay/docweb/1571-099). The Court of Rome also stated that direct marketing towards clients that have opted out may qualify as unfair commercial practice in breach of the rules set forth in the Italian Consumer Code (see *Trib. Roma*, 26 July 2007, in *Dir. informatica*, 2007, p. 859). The Garante has also

consent is the necessary requirement for any use of the personal data for commercial purposes.[55]

---

issued resolutions prohibiting the unauthorised processing of the user's email consisting in spamming (*i.e.* unsolicited emails) (see Garante's newsletter dated 17 January 2000 available at http://www.garanteprivacy.it/web/guest/home/docweb/-/docweb-display/docweb/47199; Garante's resolution dated 13 May 2008, docweb no. 1521775 available at http://www.garanteprivacy.it/web/guest/home/docweb/-/docweb-display/docweb/1521-775; and, more recently, the Garante's guidelines dated 12 September 2013, docweb no. 2629816, available at http://www.garanteprivacy.it/web/guest/home/docweb/-/docweb-display/docweb/262-9816). A number of legal scholars have also analysed the issue of unsolicited communications, see F. BRAVO, *Invio di sms commerciali e risarcimento del danno da illecito trattamento di dati personali*, in *Dir. informatica*, 2007, p. 798-814; E. TOSI, *Prime osservazioni sull'applicazione della disciplina generale della tutela dei dati personali a internet e al commercio elettronico*, in *Dir. informatica*, 1999, p. 591 *et seq.*; S. MELCHIONNA, *La tutela dei dati personali nell'ambito delle comunicazioni elettroniche*, in V. CUFFARO, R. D'ORAZIO, V. RICCIUTO, *Il codice del trattamento dei dati personali*, Torino, 2007, p. 577 *et seq.*; A. LEVI, F. ZANICHELLI, *L'utilizzo dell'e-mail a fini pubblicitari: dallo «spamming» al «permission marketing»*, in *Riv. dir. ind.*, 2001, p. 194 *et seq.*.

[55] Pursuant to Artt. 13 and 23 of the Italian Data Protection Code, the consent shall be deemed to have been validly granted only if such consent was provided on a *free* and *informed* basis (*consenso libero e informato*). The Garante has issued a number of resolutions to better clarify the meaning of the statutory requirement: the information notice the data controller provides to the data subject shall (i) contain detailed information as to the scope of the processing, (ii) precisely identify the categories of third parties to which the data subject's personal data may be communicated, (iii) clarify when the provision of data is mandatory (*e.g.* because it is necessary to perform the obligations arising from the agreement, or to comply with a statutory requirement, in which cases no consent shall be sought by the data controller) or optional (*e.g.* to perform other processing activities, which are subject to the data subject's consent). The Garante – based on the assumption that agreements between professionals and consumers are not entered into at an arm's length, and that the consumer is the weak contractual party - further clarified that data controllers, in their capacity as professionals/businesses, shall not seek a general and unconditional consent, because the consent shall be deemed to be granted on a free basis only if the data subject, in its capacity as a consumer, is not led to believe that a denial of consent to the additional processing of data may prevent the provision of the main service (see Garante's resolution dated 28 May 1997, docweb no. 40425, available at http://www.garanteprivacy.it/web/guest/home/docweb/-/docweb-display/docweb/40425; Garante's opinion dated 22 October 1997, docweb no. 1055346 http://www.garanteprivacy.it/web/guest/home/docweb/-/docweb-display/docweb/1055-346; Garante's opinion dated 8 September 1997, docweb no. 1055101, available at http://www.garanteprivacy.it/web/guest/home/docweb/-/docweb-display/docweb/1055-101; Garante's newsletter dated 21 June 1999, docweb no. 18589, available at http://www.garanteprivacy.it/web/guest/home/docweb/-/docweb-display/docweb/48589). See also R. FRAU, *Profili del consenso al trattamento dei dati personali per fini economici nell'esperienza italiana. Raffronti con la normativa spagnola*, in *Resp. civ. e prev.*, 2010, p. 2598 ss; A. Longo, *Privacy e assicurazioni*, in V. CUFFARO, R. D'ORAZIO, V. RICCIUTO, *Il codice del trattamento dei dati personali cit.*, p. 555 *et seq.*.

The user's consent to the processing of his personal data is not, however, a mere requirement to which the processing of data provided in the execution of an agreement is subject, because the processing of the user's data is *per se* at the core of the user's obligations for the purposes of the execution and performance of the agreement.[56]

The legal value of the consent to the processing of personal data is debatable: [57]

(a) some commentators sustain that it is a non-contractual authorisation act, that is substantially reconcilable with the criminal law principle of the right-owner's consent, and, on this basis, conclude that the user is entitled to withdraw and revoke his consent at any time, provided that, in any event the right-owner's (*i.e.* the data subject) rights shall be balanced with the rights granted to the lawful acquirer of the personal data (*i.e.* the data controller);[58]

(b) others ascribe a pure contractual nature to the consent, at least in cases where the user grants his consent to the processing of personal

---

In addition to the above, any profiling shall be subject to the condition precedent that the data controller has previously notified the Garante pursuant to Art. 37, para. 1, let. d) of the Italian Data Protection Code,

[56] G.M. RICCIO, *Social network e responsabilità civile*, in *Dir. informatica*, 2010, pp. 859-871; F. AGNINO, *Fino a che punto è possibile disporre contrattualmente dei propri diritti? cit.*, p. 2564.

[57] V. ZENO ZENCOVICH, *Profili negoziali degli attributi della personalità cit.*, p. 554 *et seq.* underscores that the act by which the right-owner entitles third parties to exploit a feature of his personality (*attributo della personalità*) may fall within different legal schemes and may ultimately constitute the element to an agreement; according to G. SANTINI, *I diritti della personalità nel diritto industriale*, Padova, 1959, p. 164 this act of disposition qualifies as a pure authorisation; A. DE VITA, *sub art. 10*, in A. PIZZORUSSO, R. ROMBOLI, U. BRECCIA, A. DE VITA, *Delle persone fisiche*, in *Comm. Scialoja-Branca*, Bologna, 1988, p. 505 *et seq.* and M. RICOLFI, *Il contratto di merchandising nel diritto dei segni distintivi*, Milano, 1991, p. 23 *et seq.* underscore that, to the extent the unauthorised use of a right-owner's personal feature qualifies as an unlawful conduct (with possible criminal implications), the right-owner's act of disposition qualifies as a pure consent (*consenso dell'avente diritto*).

[58] G. COMANDÈ, *Consenso. Casi di esclusione del consenso*, in E. GIANNANTONIO, M.G. LOSANO, V. ZENO ZENCOVICH, *La tutela dei dati personali*, Padova, 1999, pp. 98 *et seq.*, 115-116 qualify the grant of the data subject's consent as an "*atto non negoziale di natura autorizzatoria*."

data in his capacity as a party to an agreement.[59] Insomuch as the consent is granted in the context of a wider agreement, it appears at least arguable whether the consent can be later withdrawn; and

(c) according to an intermediate opinion, the consent may be withdrawn at any time, but cannot be retroactively enforced. This solution appears to be consistent with the rule under Art. 7, para. 4, let. a) of the Italian Data Protection Code, pursuant to which – in a scenario where the processing complies with the scope declared by the data controller – the data subject shall have the right to object to the processing of his personal data exclusively on legitimate grounds. Hence, the data subject cannot unilaterally revoke his consent, unless he has valid reasons to do so.[60]

Applying the principles above to the case of the agreement between the user and the service provider, it is worth noting that, while the Italian Data Protection Code does not set forth any explicit provision governing the possibility for the data subject to revoke his consent, Art. 7 of the same Code actually entitles the data subject to object to the processing not only in cases where the data subject has legitimate grounds to do so (Art. 7, para. 4, let. a), but also in cases where the processing is carried out for the purpose of sending advertising materials or direct selling or for the performance of market or commercial communication surveys (let. b).[61]

---

[59] F. BILOTTA, *Consenso e condizioni generali di contratto*, in V. CUFFARO, V. RICCIUTO, *La disciplina del trattamento dei dati personali*, Torino, 1997, pp. 89-91; V. CUFFARO, *A proposito del ruolo del consenso*, in V. CUFFARO, V. RICCIUTO, V. ZENO ZENCOVICH, *Trattamento dei dati e tutela della persona*, Milano, 1999, p. 201 *et seq.*.

[60] G. BUTTARELLI, *Banche dati e tutela della personalità*, Milano, 1997, p. 286.

[61] The data subject's right pursuant to Art. 7, para. 4, let. b) of the Italian Data Processing Code shall be legitimately waived by the same data subject, so that the latter might not oppose to direct marketing activities in cases where it has previously granted his consent.

It may be inferred that, by so providing, the Italian Data Protection Code is expressly taking into account the scenario where the consent was granted pursuant to a wider agreement.[62]

The general principle that the undertakings between parties shall be durable[63] would appear to be an argument in favour of the non-revocability of the consent – without prejudice to the user's opposition rights, in his capacity as a data subject, pursuant to Art. 7 of the Italian Civil Code – that was granted in the context of an agreement. Once the parties have agreed to be bound to an agreement, pursuant to which the data subject is entitled to perceive an economically valuable consideration, any data subject's right to revoke his consent shall be limited in accordance with the principle of good faith governing the interpretation and performance of contracts.[64]

On this basis, the possibility for the user to revoke his consent shall be subject to the rules governing the unilateral termination, so that the user shall be bound to the obligations arising from the contract, unless he has legitimate grounds to unilaterally terminate it.[65] This conclusion obviously relies on the assumption that, pursuant to the agreement, the user has perceived, or is due to perceive, compensation in exchange for the consent to a given processing of personal data. Should the user then oppose to the processing of his data for the purposes identified in the agreement, the user shall own no consideration

---

62 F. AGNINO, *Fino a che punto è possibile disporre contrattualmente dei propri diritti? cit.*, p. 2567.

63 «*Principio della stabilità dei rapporti*.»

64 *Ibid.*; on the applicability of the principle of good faith to the unilateral deeds with economically valuable content (*atti unilaterali a contenuto patrimoniale*) see C.M. BIANCA, *Diritto Civile, vol. III, Il Contratto*, Milano, 1984, pp. 378-379. The Italian Competition Authority (*Autorità Garante della Concorrenza e del Mercato*), in deciding a case of possible misleading advertising, incidentally established that a relationship subsisted between the gratuity of the service provided by a provider of electronic communications and the user's consent to the provider's use of his data for the purposes of profiling and marketing (see AGCM's decision no. 10276 of 20 December 2001, in *Giust. civ.*, 2002, p. 1748).

65 F. AGNINO, *Fino a che punto è possibile disporre contrattualmente dei propri diritti? cit.*, p. 2568.

to the service provider, as consideration would fail to find any contractual justification.

4.3 The grant of a license on the user-generated content[66]

The access and the use of the ISP's platform are nothing but the access and the use of a software based on a license agreement, upon the condition that another license is granted by the user on the data and the intellectual property rights that are generated by the user by means of the software, and that are subsequently stored onto the social network's servers.

It is of the outmost difficulty to fit the service agreement entered into between the user and ISP within the statutory categories of contracts (the so-called «*contratti tipici*»):[67] the licenses underlying the user-ISP agreement are in fact so peculiar in the market practice, that they shall ultimately fall within the catch-all category of contracts set forth under Art. 1322 of the Italian Civil Code (*i.e.* the "*contracts that are not of the types that are particularly regulated*," so-called «*contratti innominati*»). Most of the difficulties in determining the legal scheme that is suitably applicable to the user-provider agreement arise in connection with the fact that it does not provide for an express consideration for the supply of services to the user, a circumstance which, as discussed above, may lead to the conclusion that user-ISP

---

66 This paper aims at investigating how ISPs collect and exploit the user's data and the user-generated content, therefore any assessment of (i) the other users' rights to access and use the such data and (ii) the consequences of the user uploading third party copyrighted works or other unauthorised content shall fall out of the scope of this paper. For the issue mentioned in (i) see A. COGO, *Le regole del contratto tra social network e utente sull'uso della proprietà intellettuale del gestore, dell'utente e degli altri utenti – riflessioni a partire dall'individuazione del fenomeno, dei suoi soggetti e della funzione del contratto*, in *AIDA*, XX, 2011, pp. 320-329; the subject matter mentioned in (ii) above is extensively analysed by P. DI MICO, *Il rapporto tra diritto di autore e social network: un nuovo capitolo, ma non l'ultimo*, in *Dir. aut.*, 2010, pp. 262-276; L.MANSANI, *Contenuti generati dagli utenti*, in *AIDA*, 2010, pp. 244-257.

67 P. SAMMARCO, *op. cit.*, p. 636. The Author also considers that, if the user-ISP were to be subject to any contractual category (*contratti nominati*), there may be arguments to sustain that the user-ISP agreement falls within the scope of the lease of a movable asset (*locazione di bene mobile*).

agreements do not fit the «*do ut des*» scheme that characterises bilateral contracts, among which the license agreement.[68]

By emphasising the lack of an express consideration, the agreement would fall within the scope of the atypical gratuitous deeds («*negozi gratuiti atipici*»), because the agreement provides for the allocation of rights from one party to another, without an (express) impoverishment of the assigning party, as it would instead occur, for instance, in case of a donation. In any event, the legal scheme set forth under Art. 1322 of the Italian Civil Code does not make any distinction between gratuitous contracts or contracts for consideration; the assessment of the Art. 1322 contracts' effectiveness still relies on the contractual aim («*causa*») the parties pursue by entering into the agreement: if the aim is unlawful, the contract, as widely known, is null and void pursuant to Art. 1325 of the Italian Civil Code.[69]

It must therefore be verified whether an atypical gratuitous deed like the agreement between the user and the service provider actually pursues "*interests worthy of protection according to the legal order*" for the purposes of Art. 1322 of the Italian Civil Code,[70] which qualifies as a necessary

---

[68] P. SAMMARCO, *ibidem*. The Author, at the outset of his analysis, excludes that the user-ISP relationship shall be deemed to rely on a unilateral deed (*atto unilaterale*), pursuant to which the user would waive his rights, to the effect that these rights are consequently assigned to the ISP (*rinuncia traslativa*), because, if the waiver is aimed at granting another party an economically valuable advantage, such waiver then falls outside the scope of the unilateral deeds, and rather complies with the legal scheme of bilateral undertakings (*negozi bilaterali*i). See also F. MACIONE, *Rinuncia (dir. priv.)*, in *Enc. dir.*, vol. XL, Milano, 1989, p. 933: he who waives a right pursue the exclusive aim to dispose of a right he has no interest to retain, and the advantage a third acquirer may acquire is merely secondary and unintended by the waiving party.

[69] *Corte Suprema di Cassazione*, decision no. 10612 of 9 October 1991, in Giust. Civ., 1991, p. 2895 confirmed that "*the possibility to configure non-statutory assignment contracts is granted by the principle of contractual freedom (*autonomia contrattuale*) pursuant to Art. 1322 of the Italian Civil Code, upon the condition that the aforementioned contracts pursue a lawful aim (*causa lecita*)*."

[70] The prevailing opinion among Italian Scholars is that the requirement set forth under Art. 1322 of the Italian Civil Code corresponds to the limits imposed by the public policy (*ordine pubblico*), the common sense of moral (*buon costume*) and lawfulness (*liceità*). (G.B. FERRI, *Causa e tipo nella teoria del negozio giuridico*, Milano, 1966, p. 403; ID., *Meritevolezza dell'interesse e utilità sociale*, in *RDCo*, 1971, I, p. 89; ID., *Ancora in*

requirement for any contract to be valid. The subsistence of an economic advantage for the contracting parties ensures that there is a balance between the parties' interests.[71]

In other words, the user-ISP agreement satisfies the legal test if it can be established that any allocation of rights on a free-of-charge basis is justifiable on the basis of an economically valuable aim pursued via the agreement.[72]

---

tema di meritevolezza dell'interesse, in *RDCo*, 1979, I, p. 12; G. GORLA, *Il contratto*, I, Milano, 1954, p. 199; A. GUARNERI, *Meritevolezza dell'interesse*, in *Digesto civ.*, XI, Torino, 1994, p. 33; F. MESSINEO, *Dottrina generale del contratto*, Milano, 1952, p. 13). The majority of the Italian case law precedents are consistent with this opinion (*Cass.*, 6 February 2004, no. 2288, in *Contratti*, 2004, p. 801 and *Cass.*, 13 May 1980, no. 3142, in *Mass. Giur. It.*, 1980). Conversely, certain other case precedents express criticism as to the opportunity that Art. 1322 shall be interpreted to recall the notions of public policy, common sense of moral and lawfulness, because – by so opining – Art. 1343 would be a mere duplication of the principles set forth in Art. 1322, so that it must be inferred that the "*interest that deserves legal protection*" pursuant to Art. 1322 shall identify a somewhat wider concept, possibly including, among others, the market practice, the protection of the weaker contractual party, and the social utility of contracts (*Cass.*, 7 June 1991, no. 6496, in *Fisco*, 1991, p. 5007).

[71] P. SAMMARCO, *id.*, p. 638: as regards the contractual aim of the user-ISP agreement, the transaction entered into in the absence of a consideration must always bear an economically valuable interest that "*is referable to the assigning party who bears the sacrifice; and the interest is a part to the contractual undertakings, it defines the contractual aim and ultimately characterises the legal scheme pursuant to which the assignment is consummated. It is therefore necessary that the assigning party's allocation of rights on a free of charge basis (*liberalità*) shall be justifiable on grounds of an economically valuable interest, that is the transaction shall procure the assigning party an advantage that is suitable to be economically valued*." Lacking the assigning party's economically valuable interest, a mere allocation of rights to a third party cannot qualify as the relevant contractual aim pursued by the parties, because the requirements set forth under Art. 1322 cannot be properly assessed. It follows that an agreement providing for the assignment of a good, or the provision of a a service, is null and void due to a lack of contractual aim (*mancanza di causa*) if the parties failed to identify in the agreement the legal ground which justifies such assignment or provision. The rationale for such a serious sanction (the nullity of the contract) is that also the atypical contracts which do not pursue a donation (*contratti innominati a carattere non donativo*) cannot fail to meet the contractual aim test, *i.e.* said agreements must always pursue a determined socially-economically valuable target.

[72] F. CARINGELLA, *Alla ricerca della causa nei contratti gratuiti atipici*, in *Foro it.*, 1993, I, 1508. The Author qualifies those agreements that provide for an allocation or rights on a free-of-charge basis in the pursuit of an economically valuable aim as "*undertakings halfway between a donation and an exchange contract*"; while G. GORLA, *Il Contratto cit.*, p. 188, qualifies them as "*promises with an underlying interest (*promesse interessate*) the scheme of which falls outside the scope of both donations and exchange contracts.*"

And there can be no doubt that the user-ISP agreement ultimately qualifies as a bilateral agreement that pursues an economically valuable interest, in that the user's disposal of his privacy and intellectual property rights is functional to obtaining a right to use the web platform. Hence – contrary to the service provider's standard representations in the terms of service that there is no duty on either party to perform any activity or service – it must be inferred that the user has a right to use the software, and the social network provider is actually contractually bound to grant the user the possibility to use the software at any time, as long as ISP is entitled to collect, process and exploit the user's data.[73]

The above conclusion is confirmed by the following two circumstances.

First, in addition to granting the license on his intellectual property rights and personal data, the user also waives – and concurrently grants the ISP – the right to perceive economic consideration from third party advertisers by communicating and sharing the user's personal data and profile.[74] The execution of the agreement between the service provider and the advertiser is strictly connected with the agreement between the user and the service provider,[75] in that the provider will not be in a position to comply with its

---

[73] C. PERLINGIERI, *Gli accordi tra i siti di social networks e gli utenti cit.*, p. 214.

[74] See again the Facebook's Terms of Service: "*You give us permission to use your name, profile picture, content, and information in connection with commercial, sponsored, or related content (such as a brand you like) served or enhanced by us. This means, for example, that you permit a business or other entity to pay us to display your name and/or profile picture with your content or information, without any compensation to you. If you have selected a specific audience for your content or information, we will respect your choice when we use it.*"

[75] It is arguable whether this contractual set up is consistent with the legal scheme of the so-called connected contracts (*collegamento negoziale*), *i.e.* when negotiating a contract the parties may decide to split the execution process into different phases (instead of proceeding with the immediate and concurrent execution of the final agreement). Italian case law, for example, considers contracts to be connected if the following criteria are concurrently met: (a) the two (or more) agreements must pursue the same, unitary, economic transaction, *i.e.* the transaction is not suitable to be contemplated by means of one, single, contract; and (b) the parties' will (either expressly or tacitly) is to pursue the above-referred unitary transaction, by entering into multiple contracts. Such multi-phase process leads to a "*progressive development of the agreement*" (*formazione progressiva del contratto*), pursuant to which the execution of the final agreement(s) may occur after

contractual obligation to share the user's profile and data with the advertiser, unless the software works properly (thereby enabling the user to upload data on his profile and the advertiser to upload tailored advertisements on the user's page).[76]

Second, the provider reserves the right to unilaterally terminate the agreement, by deactivating the user's account, in cases where the user does not actively use the software for a given period, usually ranging between three and nine months.[77] The *rationale* underlying the provider's unilateral termination right is that the provider is not interested in hosting a web page, where data is not uploaded in real time, and, therefore, targeted advertising cannot be placed. Marketwise, data is not *per se* valued. Data is economically relevant only if and to the extent that it qualifies as a tradable commodity that is susceptible to circulate on the «*market of data*,» which values updated data the most.[78]

---

the signing of other agreements aimed at regulating specific aspects of the negotiation. In such a circumstance the contractual will of the parties should be construed through the interpretation of *all* the agreements that have been executed in the context of the same transaction. Hence, the answer to the question of whether the two agreements (*i.e.* user-ISP and ISP-advertiser) will very much depend on whether the assignment by the ISP to the advertiser of the user's data is considered to be a core obligation under the entire transaction contemplated under the social network agreement; to the extent that the ISP's terms of use so provide, and the user has granted his content, the conclusion shall in my view be that the two agreements are connected and pursue the same, unitary, transaction.

[76] C. PERLINGIERI, *Gli accordi tra i siti di social networks e gli utenti cit.*, p. 215.

[77] See, for example, the terms of service of Google Plus (nine months or more) or Twitter (six months).

[78] See the European Consumer Commissioner's (M. Kuneva) 9 March 2009 09/156 speech at the Roundtable on Online Data Collection, Targeting and Profiling, a summary of which is available at http://europa.eu/rapid/press-release_SPEECH-09-156_en.htm: "*Internet is an advertisement supported service and the development of marketing based on profiling and personal data is what makes it go round. Personal data is the new oil of the internet and the new currency of the digital world. We accept this reality because it is one chosen by users. Internet users have massively opted for free services offered in exchange for acceptance of advertisement. Today, advertisement online is individually targeted and increasingly based on the user's profile and behaviour. Tools must now be developed that balance the interests of business with that of the consumers. This means two things: the respect of users' right to control their public exposure; and the obligation to protect them against abusive and risky practices targeted at them.*"

Put it otherwise, the service provider does not pursue acquiring ownership over the user-generated content (most terms of service reassure the user "*You [the user] retain ownership of any intellectual property rights that you hold in that content. In short, what belongs to you stays yours*"[79]). What is material to the service provider is the ***access*** to such information, and the subsequent license to use it for the purposes of data mining, selling data to data brokers, direct advertising, *etc.*.

5. The termination of the social network agreement

Having established the contractual nature of the user-ISP relationship, it is from this angle that the termination of the social networking agreement shall be analysed.

Although the ISPs have implemented a variety of different solutions, the service provider usually has a right to suspend or interrupt the provision of services to the user, in cases where a cause for termination («*giusta causa*») occurs (meaning a contractual breach or also cases where the user has breached the policies, or the user's conduct is, or may be, detrimental to the service providers or the other users).[80]

In most cases, however, the service provider is also vested with a unilateral termination right, which can be exercised in the absence of any cause ascribable to the user («*recesso ad nutum*»).[81]

---

[79] See Google's Terms of Service (*Your Content in our Services*), available at https://www.google.com/intl/ALL/policies/terms/.

[80] See, for example, the Facebook Terms of Service (https://www.facebook.com/terms, version dated 30 January 2015): "*If you violate the letter or spirit of this Statement, or otherwise create risk or possible legal exposure for us, we can stop providing all or part of Facebook to you.*"

[81] See, for example, the Twitter Terms of Service (https://twitter.com/tos?lang=en, version dated 30 September 2016): "*We may suspend or terminate your account or cease providing you with all or part of the Services at any time for any or no reason.*" The Terms of Service of Blogster.com (http://www.blogster.com/terms) represent that "*BLOGSTER.COM may terminate this agreement at any time, without notice to you, if it believes, in its sole judgment, that you have breached or may breach any term or condition of this agreement, or it may terminate this agreement for its convenience.*"

Termination shall normally occur at any time – or, in very few cases, subsequently to a given period of the termination notice[82] – and is usually immediately effective in case of termination by the ISA, as the user will acknowledge termination at a time when he tries to login. In case of termination by the user, the effectiveness of the termination will be deferred until the ISP receives the user's e-mail containing the termination notice.

The lack of any notice when terminating the agreement with cause may prevent the user from recovering the profile-related data and content, which may result in the user suffering damages if the impossibility to recover data upon de-activation of the account results in a loss of data.[83] Therefore, to prevent any loss of the user's data, a termination with cause shall be served with an advance notice aimed at allowing the user a sufficient period of time to collect, or backup, his data that were stored on the ISP-hosted account.

The legitimacy of the service provider's unilateral termination right (without cause) shall not be taken for granted. From the user's standpoint, the creation of a profile on the social network aims at allowing the user to build a network of connections, communicate, express his opinions and – as discussed in the broader context of the right to the Internet access – exercise his constitutional freedoms. From this angle, granting the service provider a unilateral termination right may unjustifiably obstacle the user's exercise of the above-referred rights.[84]

---

[82] See, for example, the Google+ Terms of Service (https://www.google.com/intl/en/policies/terms/, version dated 14 April 2014): "*If we discontinue a Service, where reasonably possible, we will give you reasonable advance notice and a chance to get information out of that Service.*"

[83] C. PERLINGIERI, *Gli accordi tra i siti di social networks e gli utenti cit.*, p. 219. To give an idea of the risks implied by a de-activation without notice, the Author contemplates the following scenarios: "*Consider the de-activation of an account by a professional-oriented social network, or a politics-related social network shutting down – even with cause – an account during the elections period.*"

[84] See C. PERLINGIERI, *Gli accordi tra i siti di social networks e gli utenti cit.*, p. 218. To underscore the risks that are at stake, the Author makes the following example: the ISP's unilateral termination right "*obstacles the exercise of expressions that are crucial for the personal development and negatively affects the exercise of the political rights*

De-activation shall, therefore, be a justifiable and proportional remedy to the user's conduct.

Questions also arise as to whether the de-activation shall be deemed to be final or whether, subsequently to the de-activation of his account, the user can be admitted to later create a new one. The social networks' terms of service are silent on the matter, and therefore a prohibition for user to create a new account does not seem to apply, provided that the ISP will always be in a position to subsequently shut down also the newly created account, if the user fails to comply with the applicable policies again. Also, considering that children are generally allowed to register an account when they are thirteen or older,[85] the de-activation (with cause) of the account created by an underage does not appear to constitute legitimate ground for preventing the user from creating a new account when he comes of age.[86]

What matters for the purposes of this paper is, however, to assess who ultimately retains, and/or is granted access to, the user's data subsequently to the cancellation of the account: §5.1 analyses a scenario where the account

---

potentially resulting in a scenario – not so unlikely – where the results of political elections in cases where, for example, the ISP de-activates the profile of a candidate."

[85] See, for example, the Facebook Terms of Service: "*You will not use Facebook if you are under 13.*"

[86] R. CATERINA, *Cyberspazio, social network e teoria generale del contratto cit.*, p. 97 poses questions regarding the annullability, pursuant to Art. 23 of law of 31 May 1995 no. 218, of the agreements entered into with an Italian underage. Based on the Italian legal doctrine's argument that minors can effectively perform any everyday life, legally relevant, trivial actions (*atti minuti della vita quotidiana*), such as buying a bus ticket or a newspaper (see M. CINQUE, *Il minore contraente – Contesti e limiti della capacità*, Padova, 2007, p. 101), it should be consequently assessed whether the creation of a social network account, and the concurrent execution of a social network agreement, is suitable to qualify as an everyday life action that can be effectively performed by a minor. In this respect, other Scholars underscored that the minor shall be deemed to be legally capable of effectively entering into any agreements that are instrumental to enjoying the constitutional rights and fundamental freedoms, on grounds that Art. 2 of the Italian Constitution does not permit to disregard the choices made by an individual that is *compos mentis* (*capacità di intendere e di volere*) (F.D. BUSNELLI, *Capacità e incapacità di agire del minore, in Dir. fam.*, 1982, p. 54), so that R. CATERINA, *id.*, p. 98 concludes that the social network agreement could reasonably fall within those actions that the minor can effectively undertake, as it is aimed at enjoying, among others, his rights to communicate, express, *etc.*

was cancelled (either upon the user's choice, or autonomously by the service provider); §5.2 discusses the possible allocation of rights on the user's data subsequently the user's death, which imports a situation where the account (together with the data stored therein) is not automatically cancelled.

5.1 The Internet service provider's rights: issues concerning data retention following the account's deletion

The Internet service provider reserves the right – following the account's cancellation – to retain and use, either for a limited period of time, or without limit of time, the backup copies of the user's page content and to allow third parties to use the links underlying the social «*plug-in*» (*e.g.* the «*like*» or «*share*» options), and ultimately to re-assign or rename the nickname or URL («*Uniform Source Locator*») associated with the profile to allow a search engine indexing.[87]

From a purely contractually-driven analysis, any data retention and use by a service provider subsequently to the cancellation of the account may not rely on legitimate legal grounds: if we consider the user's grant of license on his data to the service provider as the consideration for the use of the software

[87] See, for example, the Facebook Help (How do I permanently delete my account, https://www.facebook.com/help/224562897555674?helpref=faq_content), whereby it is stated that "*If you don't think you'll use Facebook again, you can request to have your account permanently deleted. Please keep in mind that you won't be able to reactivate your account or retrieve anything you've added. Before you do this, you may want to download a copy of your info from Facebook. Then, if you'd like your account permanently deleted with no option for recovery, log into your account and let us know. When you delete your account, people won't be able to see it on Facebook. It may take up to 90 days from the beginning of the deletion process to delete all of the things you've posted, like your photos, status updates or other data stored in backup systems. While we are deleting this information, it is inaccessible to other people using Facebook. Some of the things you do on Facebook aren't stored in your account. For example, a friend may still have messages from you even after you delete your account. That information remains after you delete your account.*" Facebook represents that the data retention period is due to technical reasons, and aims at allowing the user a sufficient period of time to recover his data, if he wishes so. In 2011 Facebook faced allegations that its data retention goes beyond what is stated in the policies, meaning that Facebook retains more data than it shows, without the user's consent. The case was dealt with by Irish data protection authority, but did not result in any decision, hence the complaints were taken back (see: http://europe-v-facebook.org/EN/Complaints/complaints.html).

and services made available by the service provider, the termination of the agreement that is consequent to the cancellation of the account shall imply that the provider's exploitation of the user's data and the user generated content would take place in the lack of any provision of services to the user's advantage.[88] Put it otherwise, any exploitation of the user's data or profile would lack the fundamental requirement consisting in the reciprocity of obligations that is at the substance of any bilateral contract.

In addition, from a data protection standpoint, any endless, or unjustified, data retention may clash with the general principle that data can be retained for no longer than is necessary for the purposes for which the data was collected or subsequently processed (see Art. 11, para. 1, let. e) of the Italian Data Protection Code).

However, the arguments developed above at §4.1 do in my opinion support the opposite conclusion that the service provider may legitimately retain the user's data, on accounts that the user's data was the consideration for the provisions of services by the service provider.

The possibility for the service provider to retain data shall, of course, be subject to a number of boundaries. First, the user's right to prevent any data usage that is against his dignity remains unprejudiced and enforceable at any time. Second, the service provider shall retain only the data that was knowingly conferred by the user, it being among the service provider's duties to make an advance (*i.e.* in the terms of service) full disclosure of the nature and kind of data that it will retain subsequently to cancellation. Third, the service provider must not use the data for any purposes other than those of which the user was informed at a time when the user created the account.

On this basis, it becomes questionable whether the service provider, or any third party acquirer may continue providing tailored or behavioural advertising to the user subsequently to the cancellation of the account,

[88] C. PERLINGIERI, *Gli accordi tra i siti di social networks e gli utenti cit.*, p. 218.

because such activities can be legitimately performed in most jurisdictions exclusively based on the data subject's consent and, in the absence of an account, the user would not have any technical tool to opt-out, *i.e.* to withdraw his consent, which is a right the user shall be in a factual position to exercise at any time.

The user may also exercise at any time his right to access the data the ISP has retained, seeking the rectification or cancellation of such data (see Art. 7, para. 3, let. a) and b) of the Italian Data Protection Code). I do not believe that this user's right clashes with the service provider's right to further retain the user's data, in an anonymised and compiled form, within the Big Data archives the service provider has created.

5.2 The «*virtual heritage*»

Another possible scenario to be taken into account is the Internet service provider's and/or a third party's right to retain or access (as the case may be) the user's data in a scenario where the user dies.

The issue of *post mortem* allocation of the login details and rights on the web account is increasingly raising the attention of the web community, as there are reportedly more than 30 million Facebook accounts of users who have passed away.[89] However, the majority of the social networks' terms of service are almost silent on the consequences that the user's death imports on the account:[90]

(a) Twitter does not contemplate the scenario of a deceased user, and only retains the right to remove the account "*due to prolonged inactivity;*"[91]

---

[89] Source: http://www.bbc.com/future/story/20160313-the-unstoppable-rise-of-the-face-book-dead.

[90] On the contrary, the user's mere inactivity for a given period of time may result in the account's cancellation, so that the arguments developed in §5.1 may still apply in this case.

[91] See Twitter Terms of Service https://twitter.com/tos?lang=en

(b) Facebook's policy is to convert accounts of the deceased into "*memorialized*" accounts, but immediate family members with the correct documentation are able to request that an account be removed entirely. The user can also choose in advance to have his account permanently deleted, should he pass away, in which case the deletion policy described in §5.1 above shall apply;[92]

(c) LinkedIn closes the deceased user's account only if another user so informs and requests LinkedIn;[93]

(d) Yahoo! and Flickr adopt a similar approach, but refuse to pass on login and passwords to the accounts to the user's heirs;[94] and

(e) Google also only contemplates the possibility to pass on the contents of a Gmail account to the heirs, rather than passing login details, and only if specific circumstances apply.[95] However, in addition to this general policy, the "*Inactive Account Manager*" enables the users to

---

[92] See Facebook Help – Report a Deceased Person (https://www.facebook.com/help/-408583372511972/).

[93] See LinkedIn Use Agreement; Deceased LinkedIn Member – Removing Profile (https://www.linkedin.com/help/linkedin/answer/2842/deceased-linkedin-member-removing-profile?lang=en, version dated 3 November 2014): "*Unfortunately, there may be a time when you come across the profile of a colleague, classmate, or loved one who has passed away. If this happens, we can close that person's account and remove their profile on your behalf. We'll need you to gather: The member's name The URL to their LinkedIn profile Your relationship to them Member's email address Date they passed away Link to obituary Company they most recently worked at To start this process, please answer some questions about the person who has passed away by filling out this form. After you fill out this form, it will be automatically sent to us for review and we'll be in touch.*"

[94] See Yahoo! Help (Options available when a Yahoo Account owner passes away, https://help.yahoo.com/kb/SLN2021.html): "*Unfortunately, Yahoo cannot provide passwords or allow access to the deceased's account, including account content such as email. At the time of registration, all account holders agree to the Yahoo Terms (TOS). Pursuant to the Terms, neither the Yahoo account nor any of the content therein are transferable, even when the account owner is deceased.*"

[95] See Google Accounts Help (https://support.google.com/accounts/troubleshooter-/63575-90?visit_id=1-636273439418507643-2835791149&hl=en&rd=2): "*If an individual has passed away and need access to the contents of his or her email account, in rare cases we may be able to provide the Gmail account content to an authorized representative of the deceased user. We extend our condolences and appreciate your patience and understanding throughout this process*"

share "*parts of their account data or to notify someone if they've been inactive for a certain period of time*," meaning that the user can nominate trusted contacts to receive data if the user has been inactive for the time specified by him (three to eighteen months). The trusted contacts are entitled to download data the user left them. The user can also decide to only notify these contacts of the inactivity and decide to have all data deleted.[96]

Most of the abovementioned protocols governing the access to the deceased user's account and data are embodied in the "*Help*" sections of the platforms, rather than in the "T*erms of Service*," so it is arguable whether these protocols are intended to be binding on the service provider or enforceable by the deceased user's family or heirs, or if these policies merely qualify as a statement of good practice.[97] In any event, at least pursuant to the general principles of Italian law, the heirs shall be entitled to access the account, also in cases where the user has previously opted for its cancellation.

But, most of all, it is questionable whether the heirs, once they are granted access to the account of the deceased user, have a right to manage and use the account.

An increasing number of the U.S. States are considering the opportunity to adopt the Uniform Fiduciary Access to Digital Access Act (UFADAA), which was proposed in 2015 by the Nebraska lawmakers with the intended purpose to grant a deceased person's representative (*e.g.* the executor of a will, or a trustee) broad authority to access, control, and manage a deceased person's digital assets, the definition of which would also include the deceased person's web accounts: "*Not only does UFADAA give personal representatives authority to manage digital assets, it gives personal*

---

[96] See Google Accounts Help (https://support.google.com/accounts/answer/303-6546?hl=en).

[97] C. MACIEL, V. CARVALHO PEREIRA, *Digital Legacy and Interaction: Post Mortem Issues*, New York, NY, 2013, p. 129.

representatives the power to demand access to accounts and files from the companies that control them. So under UFADAA, a personal representative is authorized to use a deceased person's username and password to go into an email account. And further, if the personal representative does not have the username or password, he or she can demand access from the email provider."[98]

There still remains a legislation gap in Italy, considering that the *Codice dell'Amministrazione Digitale* encompasses within the definition of «*digital identity*» (*identità digitale*) exclusively the digital identification card (*carta di identità elettronica*) and the digital signature, while the Web 2.0 has undoubtedly reshaped the concept of digital identity to include a much wider range of interests, *i.e.* both (i) the profile and reputation an individual builds into the Web, which may result in a number of personalities associated with one single user,[99] and (ii) the "*information on an entity used by computer systems to represent*" the user,[100] that is all the web services that are suitable to identify the activities and personal identity ascribable to a user into the Web.

Any assessment on the possibility for the account and the data stored therein to circulate and be available following the death of the user is then strictly related to the so-called right to be forgotten discussed in chapter I (§2), which right to be forgotten shall evidently be interpreted to include the user's right to delete all his personal data that was made available into the Web also in cases where such personal data are not detrimental to his reputation and honour.

---

[98] Source: http://www.nolo.com/legal-encyclopedia/what-happens-digital-assets-nebraska.-html.

[99] S. RODOTÀ, *Tecnopolitica. La democrazia e le nuove tecnologie della comunicazione*, Roma-Bari, 2004, p. 139.

[100] This is the Wikipedia's definition of digital identity (https://en.wikipedia.org/-wiki/Digital_identity), which definition recalls the definition set forth in ISO/IEC 24760-1: a "*set of attributes related to an entity*."

But, also giving for granted that the heirs can seek the deletion of the deceased person's data from the Web, there remains a technical and legal possibility for the social network to retain the data the deceased person provided thereto pursuant to the social network agreement, along with the factual difficulty for the heirs to seek cancellation against all the third parties (advertisers and data brokers) which were communicated such data by the social network.

6. Conclusion

Despite the lengthy terms of service provided by the social networks and the other Internet 2.0 service providers, there remains an information asymmetry between the users and the ISPs as to what data the latter actually collects and retains.[101]

Such data, as said, feeds the ISPs' archives and databases, and is destined to be exploited in a wide variety of uses, of which the user that originally provided the data usually loses track.

Against this background, once the service provider has obtained (almost permanent) access to the users' data, the issue a service provider faces is twofold: (i) how to make the most of such a huge amount of data, which must be analysed, mined and compiled in a reasoned manner, and (ii) how to ensure protection to the results of such a collection and compilation effort. While the issue as per (i) is mostly business and technically related, the issue triggered in limb (ii) may be suitably analysed from a legal standpoint: chapter III deals with this second issue.

[101] On the disparity of conditions to which users and service providers respectively undertake to comply, see R. CLARIZIA, *Contratto informatico (per l'oggetto e per il mezzo)*, in *Enc. dir. (agg. II)*, 1998, p. 245 *et seq.* and G. FINOCCHIARO, *I contratti informatici*, Padova, 1997.

CHAPTER III

BIG DATA: EXPLOITATION AND ENFORCEMENT

1. Foreword

As anticipated in chapter II, companies tune their business decisions and their business practice, including product and services life cycles, based on what data mining shows them to be the consumers' most profitable demand trends, the most efficient processes, the foreseeable economic changes and the main business risks.

Now, as many companies' business relies on data analytics, the issue comes at identifying the strategy companies may follow to protect their investment in Big Data analysis.

The aim of this chapter is to provide an overview of the possible legal schemes under which a company's investment in Big Data analytics may find protection. However, the less rights in Big Data are suitable to fall within the traditional legal schemes of civil law, the more complicated answering this question becomes. Hence, the lack of a unitary answer by a consolidated case law and legal doctrine suggests to take a glance at the solutions that the jurists across different jurisdictions have identified to be the most suitable to protect the rights and investments in Big Data through intellectual property laws.

A relevant portion of this chapter therefore acknowledges the debate on the applicability to Big Data of patent protection *vis-à-vis* secret know-how, a debate which has developed in the United States of America following *Alice Corp v. CLS Bank Intl.*, a landmark case which will have material implications on the patentability of those data analytics software on which Big Data extensively relies.

Assessing the US approach to Big Data enforcement is, overall, quite a necessary step to take, on accounts that the greatest share of Big Data is arguably available to the US-based Internet service providers mentioned in chapter II (Google, Facebook, Yahoo!, Microsoft, *etc.*) and to the other US-

based multinational corporations, the business of which pivots around data analytics (*e.g.* Apple, IBM, Amazon, Tesla, just to name very few of them).[1]

2. How can data be legally protected?

2.1 What is «*data*» in legal terms?

The starting point for a discussion about the legal framework for Big Data is the question as to what is the nature of information and data.

Preliminarily, a distinction may be made that "*information is that which informs and is expressed or conveyed as the content of a message or arises through common observation,*" while "*data is digital information.*"[2] However, the terms «*information*» and «*data*» can be (and have actually been) used almost interchangeably in this paper, as this paper deals with those huge datasets of digital information which qualify as Big Data.

From this perspective, the lack of any boundary or limit makes data a «*non-rivalrous commodity:*» this means that one person's use of data does not necessarily prohibit or reduce the value of use of said data by another person.[3]

---

[1] In 2014 Boston Consulting Group published a study on the world's most innovative companies, whereby it concluded that "*companies who are strong at innovation are three times more likely to rely on big data analytics and data mining than their counterparts who are less adept at innovating (57% versus 19%). 67% of breakthrough innovators say their big data analytics and data mining efforts are paying off*" (see BCG's study *The Most Innovative Companies 2014: Breaking Through Is Hard to Do*, available at https://www.bcgperspectives.com/most_innovative_companies). Not surprisingly, the list of the most innovative companies relying on data mining and data analytics includes U.S. companies for the vast majority (seven in the top 10). In 2016 the website Network World provided a list of "*the real big data and analytics companies … Most focus on helping companies make sense of their oodles of data, sometimes for customer service, sometimes for IT purposes and sometimes for security reasons.*" The list includes mostly start-ups and fund-driven companies, such as StreamSets, SnapLogic, Saama, Periscope Data, Big Panda, Fuzzy Logix (see http://www.networkworld.com/article/3021350/big-data-business-intelligence/15-big-data-and-analytics-companies-to-watch.html).

[2] R. KEMP, *Legal aspects of managing Big Data*, in *Computer Law & Security Review*, 2014, p. 486.

[3] J. DE WATCHER, *Big Data and IP Business Strategy*, in *Trading Secrets – Web portal TradeSecretsLaw.com*, 2014. The Author considers that the following circumstances lead to challenge the traditional concept of data ownership: "*a) most data are generated by someone else, and b) the value of data increases by their use, not the restriction on their use.*"

Hence, as anticipated in chapter I, §4.1, subjecting a potentially «*limitless*» resource like data to legal rules about *ownership* would appear to be incompatible with the same nature of data. This is the outcome of the UK criminal law case of *Oxford vs Moss*: there is no property *in* data, as data *per se* cannot be stolen.[4]

Commentators, however, recognise that – although there are no rights *in* data, just the same as it is not possible to patent an idea – rights and obligations do arise *in relation to* data. The German Civil Act, for example, expressly provides that things which are neither «*rights*» nor «*goods*» may nevertheless be sold within a sale contract.[5]

On this basis, data can be undoubtedly sold, and data can be the subject matter of a number of transactions: above all, *disclosure*. Disclosure – which can occur in the form of a license grant – is being increasingly recognised as a means to extrapolate value from data, by granting access to those who can analyse data and find the most valuable and reliable judgments therein.[6]

---

[4] *Oxford v Moss* (1979) 68 Cr App Rep 183. The defendant, Moss, was a university student, who managed to obtain a copy of his forthcoming exam paper. He was charged with stealing information belonging to the Senate of the University. The Divisional Court, Queens Bench Division ruled that confidential information was not a form of property as defined by Section 4(1) of the Theft Act 1998, and that confidence consisted in the right to control the publication of the proof and was a right over property rather than property in itself.

[5] See Section 453 of the German Civil Act.

[6] For an ample overview of market practices in the field of disclosure of Big Data see M. MATTIOLI, *op. cit.*, pp. 535-583. In commenting the relevance of disclosing Big Data J. DE WATCHER, *op. cit.* observes that "*a lot of the value of Big Data depends … on the ability to have access, and preferably open or free access, to as much data as possible. This means that there is a natural market-driven pressure to businesses, in a Big Data environment, to prefer the use (and therefore an open approach) to data, rather than to limit or restrict use. While it is true that access to certain data can be very valuable, this approach is typically based on the assumption that one knows the data available, and understands at least the most important value considerations in respect of these data. This is where Big Data presents an important shift: not only does it become much harder to know who owns or generates which data, or what is in those data, it also becomes much riskier not to grant relatively free access to data. This is because there is a lot of relevant, but not necessarily obviously visible, value in the data. A lot of value in Big Data comes from recombining data from different sources, or approaching data in a different way (e.g. compressing data in a visual or topographic way in order to discover new patterns). As a result, businesses that open up their data are more likely to retrieve*

Hence, there must be a form of reward for the effort to create or gather data, allowing the dataright holder to block or charge for access or use.

The distinction drawn in the previous chapters between *ownership* and *access* may make more sense at this point: while it is difficult to reconcile the notion of *ownership* with that of digital information, regulating *access* to digital information may be the key to protect investment in data analytics.

2.2 The traditional approach to legal protection of data: the «*tangible property*» test

Since the *Corpus Juris Civilis* of Roman law, civil law has been based on the distinction between «*rights*» and «*goods:*» while goods were considered to be transferable from the owner to an assignee – goods being «*tangible*» and «*movable*» –, rights were not.[7]

Against this background, over the past century, discussions have focused on the conditions subject to which the concept of tangible property could be extended to intangibles, such as data.

US and European courts sometimes applied the «*tangible property*» test, according to which data deserves protection insomuch it is integrated with the physical device on which it is recorded, only the hardware being susceptible of qualifying as an «*owned asset*.».[8]

---

value from those data, and those that do, will retrieve more value from the data that is most open and accessible."

[7] The *Institutiones Justiniani* (Justinian's Institutes) make a distinction between *res corporales* (goods) and *res incorporales* (rights): among the *res incorporales* (usufruct, inheritance, obligations) there was not *ownership*, as only *res corporales* were in principle suitable to be transferred (*traditio*). See U. VINCENTI, *I modelli dell'appartenenza*, in *Diritto privato romano* (a cura di A. SCHIAVONE), Milano, 2003, pp. 273 *et seq.*.

[8] *Retail Systems Inc v. CNA Insurance Cos 469 N.W. 2d 735 (Minn. App. 1991)*: "*Other considerations also support the conclusion that the computer tape and data are tangible property under this policy. The data on the tape was of permanent value and was integrated completely with the physical property of the tape. Like a motion picture, where the information and celluloid medium are integrated, so too were the tape and data integrated at the moment the tape was lost.*"; *American Guarantee and Liability Insurance Co v. Ingram Micro, Inc 2000 WL 726789 (D. Ariz, 2000)*. For a classification as intangible, see *AOL v. St Paul Mercury Insurance 207 F. Supp. 2d 459*

In *Thyroff v. Nationwide Mutual Ins.*[9] the New York Court of Appeals applied for the first time the «*tort of conversion*»[10] model to an employee's claim that his personal data and e-mails had been illegitimately seized by the employer. While in *Sporn v. MCA Records*[11] the same Court of Appeals had previously held that "*an action for conversion will not normally lie when it involves intangible property*," in the *Tyroff* case the Court took a different view. By acknowledging that in the current society "*computers and digital information are ubiquitous and pervade all aspects of business, financial and personal communication activities*," the court affirmed that a claim for conversion of electronic records and data shall be possible under New York law. The court did not further investigate who shall be considered the proper *owner* of the electronic information at issue, but simply presumed that the plaintiff legitimately owned data based upon tangible property in the hardware and the possession of technical devices.

Following the *Thyroff* judgment, the US Bankruptcy Court of the Southern District of Texas[12] decided that the copying of seismic data stored on a computer may be subject to a conversion claim. The data "*could not*

(E.D. Va 2002); aff'd No.02-2084 (4th Cir. 2003) (St Paul Mercury). For an overview of the decisions which applied the tangible property test, see T. HOEREN, *Big data and ownership in data: recent developments in Europe*, in *European Intellectual Property Review*, 2014, pp. 751-752.

9 *Thyroff v. Nationwide Mutual Ins. Co* N.Y. 3d 283 (2007). The plaintiff was an insurance agent for the defendant employer Nationwide, who entered into an arrangement whereby Nationwide would lease him computer hardware and software (the "AOA system"). The purpose of the AOA system was "*to facilitate the collection and transfer of customer information to Nationwide*." The plaintiff used this technology for business data, but also for personal e-mails, correspondence and other relevant customer data. Nationwide uploaded daily all the information from the plaintiff's computer system on to its centralised computers and the plaintiff subsequently raised a claim protesting against the seizure of his personal information.

10 In common law jurisdictions «*conversion*» is a voluntary act by one person inconsistent with the ownership rights of another. The plaintiff has to claim ownership or immediate superior right of possession and must prove that the defendant exercised unauthorised dominion over the property in question to the exclusion of the plaintiff's rights.

11 *Sporn v. MCA Records 58* N.Y. 2d 482, 489 (1983).

12 *In re Yazoo Pipeline Co LP* 459 B.R. 636 (Bankr. S.D. Tex. 2011).

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*". The court furthermore concluded that although data was stored in an electronic format for efficiency reasons, it "*could have been represented through other, indisputably tangible, media*."

However, "*the concept of tangible property leads to nowhere, especially in the age of the Internet,*" as it might lead to strange effects if the owner of the *medium* gets the property rights to the data stored on the *medium*: such principle would validate the unacceptable conclusion that, for example, an owner of a blank DVD would become the owner of all the data stored on that DVD, irrespective of the source of said data.[13] The contrary is probably true, as devices do not *per se* have relevant business value, so long as they are deprived of the information stored therein and/or a connectivity, which makes it possible to upload information on the devices.[14]

On these accounts, the tangible property approach has been denied by UK courts recently. For instance, in *Your Response Ltd v. Datateam Business Media Ltd*[15] the UK Court of Appeal come to the conclusion that a lien (meaning, for the purposes of UK law, a right entitling a person in possession of a good to retain it in certain circumstances) does not subsist over a database. In this case, an IT maintenance company claimed a lien over the database of the defendant, pending payment of the contractually agreed fees. Contrary to what the High Court had decided in the first instance proceedings, the Court of Appeal ruled that the common law clearly distinguished between

---

[13] T. HOEREN, *op. cit.*, p. 752, who considers that "*Even if there was a «marriage» between data and their computer readable substrate this concept cannot be used in the internet world where these substrates no longer exist.*"

[14] R. MORO DI VISCONTI, *Internet delle cose, Networks e plusvalore della connettività*, in *Dir. ind.*, 2016, p. 540. See also *ID., Valutazione dei Big Data e impatto su innovazione e ditigal branding*, in *Dir. ind.*, 2016, pp. 46-53.

[15] *Your Response Ltd v. Datateam Business Media Ltd* [2014] EWCA Civ 281; [2014] 3 W.L.R. 887.

«*tangible*» and «*intangible*» property. A lien was only possible over tangible property. An electronic database did not fall within that category. The court acknowledged that there would have been sensible arguments to extend liens to digitalised materials, but as this would involve a significant departure from existing law, this would need Parliament to change the law. It may be further inferred that databases apparently are not regarded as *property* by British courts.

2.3 The «*origination*» criterion

The German courts focused on a criterion other than tangible property: data might be related to the person who generates, creates or produces it.

The German Criminal Act provides that unlawfully erasing, corrupting or altering computer data is a crime.[16] Against this provision, the Criminal Court of Appeal of Nuremberg[17] dismissed the theft charges against employees who had deleted the data stored on their company-owned laptops. As Section 303 of the German Criminal Act does not clarify the conditions subject to which data may qualify as proprietary to the employer, the Court of Appeal of Nuremberg found the borderline between lawful destruction of an individual's own data and the criminal erasure of data belonging to somebody else in the so-called «*skripturakt*» doctrine, pursuant to which property on data is vested with him who generates the data.[18]

The application of this principle to civil law cases, however, requires a certain degree of flexibility, as the «*skripturakt*» doctrine may lead to cases

---

16 See section 303 of the German Criminal Act.

17 OLG Nürnberg 1. Strafsenat, Beschluss vom 23.01.2013 — 1 Ws 445/12, CR 2013.

18 According to this principle, it would be unclear what happens in the case of automatically generated data. In this case it is not a person who is generating the data, but a "machine." Certain jurisdictions, such as the UK, provide that in the case of a literary, dramatic, musical or artistic work which is computer-generated, the author shall be taken to be the person by whom the arrangements necessary for the creation of the work are undertaken (see Art. 9(3) of the UK Copyright, Designs and Patent Act 1988). This principle may be probably applied also to automatically generated data, as outlined in chapter II, §1, above.

where an employee may be allowed to delete the data (under criminal law), but he might still be dismissed (under labour law).

As a consequence, the «*origination*» criterion might be amended by additional civil law rules, such as the principle of work (better, data) «*made for hire:*» if one has been employed or retained to generate data (for the employing company or a client), the right to retain and dispose of the data should be attributed to the employer.

On this basis, the Labour Court of Appeal of Saxony [19] (*Landesarbeitsgericht*) held that the deletion, by an employee, of the software that the employee purchased and installed on a company-owned computer was legitimate ground for dismissal. The court relied on the civil law principle [20] that, by virtue of the employment contract, the employer had become the legitimate dataright holder for all data created or acquired by the employee in the performance of his job. The employer had acquired property in the software and the related data, so that the deletion of the software was legitimate ground for dismissal.

Putting the two German cases together, it can be inferred that, as a general principle, the rights in data are attributed to the data originator. However, there are a number of circumstances and exceptions (*e.g.* the subsistence of an employment contract), which can subsequently alter this presumption and which therefore need to be assessed on a case-by-case basis. Considering that rights and duties that arise in relation to data are developing at the rapid pace

---

[19] LAG Sachsen, Urteil vom 17.01.2007 — 2 Sa 808/05, MMR 2008, 416. An employee purchased and installed Microsoft Outlook on a laptop which his employer had provided to him. Then the employer went on a sick leave for a long period of time, at the end of which his employer sought for the laptop and all the e-mails related to the employee's job. The employee returned the computer and the e-mails, but the computer did not feature Outlook, which the employee had meanwhile. The court held that this was reason for dismissing the employee: following the installation of Outlook on the employer's laptop, the employer had obtained the property in the software. By deleting Outlook from the computer, the employee had destroyed the employer's data and thus he could be dismissed.

[20] See Art. 950 of the German Civil Act.

of Big Data techniques, it becomes of essence to identify who becomes the owner of data also subsequently to the creation of said data.

In this respect, *UsedSoft GmbH v. Oracle International Corp*[21] may provide significant input for future decisions. In this landmark case the Court of Justice of the European Union 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 (provided, of course, that the licensee does not further exploit the software). The decision implies that there is a specific *ownership* right relating to intangible goods like software downloaded via the Internet. Although the applicability of this model to other digital goods remains to be considered in future court decisions, the Court of Justice has at least opened the door for a discussion on ownership in intangible assets.[22]

2.4 Finding legal protection in intellectual property law

Based on the foregoing, enforcing rights based on an alleged «*data ownership*» may prove to be quite a difficult and burdensome route to pursue for companies investing in Big Data.

On a related matter, Big Data is not an inherently "*intellectual property problem.*" But, as intellectual property law is concerned with problems of technological development and disclosure, intellectual property law may constitute useful ground for enforcing rights arising *in relation to* data. This is the (limited) scope of «*data law*» as an emerging new area in its own right around intellectual property rights.[23]

The following paragraphs, therefore, outline how intellectual property law may help find a legal basis for enforcing rights arising in relation to data

---

[21] Court of Justice of the European Union, Judgment of 3 July 2012, (C-128/11) *UsedSoft GmbH v. Oracle International Corp*

[22] T. HOEREN, *op. cit.*, p. 753.

[23] R. KEMP, *Legal aspects of managing Big Data cit.*, p. 486.

based on database and copyright (§3), patent (§4) and trade secret (§5) rights. Finally, §6 briefly discusses possible intellectual property law-based remedies for enforcing said rights.

3. Database and copyright

3.1 Copyright and *sui generis* database protection in the European Union

Database protection does not protect data as such, however it safeguards the originality embodied in, or the investment made for, collections of data which are systematically or methodically arranged and can be individually accessed.

The EU directive 96/9/EC (the "Database Protection Directive") [24] introduced a framework for copyrighted protection of database, as follows:

(a) copyright in original database (Art. 3) (see §3.2.1 below); and

(b) *sui generis* right in database in which an investment has been made (Art. 7) (see §3.2.2 below).

3.2.1 *Original database*

A compilation of data enjoys protection under copyright law as a database, which is "*a collection of independent works, data or other materials arranged in a systematic or methodical way and individually accessible by electronic or other means*" under the Database Protection Directive.

The author or creator of the database may therefore enforce copyright protection not on the content (*i.e.* the data), but on the database itself, as long as the latter "*constitutes the author's own intellectual creation*" by reason of the "*selection or arrangement*" of the data stored in the database. In *Football Dataco v. Yahoo!*, a case dealing with football match schedules,[25] the Court of Justice of the European Union ruled that the criterion of author's own intellectual creation for database copyright requires that the author expresses

[24] Directive 96/9/EC of the European Parliament and of the Council of 11 March 1996 on the legal protection of databases.

[25] Court of Justice of the European Union, Judgment of 1 March 2012, (C-604/10) *Football Dataco v. Yahoo!*

his "*creative ability in an original manner by making free and creative choices.*" The requirement, however, is not met in cases where the setting up of the database is dictated by technical constraints which leave no room for creative freedom.

Based on these criteria, the Court of Justice further ruled that Art. 3 of the Database Protection Directive is suitable to cover, among others, an anthology of poems, which was used to produce a CD-Rom of a collection of poems (*Directmedia v. Albert-Ludwigs-Universitaet*).[26]

Protection in the database applies automatically once the database exists in material form and will last for 70 years after the end of the calendar year in which the author died.

Temporary or permanent reproduction, translation, adaptation or alteration of the database, or distribution or communication to the public of copies thereof, without authorisation of the owner will result in infringement of copyright in the database.

3.2.2 *Database in which an investment has been made*

The person who "*takes the initiative and the risk of investing*" – or, if the work is subcontracted, the commissioner of the sub-contract[27] – deserves a *sui generis* protection over the database which involves investment under Art. 7 of the Database Protection Directive.

As clarified by the Court of Justice in *British Horseracing Board v William Hill,*[28] in order to qualify for protection, there must be substantial investment in the database: resources spent in creating the data in the database cannot be taken into account, while proof of the costs incurred, and efforts made, to seek existing independent materials and collecting them in the

---

[26] Court of Justice of the European Union, Judgment of 9 October 2008, (C-304/07) *Directmedia v. Albert-Ludwigs-Universitaet*

[27] See Recital 41 of the Database Protection Directive.

[28] Court of Justice of the European Union, Judgment of 9 November 2004, (C-203/02) *British Horseracing Board v. William Hill*

database is required. The maker or rightholder of the database (where a company or firm) must have its registered office, central administration or principal place of business within the European Union.

On this basis, the *sui generis* right is likely to cover marketing databases (irrespective of their complexity, which would be relevant for the purposes of copyright protection under Art. 3), a record of live actions occurring during a sporting event (*Football Dataco v. Sportradar*)[29], case law databases, databases of rainfall in particular locations, databases of contact details for people working in doctors' surgeries, collection of car advertisements hosted on a website.[30]

The right arises upon creation of the database and lasts 15 years from its creation or (if later) it being made available to the public. A new 15-year period right commences when substantial updates or changes are made to the database.

It will be an infringement of the Art. 7 *sui generis* right in database to extract («*extraction*» meaning the permanent or temporary transfer of all or a substantial part of the contents of a database to another *medium* by any means

---

[29] Court of Justice of the European Union, Judgment of 18 October 2012, (C-173/11) *Football Dataco v. Sportradar*. The Court of Justice, in a case involving a database of data gathered live during football matches, held that a database owner can sue an online infringer of database right where Internet users whom the infringer intends to target are located.

[30] See DLA Piper's 2014 report *IP Rights in Data Handbook*, available at https://www.dlapiper.com/en/uk/insights/publications/2014/09/ip-rights-in-data-hand-book/. See also the following decisions of the **French** courts: 1) In *Cour de cassation*, 5 March 2009, *Precom and Ouest France v. Direct Annonce* the French Supreme Court refused the protection of *sui generis* database right to a database of real estate advertisements included in the different versions of a newspaper, because the investment was not related to obtaining the contents of the database but to the creation of the items included in this database and the purely formal verification operations, during this creation phase; 2) in *Cour d'appel de Paris*, 15 November 2013, *Pressimmo online v. Yakaz and Gloobot*, the Court of Appeal of Paris refused the protection of *sui generis* database right to a website of real estate advertisements, because the website owner only alleged having made substantial investments and did not break down the investments (obtaining, verification and presentation) and did not prove the substantiality of such investments. Moreover, the Court stated that the website owner could use the unfair competition regime to bypass the sui generis database right.

or in any form) or re-utilise («*re-utilisation*» meaning 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 online or other forms of transmission) the whole or a substantial part of the contents of that database, without the owner's consent.[31] In the abovementioned *Directmedia v. Albert-Ludwigs-Universitaet* decision the Court of Justice explained that the "*extraction*" act of infringement in relation to database right should be given a wide definition and would include 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 it.[32]

---

[31] According to C. GALLI, *Banche dati e giornali*, in AIDA, 1997, pp. 29-44 the notion of extraction should be considered from a qualitative (rather than merely quantitative) perspective. Accordingly, the parameter to ascertain whether a substantial part of the database was extracted should be the suitability that the data that was extracted may provide a complete, comprehensive, useful piece of information, such as a proven inference between two phenomena, or data relating to a group of users with homogeneous interests in view of a future profiling.

[32] In *Automobil-Onlinebörse (I ZR 159/10)* **t**he **German** Federal Court of Justice (*Bundesgerichtshof* – BGH) held that a meta-search engine, which enables individual users to conduct specific searches of protected databases, (in the case at hand an online automobile-market), does not infringe the database maker's exclusive right to "*distribute and reproduce its database*," since: (i) the meta-search engine does not reproduce a substantial part of the database (single queries only) and (ii) the repeated use of insubstantial parts of the database does "*not amount to an illegitimate reproduction of a substantial part*" (meta-search engine does not intend to create a copy of the full/substantial part of the database). This BGH's decision of 2011 contradicts the findings of the EU Court of Justice's decision on (C-202/12) *Innoweb v. Wegener* concerning the interpretation of Art. 7 of the Database Directive. The Court of Justice stated that a dedicated meta search engine may "*re-utilise the whole or a substantial part of the contents of a protected database*" – and therefore may infringe the rights of a database maker – where the dedicated meta search engine: (i) provides the end user with a search form which essentially offers the same range of functionality as the search form on the database site; (ii) "*translates*" queries from end users into the search engine for the database site "*in real time,*" so that all the information on that database is searched through; and (iii) presents the results to the end user using the format of its website, grouping duplications together into a single block item but in an order that reflects criteria comparable to those used by the search engine of the database site concerned for presenting results. The decision of the Court of Justice will most likely have an impact on any future judgments of the BGHV concerning this so-called 'screen scraping'.

3.2 Database protection in the U.S.A.

Overseas, enforcement of copyright on database relies on similar grounds as the Database Protection Directive. The Copyright Act grants protection over a «*compilation*,» which is a "*work formed by the collection and assembling of pre-existing materials or of data that are selected, coordinated, or arranged in such a way that the resulting work as a whole constitutes an original work of authorship*."[33] The copyright in a compilation is limited only to the compilation itself, and not to the underlying materials or data.[34]

As pointed out by the Supreme Court in *Feist Publications, Inc v. Rural Telephone Service Company*,[35] in order for a compilation of data to be protected, its selection, coordination, and arrangement must contain a degree of creativity, provided that "*even a slight amount will suffice*." Based on this principle, the Supreme Court concluded that a selection of data (names, towns, and telephone numbers) is obvious and lacks creativity in cases where names are ordered alphabetically. In so ruling, the Supreme Court rejected the so-called «*sweat of the brow*» doctrine, according to which copyright

---

[33] See Art. 17 USC 5 101 of the Copyright Act.

[34] See Art. 17 USC 5 103(b) of the Copyright Act. L.H. GUTTENPLAN, L. JOHNSTON, *Big brother, Big Data: how museums are collecting, using and storing ditigal data*, ALI CLE Course of Study Materials, 2014, pp. 2-3 explain that "*database may still be protected as a «compilation» work under federal copyright law if the selection, coordination, or arrangement of the underlying material is sufficiently creative to constitute an original work of authorship. The creativity threshold for originality is low and will likely be met as long as selection and arrangement decisions are not «so mechanical or routine that no creativity exists.» Note, that copyright protection for databases is «thin,» in that it only extends to the creative elements of the database (*i.e. *creative selection and arrangement decisions). In addition, databases have been protected as trade secrets or patents and their misuse has been prohibited under other laws such as unfair competition, misappropriation, trespass, or computer fraud (*The Computer Fraud and Abuse Act*)*."

[35] *Feist Publications, Inc., v. Rural Telephone Service Co.* 499 U.S. 340 (1991). Rural Telephone Service, a local telephone company, published telephone directories based on data from its subscribers. Feist used Rural's data to publish a "*white pages*" encompassing a much larger geographic area. Rural sued Feist for copyright infringement. The Court ruled that compilations and databases are pro tectable only when arranged and selected in an original manner. On the facts before it, the Court held that Rural's data compilation was not copyrightable: ordering names alphabetically is commonplace, and does not meet the originality threshold.

protection should be afforded to compilations of data and other materials based simply on the effort used to create the compilation.[36]

3.3 *Sui generis* right in Big Data databases

Big Data, as said, consists of large datasets collected, among others, through IoT sensors and Web platforms. To the extent input data is uploaded in the database in an automatic and "*passive*" fashion, and data is subsequently analysed and processed (again, automatically) by software, Big Data databases do not appear to meet the creativity and selectivity requirements pursuant to Art. 3 of the Database Protection Directive.

Italian commentators have excluded that the scope of application of Art. 2, para. 9, of law of 22 April 1941, no. 633 (the "Italian Copyright Act") – which implements Art. 3 of the Database Protection Directive – may include a Big Data collection, in cases where data is not "*systematically or methodically arranged*." One may still argue that the data compilation for which the software is responsible may, in certain cases, be considered itself to be a «*work*,» in that the software is capable of cataloguing – and finding relations between – data in a fashion that is original and not commonplace.[37] However, this argument may ultimately be inconsistent with the *rationale* underlying Art. 2, para. 9, of the Italian Copyright Act, according to which copyright protection shall be granted exclusively to those databases that result from an

[36] Unlike the US and the EU, there is no database protection right in Canada, so that no additional protection is afforded to creative databases. Protection to databases is afforded under copyright law and case law. See J. TSOUKAS, N.J. VERMETTE, *Intellectual Property Protection for Big Data in Canada and the United States*, in *Information Law Journal*, 2016, 7(3), p. 19.

[37] See *Trib. Bologna Sez. spec. propr. industr. ed intell.*, 10 August 2011 (*Porfiri G. c. Franca Cosimo Panini Editore S.p.A*). Although there are no Italian case precedents focussing on Big Data, the Italian case law is consistent in stating that mere collections are not databases for the purposes of Art. 2, para. 9 of the Italian Copyright Act: for instance, in the cited case, the Court of Bologna ultimately did not find a collection of funny and ironic SMS messages to be a database under the Italian Copyright Act.

original selection of data and materials, with a complex structure that is truly expressive of the author's personality.[38]

Conversely, Big Data is likely to enjoy the *sui generis* protection afforded under Art. 7 of the Database Protection Directive, which, as said, recognises investment in data gathering and assembling.[39]

Art. 7 of the abovementioned directive was implemented by Art. 102-*bis* of the Italian Copyright Act, which was used by a 2011 Court of Appeal of Milan's decision to protect a database consisting of a book containing genealogical information.[40] The Court of Appeal held that such a book could be protected by database rights, provided that there had been substantial investment in obtaining the contents of the book. The *rationale* underlying the *sui generis* right is that, despite the lack of a creative effort, protection must be granted in those collections of "*a relevant number (which cannot be abstractly quantified in advance)*" of data, so long as "*the mere fact that such data was put together satisfies a socially-relevant «informative need.»*"[41]

On this basis, Big Data consisting of data relating to the activities performed by a user of a search engine or e-commerce shop may qualify as a

---

[38] M. BOGNI, A. DEFANT, *Big data e diritti IP: problemi di privacy, contrasto della contraffazione e protezione dei database e delle pagine dei social networks*, *Dir. ind.*, 2015, p. 119.

[39] M. REBEIRO, M. EVANS, *Big Data: protecting rights and extracting value*, in *Practical Law*¸ 2015, pp. 9-10, according to whom "*The database protection regime has limitations for generators of data. The European Court of Justice has construed the Database Directive literally, denying any database protection to the maker of a database whose investment primarily related to the creation of the underlying data (*British Horseracing Board Ltd v William Hill Organisation Ltd C-203/02*). However, that person should still be able to avail itself of the database protection if it can show that substantial investment has been made in verifying or presenting that original data. … In some jurisdictions (for example, the* ***UK****), the database and copyright protection that is available is subject to an exception that entitles researchers who have lawful access to protected material to carry out text and data analysis for non-commercial purposes, provided that sufficient acknowledgement is given where possible (*The Copyright and Rights in Performances (Research, Education, Libraries and Archives) Regulations *2014 (SI/2014/1372)).*"

[40] *App. Milano Sez. spec. propr. industr. ed intell*., 21 November 2011, no. 3206.

[41] G. GUGLIELMETTI, *La tutela delle banche dati con diritto sui generis nella Direttiva 96/9/CE*, in *Contr. e impr. Europa*, 1997, p. 181.

database in which an investment has been made. The dataright holder (*i.e.* the service provider) may therefore enforce its *sui generis* right against any unauthorised third party's extraction or re-utilisation of the entire database or a substantial part thereof. The dataright holder's right arises in the very moment where the data is uploaded in the database and "*indexed ... or encoded in such a way as to enable [the dataright holder] to retrieve or catalogue the data in accordance with a given criterion.*"[42] This means that data shall be stored in a fashion that permits the dataright holder to find the same data at a later stage, based on different criteria, and in accordance with the purposes to which data collection was destined.[43]

When compared to the European approach, the US would appear to be more restrictive, in that the US courts have progressively abandoned the above-referred «*sweat of the brow*» doctrine, in those cases where the database lacked the creativity step. The possibility for a Big Data database to enjoy copyright protection then lays with the dataright holder's ability to prove that the analysis performed by the software is creative by nature, an argument that, as said, would stand very low chances of success before the European courts.

Hence, while the database scheme may appear as a system specifically designed for Big Data, reality shows otherwise.

Database rights protect (sometimes at narrow conditions) the way in which data are organised or represented – and not data itself. So, in a typical Big Data situation, database rights "*would apply to the structured result of an algorithmic analysis of a dataset. Or they could apply to a relational database model, the way in which an application will sort data that is*

---

42 G. GUGLIELMETTI, *id.*, p. 180.

43 M. BOGNI, A. DEFANT, *id.*, p. 120. The Authors clarify that, for example, in cases where the data were collected in view of a subsequent profiling of the data subjects, the dataright holder shall be in a position to either find the data relating to a single, specific, user, or to a category of users (*e.g.* all the users who have purchased a particular product) to be able to send dedicated, tailored, advertisement.

delivered to it, before specific functionality is applied to it." The value of such databases directly derives from (i) access to the underlying data and (ii) the algorithmic process of selecting and manipulating the data, both of which are not covered by database rights. The end result of the exercise, "*as a snapshot, is covered by database rights. But the logic of importing, selecting and other functions on them, are not.*"[44]

3.4 Can copyright grant protection to Big Data beyond database?

The protection granted by database to Big Data is of limited scope. But, would copyright afford a greater level of protection to Big Data?

Copyright does not apply to the semantic content or meaning of text written by (human) authors: copyright covers the *way* the message is formulated, not the message underlying. If only one formulation is possible, then there is no copyright protection, because there is no creative choice possible.[45]

This means that data itself – *i.e.* the content – is not copyrightable.

A large part of Big Data is therefore not copyrightable in principle, *i.e.* data generated by machines or sensors, or statistical or mathematical data. Their nature as to digital information is susceptible to enjoy copyright protection, only if the dataright holder has re-elaborated data as a result of a creative effort, which is quite difficult to configure.

Another large portion of Big Data is, *in theory*, covered by copyright: user-generated content – such as blogs, pictures, videos, drawings, tweets posted online by the social network's user – does qualify as original or derivative work and may therefore enjoy copyright protection. However, *in practice*, copyright is never actually used on this kind of content: as discussed in chapter II, §2.1, in the vast majority of cases users grant a license over their

---

[44] J. DE WATCHER, *Big Data and IP Business Strategy*, in *Trading Secrets – Web portal TradeSecretsLaw.com*, 2014.

[45] S. ERCOLANI, *L'oggetto della protezione: l'opera dell'ingegno*, in C. GALLI, A.M. GAMBINO, *Codice commentato della Proprietà Industriale ed Intellettuale*, Torino, 2011, pp. 2837 *et seq.*.

user-generated data and content. The Internet service providers' terms of service include extremely liberal licenses, pursuant to which the service provider is vested with the widest rights of economic exploitation on the user-generated works.

In any event, even if copyright over user-generated content were enforceable, copyright would likely be of little use for the purposes of Big Data. As discussed above in this chapter, §2.1, the value of Big Data like user-generated content lies in disclosure. User-generated content, in order to have value, "*must be freely available to copy and paste, tag, adapt, create derivatives of, and, fundamentally, share without limitation*." It is the opposite of what copyright tries to achieve (*i.e.* a system of limited and controlled distribution and copying [46]). "*Most business value in using Big Data will be in open breach of copyright, typically by ignoring it or, at best, pay some lip service to it (as e.g. Facebook or other large social media do), or will be dealing with data that are not under copyright, but have not necessarily been recognized yet as such by the court system*."[47]

Against this background, one may argue that a distinction must be drawn between merely compiling raw data *vis-à-vis* mining Big Data, which is the result of working with raw data to extrapolate patterns or trends.[48]

---

[46] The owner of the copyright has the exclusive right to reproduce, distribute, publicly perform or display the work; to prepare derivative works; or to license others to engage in the same acts under specific terms and conditions.

[47] J. DE WATCHER, *ibid.*: "*As a result, the copyright aspect of any IP strategy in Big Data will first and foremost have to make the analysis of a) whether copyright applies, and b) whether it adds any business value. Since the applicability of copyright on machine or user generated content is partly in legal limbo, an appropriate solution for some businesses may be to use the creative commons approach. It helps to ensure that data are shared and re-used, hence increasing their value, and allows, from a practical perspective, to ignore the question whether or not copyright applies. If it applies, the creative commons license solves the problem. If it does not apply, and the data can be freely used, the end result will be, from a business perspective, similar.*"

[48] According to J. BICK, *Big Data rights protection found in Internet copyright law*, J. Bick, *Big Data rights protection found in Internet copyright law*, New Jersey Law Journal April 15, 2015, "*Here are some sample data sources and, thus, copyright creators: standard report users, ad hoc query users, remote partners/suppliers, power users, business executives, statisticians/data scientists, executives and the board of*

In the abovementioned *Feist Publications, Inc v. Rural Telephone Service Company, Inc.*, the United States Supreme Court concluded that names, cities and telephone numbers of subscribers – sorted in alphabetical order in local "*white pages*" telephone directories – were not copyrightable facts as the dataright holder had not compiled and arranged these data «*in an original way*.» Hence, even though the actual circumstances of the case led the Supreme Court to deny copyright protection, the court's line of reasoning may theoretically support the contrary argument that individual pieces of data, if arranged or coordinated in an original way (such as through an algorithm), may enjoy copyright protection.[49]

In 2012 the Canadian postal service, Canada Post, filed a copyright claim against Geolytica Inc.. The Canada Post's claim seems to follow the logic of *Feist Publications, Inc v. Rural Telephone Service Company, Inc.* to an extent.[50]

---

directors, and human resources. The user who is responsible for making the data tangible is considered the owner."

[49] J. TSOUKAS, N.J. VERMETTE, *op. cit.*, pp. 19-20.

[50] By way of background, the criteria to access copyright protection in **Canada** are stricter than in the US. According to J. TSOUKAS, N.J. VERMETTE, *op. cit.*, pp. 19-20, the Canada Copyright Act states defines a «*compilation*» as a work resulting from the selection or arrangement of literary, dramatic, musical or artistic works or of parts thereof, or a work resulting from the selection or arrangement of data. As a general rule, data compilations are subject to copyright protection in Canada if they are original works. Art. 5 of the Copyright Act states that, in Canada, copyright shall subsist "*in every original literary, dramatic, musical and artistic work.*" The Copyright Act does not define the meaning of term «*original*.» For a work to be considered original, within the meaning of the Copyright Act, the Supreme Court of Canada reasoned that it "*must be more than a mere copy of another work. At the same time, it need not be creative, in the sense of being novel or unique. What is required to attract copyright protection in the expression of an idea is an exercise of skill and judgment.*" (*CCH Canadian Ltd. v. Law Society of Upper Canada*, 2004 SCC 13) By skill, the Supreme Court meant the "*use of one's knowledge, developed aptitude or practised ability in producing the work.*" By judgment, it meant the "*use of one's capacity for discernment or ability to form an opinion or evaluation by comparing different possible options in producing the work.*" Finally, the exercise of skill and judgment required to produce the work must not be a purely mechanical exercise. As is the case in Canada, to qualify for copyright protection in the United States, a work must be original to the author. While the laws are similar in the United States and Canada, the originality standard set out by the Supreme Court of Canada differs from that established by the United States Supreme Court. In *Feist Publications, Inc. v. Rural Telephone Service Co., Inc.*, the US Supreme Court

Geolytica operates the website GeoCoder.ca, a website that provides several geo-coding services including free access to a crowd-sourced compiled database of Canadian postal codes.[51] Canada Post claimed to be the exclusive copyright holder of all Canadian postal codes and accused Geolytica of "*appropriating*" the Canada Post database and "*making unauthorised reproductions*." Geolytica argued that (i) its database does not rely on the Canada Post's database; (ii) an address, including the postal code, is a fact and not an original work within the meaning of the Canada Copyright Act, so that Canada Post Corporation by no means owns copyright in addresses, or in postal codes that form a part of addresses; and, in any event, (iii) the selection and arrangement of data into the Canada Post's database is commonplace.[52]

The parties ultimately entered into a settlement agreement – the terms of which are confidential – whereby, according to the joint public statement, Canada Post discontinued the claim and Geocoder continued to make its database available to the public.[53] In any event, *Canada Post v. Geolytica*

---

maintained that the term «*original*,» as used in copyright, means only that the work was "*independently created by the author (as opposed to copied from other works), and that it possesses at least some minimal degree of creativity*." The Canadian threshold for copyright protection of an original work appears to be more demanding than the standard required in the United States. The exercise of skill and judgement being arguably a higher standard than a work being independently created and having a minimal degree of creativity.

[51] Geolytica describes itself as follows: "*Geocoder.ca is a public web service providing both free and commercial geocoding services for North America: Canada and the USA. Geocoding is the process of computing the latitude and longitude of a location. ... When you make a query to geocoder containing for example this information "1435 Prince of Wales, Ottawa, ON K2C 1N5," we then extract the postal code "K2C 1N5" and insert it into the database that you may download for free on this website*" (www.geocoder.ca).

[52] As Geolytica put it in its statement of defence, "*Although there may have been an exertion of labour to establish a postal code designation system, Canada Post Corporation did not, and does not, exert a non-trivial amount of skill and judgment to create and maintain the CPC Database.*"

[53] See the Canada Post and Geolytica joint press release: "*Canada Post commenced court proceedings in 2012 against Geolytica Inc. for copyright infringement in relation to Geolytica Inc.'s Canadian Postal Code Geocoded Dataset and related services offered on its website at geocoder.ca. The parties have now settled their dispute and Canada Post will discontinue the court proceedings. The postal codes returned by various*

suggests that, although copyright may apply to data at very narrow conditions, businesses may still not refrain from claiming copyright thereon. Claiming copyright is relatively easy: there is no registration system, and there is no sanction for wrongfully claiming copyright or claiming copyright on something that cannot be covered by copyright (*e.g.* machine-generated data).[54]

4. Software

Having established that a compilation of Big Data may, at certain conditions, be subject to database protection, the following paragraphs discuss whether the actual process of analysing data is suitable to enjoy patent protection. Data analytics extensively rely on software and algorithms, which transform raw data into valuable pieces of information, so that a great deal of discussion has arisen, in particular in the US, as to the possibility to claim patent protection on these software and algorithms.

4.1 Software protection within the European Union: patentability requirements for computer-generated inventions

Source codes and object codes themselves, rather than their functionalities, in software can be protected by copyright law.[55] As in the European Union regulations software is equalled to literary works, copyright

---

geocoder interface APIs and downloadable on geocoder.ca, are estimated via a crowdsourcing process. They are not licensed by geocoder.ca from Canada Post, the entity responsible for assigning postal codes to street addresses. Geolytica continues to offer its products and services, using the postal code data it has collected via a crowdsourcing process which it created" (http://geocoder.ca/?sued=1).

54 J. DE WATCHER, *op cit.*

55 For a comprehensive analysis of the pros and cons that the copyright and patent protection can afford to software, see L. SCHIUMA, *Il Software tra brevetto e diritto d'autore*, Relazione al Convegno "La tutela del software tra brevetto e diritto d'autore", Facoltà di Giurisprudenza dell'Università Lumsa di Roma, Roma, 24 giugno 2004, in *Riv. dir. civ.*, 2007, pp. 683-707.

protection is conditional on the originality criterion being met.[56] In this respect the same considerations made in §3 apply.

In the European Union, however, the *functional features* within a computer program may be eligible for patent protection if, pursuant to Art. 52 of the European Patent Convention, those functionalities are "*new, inventive and make a technical contribution.*"[57]

Following the EPO Board of Appeal's decision in *Computer program product/IBM* (T 1173/97) – which affirmed that if "*technical features in addition to the computer program were claimed,*" patent protection may apply – Courts have in principle resolved for the patentability of computer-implemented invention. Patentability is conditional on a multiple-step legal test,[58] which shall ultimately establish if the contribution offered by the

---

[56] As known, the assimilation of software to literary works is acknowledged by directive 91/250/EEC of 14 May 1991 on the legal protection of computer programs, as the outcome of a joint reading of Art. 2 of the 1886 Berne Convention for the protection of literary and artistic works; Art. 10, para. 1, of the 1994 WTO Agreement on Trade-Related Aspects of Intellectual Property Rights (TRIPS Agreement); Art. 4 of the 1996 WIPO Copright Treaty. Art. 2, para. 1, no. 8 implemented the directive as follows: "*[protection shall extend to ...] computer programs, in whatever form they are expressed, provided that they are original and result from the author's own intellectual creation. Ideas and principles which underlie any element of a computer program, including those which underlie its interfaces, shall be excluded from the protection afforded by this Law. The term "computer program" shall include their preparatory design materials.*"

[57] Art. 52 of the European Patent Convention provides that "*(1) European patents shall be granted for any inventions, in all fields of technology, provided that they are new, involve an inventive step and are susceptible of industrial application. ... (2) The following in particular shall not be regarded as inventions within the meaning of paragraph 1: ... (c) ... programs for computers. (3) Paragraph 2 shall exclude the patentability of the subject-matter or activities referred to therein only to the extent to which a European patent application or European patent relates to such subject-matter or activities as such.*" Remarkably, in 2005 the European Parliament rejected the draft directive on computer implemented inventions was dismissed by European Parliament in 2005, because Member States could not agree on its impact on software use and protection; see Practical Law's 25 July 20005 press release (available at www.practicallaw.com/7-201-0664).

[58] The vast majority of case law on the patentability of computer-implemented inventions lays at an European Patent Office's (EPO) level. As summarised in H. PEARSON, A. DUFFUS, P. WARD, R. CLABAUGH, F. BOURGUET, *Software and business methods: should you be patenting them?*, in *Practical Law*¸ 2017 (available at www.practicallaw.com/4-200-9964), the current EPO's position on software-related

invention is «*technical*» (although a definition for the term is not provided) and if the functional features of the software make the computer «*better*,» *i.e.* the software enables the computer to run more efficiently and effectively.[59]

---

inventions can be summarised as follows: "*While «programs for computers» and «methods for doing business» are among the items excluded from patentability under Article 52(2) of the European Patent Convention, if the claimed subject-matter has a technical character it may be patentable if it meets the other requirements for patentability … Although the execution of a computer program always involves physical effects (for example, electrical currents), such effects are not sufficient to give the program technical character. A program will be patentable if it produces a further technical effect going beyond the normal physical interactions between the program and the computer. Claims to an invention which is realised by means of a computer program may take the form of a method of operating apparatus, the apparatus set up to execute the method, or the program itself. When examining an invention for patentability, the EPO Guidelines for Examination (the detailed and technical guide used by patent examiners in the EPO) state that the EPO will not draw any distinctions based on the overall purpose of the invention, that is, whether it is intended to fill a business niche, provide new entertainment and so on.*" On the other hand, as clarified in L.BOSOTTI, O. CAPASSO, *Oggetto del brevetto ed esclusioni della brevettabilità*, in C. GALLI, A.M. GAMBINO, *Codice commentato della Proprietà Industriale ed Intellettuale cit.*, pp. 557-558, national courts' case law precedents on the patentability of computer-implemented inventions are relatively few, because, as said, national courts deal with the matter only at a stage (*i.e.* litigation on national portions of EPO patent applications), where the EPO has already assessed the patentability issue. Against this background, an arbitration panel (award dated 13 December 2004) concluded that "*an invention concerning the method to lease real estates for holidays implemented through electronic elaborators and the Internet Web is patentable* ."

[59] In *Aerotel Ltd v. Telco Holdings Ltd* [2006] EWCA Civ 1371, whereby the Court of Appeal of **England and Wales** assessed the patentability of a telephone system whereby "*a user prepays, obtains a code, and both the code and prepayment amount are stored in a memory in a special exchange. The user then calls the special exchange, inputs their code and is connected to the desired telephone number. During the call the exchange monitors the user's credit and disconnects the call when the credit reaches zero.*" The Court of Appeal set forth a four-step process to test the patentability of a computer-implemented invention, as follows: "*1. Properly construe what it is that the patent claims as the invention. 2. Identify the actual contribution: what has the inventor really added to human knowledge? 3. Ask whether the contribution falls solely within the subject matter excluded from protection by Article 52. 4. Check whether the actual or alleged contribution is actually technical in nature*." The Court of Appeal held that, although the system could be implemented using conventional computers, the contribution was a "*new physical combination of hardware*." It was therefore "*more than just a computer loaded with a program and it was also more than conducting a method of doing business*." As a result, it did not fall solely within the subject matter excluded by Article 52, and passed step 3 of the four-step approach. It was also "*clearly technical in nature and so passed step 4.*" It was patentable subject matter. See P. ENGLAND, H. SMITH, P*atenting inventions: the Court of Appeal's four-step programme*, in *Practical Law*, 2006 (available at www.practicallaw.com/5-206-3960), which contains also reference to the Macrossan's patent application for a computer-automated method of acquiring documents for incorporating a company: in this case the same

These criteria represent common ground for the analysis of the same patentability issue in the US. As discussed in §4.2 below, the patentability of software in the US relies on a similar principle that software constitutes an improvement of the functioning of a computer, or an improvement of any other technology or technical field.

4.2 Software protection within the U.S.A.: *Alice Corp v. CLS Bank Int'l*

Pursuant to Section 101 of the US Patent Act, "*Whoever invents or discovers any new and useful process, machine, manufacture, or composition of matter, or any new and useful improvement thereof, may obtain a patent therefor … .*" It is an established principle of US patent law that laws of nature, natural phenomena, and abstract ideas are not patentable, as clarified by the Supreme Court in *Association for Molecular Pathology v. Myriad Genetics, Inc.*.[60]

Although the US patent law does not expressly mention software among the patentable inventions, since *Gottschalk v. Benson* in 1972 the Supreme Court has not denied that software is in principle patentable, but has in fact refused to grant protection to software-related inventions, on grounds that the subject matters of those inventions, *i.e.* the underlying mathematical algorithm, was not patentable.[61]

Court of Appeal of England and Wales concluded that the contribution under step 2 was merely an up-and-running computer program and so was excluded subject matter. For an overview of the **German** and **French** precedents on the patentability of computer-implemented inventions, see H. PEARSON, A. DUFFUS, P. WARD, R. CLABAUGH, F. BOURGUET, *op. cit.*.

[60] *Association for Molecular Pathology v. Myriad Genetics* [1] No. 12-398 (569 U.S. ___ June 13, 2013)

[61] Based on the principle that an idea is not patentable, the Supreme Court found the ineligibility of patent claims involving (i) an algorithm for converting binary-coded decimal numerals into pure binary form (*Gottschalk v. Benson* 409 U.S. 63 (1972)); (ii) a mathematical formula for computing alarm limits in a catalytic conversion process (*Parker v. Flook* 437 U.S. 584 (1978)); and (iii) a method for hedging against the financial risk of price fluctuations (*Bilski v. Kappos* 561 U.S. 593 (2010)).

The 2014 landmark decision on *Alice Corp v. CLS Bank Int'l* confirmed the Supreme Court's trend, but offered new, interesting arguments to develop a patentability case in future.[62]

By way of background, Alice Corporation sought protection for a computer-implemented process aimed at minimising settlement risk, *i.e.* the risk that only one party to a trade of financial instruments will perform its obligation. The patent claims were designed to facilitate the exchange of financial obligations between the parties by means of a third party computer intermediary system.[63]

Consistently with the patentability test used in *Mayo Collaborative Services v. Prometheus Laboratories, Inc.*,[64] the Supreme Court held that the patentability analysis for software patents must consider a two-part test, assessing (a) whether the claims only encompass abstract ideas and (b) if encompassing an abstract idea, whether the claims include some additional inventive step showing an application of the abstract idea. According to the Supreme Court's *en banc* judgment, Alice Corporation's software did not satisfy the *Mayo*'s two-step patentability test, for the following reasons:

(a) *abstract idea:* the concept of intermediated settlement is – like the hedging system at the core of *Gottschalk v. Benson* – "*a fundamental economic practice long prevalent in [the] system of commerce ... and a use of a third-party intermediary (or «clearing house») is building block of the modern economy. Thus, intermediated settlement, like hedging, is an abstract idea*"; and

(b) *inventive step:* the method claims of Alice Corporation's patent fail to transform the abstract idea into a patent-eligible invention. The

---

62 *Alice Corp. v. CLS Bank International* 573 U.S. __, 134 S. Ct. 2347 (2014)

63 In particular, Alice Corportation's US patent no. 5,970,479 claimed "*(1) a method for exchanging financial obligations, (2) a computer system configured to carry out the method for exchanging obligations, and (3) a computer-readable medium containing program code for performing the method of exchanging obligations.*"

64 *Mayo Collaborative Services v. Prometheus Laboratories, Inc.* 132 S. Ct. 1289 (2012).

introduction of a computer into the claims does not qualify as an inventive concept. Alice Corporation's patent "*does no more than simply instruct the practitioner to implement the abstract idea of intermediated settlement on a generic computer.*" The function performed by the computer at each step – *i.e.* creating "*shadow*" accounts, obtaining data, adjusting account balances, and issuing automated instructions – is conventional. Alice Corporation's invention does not "*purport to improve the functioning of the computer itself or effect an improvement in any other technology or technical field. An instruction to apply the abstract idea of intermediated settlement using some unspecified, generic computer is not enough to transform the abstract idea into a patent-eligible invention.*"

The *Alice* decision resulted in an increased scrutiny on software: both Federal Circuit and district court's most recent decisions have applied the two-part *Alice* test to deny the patentability of inventions aimed at manipulating information using a computer (*e.g.* by collecting user gaming data and then generating statistical information), [65] on grounds that manipulating and processing data is an abstract idea and the very same tasks solved by the software might have been performed by a human hand, so that the computerisation of a process does not hit the inventive threshold.[66]

---

[65] See *GC Technology Development, LLC v. Big Fish Games, Inc.* 2:16-cv-00857-RCJ-VCF (D Nev 29 August 2016)

[66] J. TSOUKAS, N.J. VERMETTE, *op cit.*, pp. 21-22 explain that similar principles apply in **Canada**. In its 2011 decision, the Canada Federal Court of Appeal (*(Attorney General) v. Amazon.com, Inc.* 2011 FCA 328) supported the position that, if, upon conducting a «*purposive construction*» of the claim elements, the only novel aspect of a claimed invention is an "*abstract idea such as an algorithm*," then the invention is not patentable. On the other hand, "*if the algorithm forms part of a novel combination of essential elements then that invention would be eligible for patent protection.*"

4.3 Patenting Big Data analytical software

The recent developments in patent law suggest that protecting Big Data analytical software with patent should be at least reconsidered. While software appears to remain patentable in principle (the issue was not openly addressed in *Alice*), a number of concerns arise.

Firstly, merely assembling, organising or manipulating data is not itself eligible for patenting.[67]

Secondly, algorithms remain very difficult to protect by way of patents, because "*they generally involve a series of calculation steps and do nothing «technical» to a computer*," which may result in courts denying protection on grounds that "*an algorithm does not normally make a computer faster, more reliable or have a higher resolution.*"[68] Hence, as long as Big Data analytics focuses on source codes and algorithmic manipulation of datasets, patent protection may be difficult to pursue: even if a patent is issued, infringers may routinely raise defences that the subject matter of the patent was not eligible for a patent grant.[69] These issues may be mitigated, in general terms, by employing unequivocal claims language, but full disclosure may in turn trigger a risk that *too much* information is shared with competitors – a risk for which a reward may ultimately lack, should the algorithm fail to enjoy patent protection, and that may ultimately disincentive major patent investment in the field of Big Data analytics.

---

[67] T. VARE, M. MATTIOLI, *Big business, big government and big legal questions*, in *Managing IP*, October 2014, p. 47

[68] M. REBEIRO, M. EVANS, *Big Data: protecting rights and extracting value*, in *Practical Law*¸ 2015, p. 9. R. MERGES, *Symposium: Go ask Alice – what can you patent after Alice v. CLS Bank*, 20 June 2014 post on the SCOTUS blog (available at http://www.scotusblog.com/2014/06/symposium-go-ask-alice-what-can-you-patent-after-alice-v-cls-bank/) criticises the *Alice* decision: by recalling the principle expressed in *Benson* that algorithms are species of abstract ideas, the court "*misrepresentes the nature of algorithms (which simply do not grow on threes),*" meaning "*an entire shelf full of discredited cases on the metaphysics of what is and is not algorithm must now be dusted off.*"

[69] D.A. PRANGE, *Navigating the protection of big data cit.*, p. 55.

What appears to be crucial going forward is, however, a deeper analysis of the role played by software towards the solution of a technical issue. From this angle, after *Alice* some classes of software look highly unlikely for patent protection at their outset, to the extent that they merely apply a computer process to a pre-existing business principle (*e.g.* financial settlement risk, hedging).[70] On the other hand, the *Alice* decision may support arguments that "*complex software and/or hardware solutions to analyse, manipulate or store big data may be less vulnerable to the kind of attack that cost* Alice *its patent. And at the very least, courts will be confronted again with the question of what is and is not an algorithm.*"[71]

The *Diamond v. Diehr* case cited in the *Alice* decision may provide useful guidance. According to the Supreme Court in *Alice*, the invention in *Diehr* was patentable, because it "*transformed the [pre-existing] process into an inventive application of the formula*."[72] In so deciding, the Supreme Court appears to set the patentability threshold in accordance with the capability for the software analytics of «*doing something*,» *i.e.* "*performing action and demonstrating function beyond merely informing.*"[73] On this basis, when Big Data solutions act to improve existing technological processes and solve current technological problems – such as reformatting data from disparate

---

[70] This may be the reason why Goldman Sachs & Co. used to maintain their high-speed trading algorithms secret. The steps used by Goldman Sachs to keep the algorithm under confidentiality are described in *United States v. Aleynikov*, 10 Crim. 96 (S.D.N.Y. Feb 10, 2010), 2010 WL 4000356.

[71] A. GOTHING, A. MUÑOZ-KAPHING, *Big data and Alice v. CLS: predicting what's next*, in *Intellectual Property Magazine*, 7/2014, pp. 21-22.

[72] In *Diamond v. Diehr* 450 U.S. 175 (1981) the Supreme Court confirmed the patentability of a computer-implemented process for curing rubber. The invention relied on the exploitation of a well-known formula by a device that recorded constant temperature measurement inside the rubber mould to provide real time calculations of remaining cure time.

[73] E. MICHIKO MORRIS, *Alice, Artifice and Action - and Ultramercial*, 8 July 2014 post on the Patently-O blog, available at http://fstp-expert-system.typepad.com/files/94-e.-morris_alice-artifice-and-action-and-ultramercial_iu-i.n.-08.07.2014.pdf; E. MICHIKO MORRIS, *What is Technology?, V 20 Boston University Journal of Science and Technology Law*, 2014.

sources, creating new datasets for easier storage, or reconfiguring data into different display sets – it is the extent of the solution that is affected that will have the greatest impact on patentability.[74]

Without prejudice to the arguments above, a material degree of uncertainty is obviously going to lay with the patentability of an algorithm for a long time. Keeping the algorithm secret, and relying on the law of confidential information, may therefore be a safer harbour, at least in the near future.

5. Trade secret

5.1 Recent developments in trade secret regulation in the European Union and in the U.S.A.

Until very recent years trade secret was afforded protection in the European Union and in the US mostly through un-harmonised local legislations,[75] resulting in a "*fragmentation of the internal market*" and "*weakening of the overall deterrent effect of the relevant rules.*"[76]

---

[74] A. GOTHING, A. MUÑOZ-KAPHING, *op. cit.*, p. 22.

[75] Before the adoption of the Trade Secret Directive (as defined below), a uniform definition for «*trade secret*» lacked across the territory of the **EU**, some Member States not even having a specific trade secret national legislation: the **Netherlands** relied on tort law, **France** and **Germany** relied on contract (*i.e.* there is no right to prevent misappropriation/misuse of confidential information, other than via contractual obligations, under non-disclosure agreements; Germany has a set of criminal law provisions prohibiting disclosure of confidential information, but, again, no parallel civil rights of action, unless a non-disclosure agreement is breached), whilst the **UK** and **Ireland** relied on common law (in particular, equitable action for breach of confidence does not protect the information *per se*, but rather the unconscionable use or threatened use of such information when it is imparted in circumstances of confidence, as clarified in *Coco v. AN Clark (Engineers) Ltd* [1968] F.S.R. 415). In **Italy**, company information and technical/industrial know-how, including commercial information, are protected pursuant to Art. 98 of legislative decree of 10 February 2005 no. 30 (the "Italian IP Code") upon the concurrent conditions that information/know-how (i) is «*secret*» (meaning not generally known or accessible to the public); (ii) must derive economic value from being kept confidential; and (iii) is subject to security/confidentiality measures which are reasonably adequate to keep information/know-how secret. Overseas, before the entry into force of the DTSA (as defined below) a number of **US** States had adopted federal laws in the form of the Economic Espionage Act of 1996.

[76] See Recital 8 of the Trade Secret Directive (as defined below).

At the end of a lengthy and complex process, in 2016 both the European and US lawmakers adopted harmonised regulations for trade secret, as follows:

(a) the EU Member States must implement directive 2016/943/EU (the "Trade Secret Directive") of 8 June 2016 by the end of 2018;[77] and

(b) former US President Obama signed the Defend Trade Secrets Act ("DTSA") on 11 May 2016, and the law took effect immediately.[78]

Both the Trade Secret Directive and the DTSA are aimed at protecting commercial confidential information.

Art. 2 of the Trade Secret Directive defines a trade secret as information which meets all of the following requirements:

(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. [79]

The DTSA defines «*trade secret*» as "*all forms and types of financial, business, scientific, technical, economic, or engineering information, including patterns, plans, compilations, program devices, formulas, designs, prototypes, methods, techniques, processes, procedures, programs, or codes, whether tangible or intangible, and whether or how stored, compiled, or*

[77] Directive 2016/943/EU of 8 June 2016 on the protection of undisclosed know-how and business information (trade secrets) against their unlawful acquisition, use and disclosure.

[78] S.1890 - Defend Trade Secrets Act of 2016.

[79] This definition is a consistent with the definition for «*trade secret*» pursuant to Art. 39(2) of the TRIPs Agreement, pursuant to which the signatories undertake to grant some level of protection for confidential information.

memorialized physically, electronically, graphically, photographically, or in writing if-

1. *the owner thereof has taken reasonable measures to keep such information secret; and*
2. *the information derives independent economic value, actual or potential, from not being generally known to, and not being readily ascertainable through proper means by, the another person who can obtain economic value from the disclosure or use of the information.*"

The above definitions provide a substantially common ground for the protection of trade secret in the European Union and in the US: to be considered a trade secret the information must be kept confidential and derive an economic value from the fact that it is confidential. What is not required is that the information be entirely novel, a distinction from other forms of intellectual property. Further, the Trade Secret Directive makes explicit that combinations of otherwise publicly available information can be protected provided they are not readily accessible or generally known.[80]

In the US, the limitation period for the rightholder to bring an action under the DTSA is three years after the date on which the misappropriation is discovered (or could have been discovered with reasonable diligence). The Trade Secret Directive gives the Member States the freedom to set the limitation period for their respective national laws, but sets the maximum

[80] In addition, both the Trade Secret Directive and the DTSA set out provisions aimed at ensuring confidentiality during legal proceedings for trade secret misappropriation. The Trade Secret Directive requires an applicant to file a "*duly reasoned*" application clarifying why information must be kept confidential; similarly, the DTSA states that a court may not authorise disclosure of information, unless the information owner is first given the opportunity to file a submission describing the interest in keeping the information confidential. For a more in-depth comparison between the Trade Secret Directive and the DTSA see D. CROUCH, A Comparison of the EU Trade Secrets Directive and the US Defend Trade Secrets Act, 16 May 2016 post on the Patently-O blog, available at https://patentlyo.com/patent/2016/05/comparison-secrets-directive.html; J. EXTEN-WRIGHT, *The Trade Secrets Directive: New protection for businesses?*, 18 July 2016, available at https://www.dlapiper.com/en/abudhabi/-insights/publications/2016/07/the-trade-secrets-directive/.

period at six years (but Member States have discretion as to when the period starts running).

5.2 Maintaining secrecy over Big Data

The recent regulatory changes to the trade secret discipline, together with the issues arising in connection with patenting Big Data software analytics, suggest that trade secret may become more attractive for protecting a company's Big Data.

Big Data relies on a wide variety of data compilations, processes, *formulae*, including algorithms, data structures or methods for analysing or delivering tailor-made content. All these data compilations and processes are suitable to find protection under trade secret law,[81] to the extent that they have «*commercial value*» for the purposes of the Trade Secret Directive and the DTSA. As discussed above, the commercial value of Big Data lays with the fact that "*acquisition, use or disclosure is likely to harm the interests of the person lawfully controlling it, in that it undermines his or her scientific and*

[81] A. GOTHING, A. MUÑOZ-KAPHING, *op. cit.*, p. 22. **Italian** case law has clarified that trade secret covers both «*industrial*» and «*commercial*» information. Industrial information means all materials and know-how which, as a whole, identify the necessary steps to run a process (see *C. App. Milano*, 13 June 2007); commercial information includes, among others, lists of clients (but not a mere mailing list), marketing or discount strategies, advertising or promotional materials, market analysis reports, business methods. For an overview of the Italian case law see C. PASCHI, *L'oggetto della tutela delle informazioni riservate*, in C. GALLI, A.M. GAMBINO, *Codice commentato della Proprietà Industriale ed Intellettuale cit.*, pp. 900-902. As regards **US** trade secret law, see V. CHIAPPETTA, *Myth, Chameleon or Intellectual Property Olympian? A Normative Framework Supporting Trade Secret Law*, 8 GEO. MASON L. REV. 69, 76 (1999): "*Trade secret law ... extends to technical and nontechnical information, expression, ideas and facts, embracing such things as customer and supplier lists, financial information, methods of doing business, future marketing, sales and product plans and even employee names, job responsibilities and phone numbers.*" A well-known example of Big Data-related process that is currently being kept confidential as trade secret is Google *PageRank*, *i.e.* the algorithm Google Search uses to rank websites in their search engine results. As explained PageRank operates "*by counting the number and quality of links to a page to determine a rough estimate of how important the website is. The underlying assumption is that more important websites are likely to receive more links from other websites*" (see https://web.archive.org/web/20111104131332/https://www.google.com/competition/howgooglesearchworks.html).

technical potential, business or financial interests, strategic positions or ability to compete."[82]

However, issues may arise when assessing the «*confidentiality*» requirement. This is the case, for example, of data sourced by publicly available sources (such as «*open data*»), or of data being licensed to third parties. In both cases – which happen quite frequently (see above in this chapter, §2.1) – extrapolation of raw data extends beyond internal sources or uses, so that it may be difficult for the dataright holder to demonstrate that the information is kept confidential.[83]

Overall, it may be that one of the «*V's*» usually associated with Big Data, *volume*, is by nature hard to reconcile with the concept of confidentiality:[84] Big Data relies on massive, constant flows of data; data scientists are constantly immersed in data, which they analyse, compile and mine as a job; Big Data software and data assets can also be easily moved from computer to computer.

Big volumes and easy transportability inevitably raise the bar for proving that the dataright holder has undertaken reasonable steps to keep data secret. This may occur by combining physical measures with strict non-disclosure contractual provisions for employees, as follows:

(a) *Physical measures:* standard precautionary measures – such as network password protection, access limitations, and computer port limitations (for example, preventing flash drive use) – may not be

---

[82] See Recital 14 of the Trade Secret Directive.

[83] M. REBEIRO, M. EVANS, *op. cit.*, p. 10 considers that "*Whether any information contained in the data has the quality of confidence necessary for protection under the law of confidence will depend on the facts surrounding the case, such as an unusually high price paid for the data, complexity of technical measures put in place to control access to the data, or how access to the data is monitored.*"

[84] J. DE WATCHER, *op. cit.*, considers that "*secrecy has a major downside: it means you can't talk about, use or disclose whatever is secret in a way that allows others to find out about it. The challenge here is that a lot of the value of Big Data depends, as we have seen repeatedly, on the ability to have access, and preferably open or free access, to as much data as possible.*"

deemed to be sufficient for protecting Big Data. Other physical protections may be necessary, including "*limited networks and avoiding cloud storage of information (unless the cloud itself is proprietarily protected),*" and "*actively cataloguing the types of information protected as a trade secret in order to evaluate the existence of redundant protective measures*;"[85] and

(b) *Employee-related restrictions:* a relevant portion of the data mining processes are performed by a new category of workers, for which commentators forecast there will be a massive market demand in the near future: data scientists.[86] Placing restrictions on data scientists' employment opportunities may potentially conflict with the principle of free movement of workers endorsed in both the Trade Secret Directive[87] and the DTSA,[88] and may be therefore difficult to enforce. Hence, while companies with Big Data assets may not be able to simply prohibit employees from moving to competitors, companies

---

85 D.A. PRANGE, *Big data and trade secrets: part 2*, in *Intellectual Property Magazine*, 2/2017, pp. 41-42 comments that cataloguing, "*if documented, may be used at trial as an exhibit to help explain to the jury the different processes that are used to protect a company's trade secret information. But doing such a study also has risks if the conclusions are not followed; a party challenging a trade secret may point to the lack of diligence as proof that reasonable measures were not taken to protect asserted trade secret information*."

86 D.A. PRANGE, *Big data and trade secrets: part 2*, p. 41.

87 See Recital 13 of the Trade Secret Directive.

88 D.A. PRANGE, *ibid.* highlights that, in the pursuit of a balance between confidentiality and free movement of workers, "*some US states recognise an «inevitable disclosure» doctrine* [see, for example, *WL Gore & Assoc, Inc v. Wu* 2006 WL 2692584 (Del Ch 2006)] *providing that an employee with knowledge of a trade secret can be prevented from working for a competitor on the theory that the individual's work at the competitor will result* in the «inevitable disclosure» of the trade secrets learned at the former employer. Other states have not endorsed this doctrine [see, for example, *Whyte v. Schlage Lock Co* 125 Cal Rptr 2d 277 (4th Dist 2002)], *or taken a middle position in which application of the doctrine depends on the type of individual or if there is a separate employment agreement. The DTSA takes an intermediate position in allowing for an injunction award against an individual. To enjoin an individual based on statutory trade secret misappropriation, there must be actual or threatened misappropriation as opposed to the individual possessing general knowledge. It can be more difficult to prevent a data scientist from leaving for a competitor if there is no specific misappropriation identified.*"

may still effectively limit former employee activities related to the company Big Data trade secret assets to which the employee had access. Companies should strive to prevent data dissemination by accurately drafting employment agreements, whereby (i) the trade secrets are identified as precisely as possible, (ii) the employee acknowledges ownership and secrecy of data, and acknowledges the fact that the employer has taken reasonable efforts to keep the information secret, (iii) the employee undertakes to return all company property, including information, upon termination of employment relationship, and (iv) the employee undertakes to keep information secret also after the termination of the employment contract.

6. Remedies against infringement

As repeated frequently in this paper, the output of data analytics is data again. Output data is of course more useful than raw input data, because output data is the result of analysis that entails personal judgment, which is is normally applied by data scientist.

Notably, personal judgment is often key to understanding how valuable output data is. So, in a scenario where an unauthorised third party were to access Big Data, the first remedy that may be available to dataright holder would be of a very practical and technical nature. If the infringer does not know which criteria and personal judgments were applied in the data mining phase, the infringer may not be able to exploit data at all.

Turning to the legal remedies, the dataright holder may seek those remedies that intellectual law offers in response to a breach of the exclusivity granted by the rights mentioned in §§3, 4, 5 above (*i.e.* copyright, patents and trade secret). The IP Enforcement Directive 2004/48/EC[89] empowers the right holder to seek the following reliefs against infringement: injunctions (both

[89] Directive 2004/48/EC of the European Parliament and of the Council of 29 April 2004 on the enforcement of intellectual property rights.

interim and final), recall or destruction of infringing materials or materials used to create infringing copies, damages and payment of legal costs.

Although the application of traditional intellectual property law remedies to Big Data is still untested before the judicial courts, commentators tend to delimit injunctions against unauthorised uses of Big Data to a limited scope, in the sense that the dataright holder may not be able to prevent reproductions or distributions of the subject matter of the data analysis subsequently to the unauthorised access.[90] The *rationale* underlying this consideration is that data, as content and ideas, is not subject to exclusivity, hence the public interest in having widespread access to data shall ultimately prevail.[91]

As concerns damages, the relief-from-royalty criterion elaborated by the national courts may provide useful guidance in cases where the breach concerned datasets that were destined to disclosure anyway (*e.g.* in the context of a license grant). Conversely, if data was confidential, alternative valuation methods may apply on a case-by-case basis, *e.g.* (i) cost-related methods, which rely on the costs incurred by the dataright holder to collect and analyse Big Data; (ii) earnings-based methods, such as excess earnings or the abovementioned relief-from-royalty, aimed at quantifying the added business value that is ascribable to Big Data analytics, or (iii) other empirical

---

[90] M. MATTIOLI, *Disclosing Big Data cit.*, p. 579:"*we might wish for dataright holders to be entitled to sue unauthorised users of their data for injunctive relief for some limited period of time. An exemplary «use» of data would be, for instance, applying a dataset to analysis in order to study a new problem or phenomenon. We might wish for dataright holders to not be entitled to prevent third parties from reproducing or distributing descriptions of the subject matter itself. Thus, underlying data could be freely reproduced and distributed barring any additional restrictions imposed by publishers through, for instance, contracts.*"

[91] J. TSOUKAS, N.J. VERMETTE, *op. cit.*, p. 23 note that the US "Data.gov" and the EU "open data" portals are two major examples that governments are increasingly taking open and public data into consideration as these data may be beneficial and necessary to the public good.

methods that offer credible measurements of the return on investment on Big Data.[92]

[92] For an overview of the valuation methods that may be employed to assess the value of data provided by IoT see R. MORO DI VISCONTI, *op cit.*, p. 542-544. For an analysis of the different provisions governing compensation for damages see C. GALLI, *Diritti IP e risarcimento dei danni: un rapporto difficile che sta cambiando*, in *Dir. ind.*, 2012, pp. 105-106; C. GALLI, *Risarcimento del danno e retroversione degli utili: le diverse voci di danno*, in *Dir. ind.*, 2012, pp. 109-120; R. ROMANO, *Diritto d'autore, tecniche di tutela e risarcimento del danno*, in *Dir. aut.*, 2003, pp. 406-415; S. CORONA, *Le misure risarcitorie e indennitarie*, in C. GALLI, A.M. GAMBINO, *Codice commentato della Proprietà Industriale ed Intellettuale cit.*, pp. 1098-1125.

CONCLUSION

The concept of Big Data does not easily lend itself to formal intellectual property protection. Digital technology is evolving at an unprecedented rate and the legal framework for intellectual property protection has yet to fully embrace to the Big Data era.

Commentators underscore the importance of a business-oriented cultural shift towards the definition of a more tailor-made legal framework for Big Data: while intellectual property law must undoubtedly continue to pursue the protection of investment and innovation, the traditional approach to business secrecy may need to be reconsidered. The need to keep information confidential arises when the information is scarcely available on the market. Big Data is not: digital information abounds on the Internet, as users have proven to be extremely willing to share it. With this in mind, companies need to consider that IP strategy is not only about protecting or restricting access to intellectual capital, but also about positive use of that intellectual property, meaning that, at certain conditions, disclosure may generate more business opportunities than confidentiality.[93]

Data sharing, however, needs substantial regulation. And contract law has so far emerged to be the most effective legal tool to govern data sharing, both in the *upstream* and *downstream* phases of the data flows, as follows.

A. Upstream phase (input data)

As discussed in chapter II, the provision of input data from the Web users to the service providers follows the contractual rules set forth in the terms of service that are unilaterally drafted by the service providers. This means that, until now, the Web is governed by contract law for the most part. Regulation

[93] J. DE WATCHER, *op. cit.*: "*From an IP strategy point of view, ... understanding and selecting those intangible assets that have more value as a secret than as an open, accessible intangible asset will become more difficult, but, arguably, also more important. On the other hand, businesses that reject the knee-jerk reaction to keep as much as possible hidden or secret, may find that they evolve faster and generate more new business opportunities. It is not a coincidence that Open Innovation has become such a tremendous success. Big Data is likely to reinforce that evolution.*"

(*e.g.* data protection and antitrust) will increasingly set the boundaries for data collection and raise awareness among the users, but – as for now – users may have a hard time trying to track who ultimately processes the personal data they have shared online.

B. Downstream phase (data reuse)

As discussed in chapter III, when it comes for the dataright holder to re-use Big Data for its business, a distinction may be usefully made between the data itself, the algorithm that serves for the purposes of compiling data, and the compilation that results from the analysis of Big Data.

B.1 Algorithm

The *Alice Corp v. CLS Bank Int'l* judgment has made it clear that software innovators may be subject to severe scrutiny when trying to enforce business analytics software. Courts subject patentability to the narrow condition that software solves a technical issue or improves a process (*e.g.* by reconfiguring data into different display sets, or by reformatting data from disparate sources), and not just executes routine operations (*e.g.* collecting or listing data) which an individual may do by hand.

As the patentability of computer-implemented remains difficult to pursue – and patents, where granted, will not cover the most abstract concept underlying Big Data analysis anyway – trade secret practices may effectively ensure confidentiality over algorithms and processes.

Big Data processes are relatively easy secrets to keep, as data mining often entails a high degree of subjective judgment. Because judgments are performed in an *ad hoc* fashion in response to unique circumstances in which a given dataset is initially gathered, they are a mystery to downstream users other than the dataright holder.[94] Non-disclosure covenants, in any event, allow additional protection and easy-to-enforce contractual remedies against breaches of confidential processes.

[94] M. MATTIOLI, *Disclosing Big Data cit.*, p. 570.

B.2 Data

Data, as said, cannot be *owned*. However, it can be *accessed.*

Hence, when granting access to a third party pursuant to a license agreement – in order to maximise the value of Big Data – contract law is suitable to provide legal basis for the dataright holder to both seek reward for its investment and maintain control over data, thereby overcoming the hurdles that the traditional legal schemes of civil law pose against the enforcement of proprietary rights in data.

B.3 Compilations of data

Most of the compilations of Big Data are the result of an automatic uploading by a software on a given database, which permits the dataright holder to retrieve the same data at a later stage, based on the criterion set by the dataright holder.

On this basis, compilations of data appear to be suitable to be covered by a *sui generis* right in database, so that dataright holders may prevent any unauthorised third party's extraction or re-utilisation of the database (or a substantial part thereof).

***

All the above considered, companies investing in Big Data may pursue an IP strategy that is some sort of a "*patchwork*" of multiple intellectual property rights: **secrecy for algorithms**, **contract for data**, and ***sui generis* copyright in database for compilations of data**.

The high enforcement costs that such a "*catch-all*" IP strategy may generate suggest that a preference for one or another solution may vary on a case-by-case basis. "*It depends*" may not be the most appealing answer to the question of how to protect investment in Big Data but, considering the different rates at which technology and regulation move, it may be also the most accurate in the circumstances.

BIBLIOGRAPHY

A. Encyclopaedias, commentaries, monographies and manuals

AA.VV., *Commentario Scialoja-Branca*, Bologna, 1988

AA.VV., *Digesto civile*, XI, Torino, 1994

AA.VV., *Diritto privato romano* (a cura di A. SCHIAVONE), Milano, 2003

AA.VV., *Enciclopedia del diritto*, Milano, 1989

AA.VV., *Internet e diritto civile, Atti del convegno svoltosi a Camerino il 26 e 27 settembre 2014* a cura di C. PERLINGIERI e L. RUGGERI, Napoli, 2015

C.M. BIANCA, *Diritto Civile, vol. III, Il Contratto*, Milano, 1984

G. BUTTARELLI, *Banche dati e tutela della personalità*, Milano, 1997

G. CASSANO, *Diritto dell'internet*, Milano, 2005

M. CINQUE, *Il minore contraente – Contesti e limiti della capacità*, Padova, 2007

V. CUFFARO, R. D'ORAZIO, V. RICCIUTO, *Il codice del trattamento dei dati personali*, Torino, 2007

V. CUFFARO, V. RICCIUTO, *La disciplina del trattamento dei dati personali*, Torino, 1997

V. CUFFARO, V. RICCIUTO, V. ZENO ZENCOVICH, *Trattamento dei dati e tutela della persona*, Milano, 1999

N. ELKIN-KOREN, N. WEINSTOCK NETANEL (eds.), in *The Commodification of Information*, New York, NY, 2002

G. FINOCCHIARO, *I contratti informatici*, Padova, 1997

E. GIANNANTONIO, M.G. LOSANO, V. ZENO ZENCOVICH, *La tutela dei dati personali*, Padova, 1999

A. LEVI, F. ZANICHELLI, *L'utilizzo dell'e-mail a fini pubblicitari: dallo «spamming» al «permission marketing»*, in *Riv. dir. ind.*, 2001, p. 194 ss.

D. MESSINETTI, *Oggettività giuridica delle cose incorporali*, Milano, 1970

G. DE NOVA, *Il contratto alieno*, Torino, 2008

P. FEMIA, *Interessi e conflitti culturali nell'autonomia privata e nella responsabilità civile*, Napoli, 1996

G.B. FERRI, *Causa e tipo nella teoria del negozio giuridico*, Milano, 1966

C. GALLI (a cura di), *Le nuove frontiere del diritto dei brevetti*, Torino, 2003

C. GALLI, A.M. GAMBINO (a cura di), *Codice commentato della Proprietà Industriale ed Intellettuale*, Torino, 2011

C. GALLI, A. ZAMA, *Stampa 3D. Una rivoluzione che cambierà il mondo?*, Bologna, 2014

G. GORLA, *Il contratto*, Milano, 1954

C. MACIEL, V. CARVALHO PEREIRA, *Digital Legacy and Interaction: Post Mortem Issues*, New York, NY, 2013

F. MESSINEO, *Dottrina generale del contratto*, Milano, 1952

M. RICOLFI, *Il contratto di merchandising nel diritto dei segni distintivi*, Milano

S. RODOTÀ, *Tecnologie e diritti*, Bologna, 1995

S. RODOTÀ, *Tecnopolitica. La democrazia e le nuove tecnologie della comunicazione*, Roma-Bari, 2004

G. SANTINI, *I diritti della personalità nel diritto industriale*, Padova, 1959

I. SIDHU, T.C. DOYLE, *The Digital Revolution: How Connected Digital Innovations Are Transforming Your Industry, Company & Career*, London, 2016

B. Articles

B.1 EU

J. EXTEN-WRIGHT, *The Trade Secrets Directive: New protection for businesses?*, 18 July 2016, available https://www.dlapiper.com/en/abudhabi/insights/publications/-2016/07/the-trade-secrets-directive/

R. KEMP, *Legal aspects of managing Big Data*, in *Computer Law & Security Review*, 2014, pp. 482-491

R. KEMP, *Legal Rights in Data*, *Computer Law & Security Review*, 2011, pp. 139-151

I. GRAEF, *Data portability at the crossroads of data protection and competition policy*, available at http://www.agcm.it/component/joomdoc/eventi/convegni/201611-09_07.pdf/download.html

T. HOEREN, *Big data and ownership in data: recent developments in Europe*, in *European Intellectual Property Review*, 2014, pp. 751-754

A. LAMADRID, S. VILLIERS, *Big Data, privacy and competition law: do competition authorities know how to do it?*, in *CPI Antitrust Chronicle*, 2017, pp. 7-10

R. MAHNKE, *Big Data as a Barrier to Entry*, in *CPI Antitrust Chronicle*, 2015, pp. 2-6

M. MATTIOLI, *Disclosing Big Data*, in *Minnesota Law Review*, 2014, pp. 535-583

H. PEARSON, A. DUFFUS, P. WARD, R. CLABAUGH, F. BOURGUET, *Software and business methods: should you be patenting them?*, in *Practical Law*¸ 2017 (available at www.practicallaw.com/4-200-9964)

C. PRINS, *Property and Privacy: European prospectives and the commodification of our identity*, in *Information law series*, available at https://papers.ssrn.com/-sol3/papers.cfm?abstract_id=929668

M. REBEIRO, M. EVANS, *Big Data: protecting rights and extracting value*, in *Practical Law*¸ 2015, pp. 1-16

B. SEGALIS, N. SHAH, *FTC Looks to Link Do-Not-Track, Big Data Privacy Concerns; Seeks Solutions*, available at http://www.infolawgroup.com/2012/03/articles/data-privacy-law-or-regulation/ftc-looks-to-link-donottrack-big-data-privacy-concerns-seeks-solutions/

B.2 Italy

F. AGNINO, *Fino a che punto è possibile disporre contrattualmente dei propri diritti? (Vedi contratto FB)*, in *Giur. mer.*, 2012, pp. 2555-2568

F. ASTONE, *Il rapporto tra gestore e singolo utente: questioni generali*, in *Ann. it. dir. ind.*, 2011, pp. 102-124

G. AZZARITI, *Internet e Costituzione*, 2011, available at http://www.costituzionalismo.it/articoli/392/

M. BASSINI, *Point and click: la tutela del consumatore nel commercio elettronico*, ILSU Working Paper no. 2008-13/IT, available at *web www.diritto.it*

M. BOGNI, A. DEFANT, *Big data e diritti IP: problemi di privacy, contrasto della contraffazione e protezione dei database e delle pagine dei social networks*, *Dir. ind.*, 2015, pp. 117-126

S.F. BONETTI, *La tutela dei consumatori nei contratti di accesso ad internet: i contratti dei consumatori e la privacy tra fattispecie giuridiche e modelli contrattuali italiani e statunitensi*, in *Dir. inf.*, 2002, pp. 1087-1140

L. BOSOTTI, *I limiti di brevettabilità nelle innovazioni della rete*, in *Dir. ind.*, 2015, pp. 137-145

M. BOTTI, *La tutela del software: l'attuale posizione dell'EPO*, in *Dir. ind.*, 2015, pp. 146-148

F. BRAVO, *Invio di sms commerciali e risarcimento del danno da illecito trattamento di dati personali*, in *Dir. informatica*, 2007, p. 798-814.

F.D. BUSNELLI, *Capacità e incapacità di agire del minore*, *in Dir. fam.*, 1982, pp. 54-70

F. CARINGELLA, *Alla ricerca della causa nei contratti gratuiti atipici*, in *Foro it.*, 1993, I, pp. 1508-1522

R. CATERINA, *Cyberspazio, social network e teoria generale del contratto*, in *AIDA*, 2011, pp. 93-101

S.A. CERRATO, *I rapporti contrattuali (anche associativi) tra i soggetti del social network*, in *AIDA*, 2011, pp. 168-218

A. COGO, *Le regole del contratto tra social network e utente sull'uso della proprietà intellettuale del gestore, dell'utente e degli altri utenti – riflessioni a partire dall'individuazione del fenomeno, dei suoi soggetti e della funzione del contratto*, in *AIDA*, XX, 2011, pp. 305-341

P. CRUGNOLA, *Problemi giuridici relativi all'uso di fotografie per la pubblicità commerciale*, in Dir. aut., 1973, pp. 418-440

A. DE FRANCESCHI, M. LEHMANN, *Data as a Tradeable Commodity and the new Steps for their Protection*, in *The Italian Law Journal*, 1, 2015, avaible at https://www.academia.edu/12144964/Alberto_De_Franceschi_Michael_Lehmann_-_Data_as_Tradeable_-Commodity_and_New_Measures_for_their_Protection

P. DI MICO, *Il rapporto tra diritto di autore e social network: un nuovo capitolo, ma non l'ultimo*, in *Dir. aut.*, 2010, pp. 262-276

F. FAINI, *Open data, big data e mercati*, in *Rivista Elettronica di Diritto, Economia, Management*, 2016, pp. 29-41

V. FALCE, *Standard e cloud computing*, in *Dir. ind.*, 2015, pp. 155-159

G.B. FERRI, *Meritevolezza dell'interesse e utilità sociale*, in *RDCo*, 1971, I, pp. 81-97

G.B. FERRI, *Ancora in tema di meritevolezza dell'interesse*, in *RDCo*, 1979, I, pp. 1-14

R. FRAU, *Profili del consenso al trattamento dei dati personali per fini economici nell'esperienza italiana. Raffronti con la normativa spagnola*, in *Resp. civ. e prev.*, 2010, pp. 2598-2619

F. GALGANO, *Diritto ed economia alle soglie del nuovo millennio*, in *Contr. impr.*, 2000, pp. 189-205

C. GALLI, *Banche dati e giornali*, in *AIDA*, 1997, pp. 29-44

C. GALLI, *Diritti IP e risarcimento dei danni: un rapporto difficile che sta cambiando*, (Relazione al convegno "La contraffazione non paga. Risarcimento del danno e reversione degli utili del contraffattore tra problemi applicativi e strategie per le imprese", Parma, 21 ottobre 2011), in *Dir. ind.*, 2012, pp. 105-106

C. GALLI, *Diritti di proprietà intellettuale e remunerazione degli investimenti*, (Relazione al Convegno "IP e costituzioni", Università di Pavia, 23-24 settembre 2005), in *AIDA*, 2005, pp. 68-79

C. GALLI, *I diritti IP nel mercato globale e nella nuova economia digitale: le ragioni di un Convegno*, in *Dir. ind.*, 2015, pp. 105-106

C. GALLI, *Le sfide del commercio elettronico al sistema della moda*, in *Dir. ind.*, 2013, pp. 342-353

C. GALLI, *L'innovazione nel web: opportunità e problematiche giuridiche*, *Dir. ind.*, 2015, pp. 127-136

C. GALLI, *Risarcimento del danno e retroversione degli utili: le diverse voci di danno*, in *Dir. ind.*, 2012, pp. 109-120

C. GALLI, P. PAGANINI, *How Italy successfully improved its approach to intellectual property rights – Case study on Italy*, in *2014 International Property Rights Index*, available at www.internationalpropertyrightsindex.org

M. GRANIERI, *Le clausole ricorrenti nei contratti dei social network dal punto di vista della disciplina consumeristica dell'Unione europea*, in *AIDA*, 2011, pp. 125-142

G. GUGLIELMETTI, *La tutela delle banche dati con diritto sui generis nella Direttiva 96/9/CE*, in *Contr. e impr. Europa*, 1997, pp. 177-195

N. IRTI, *Scambi senza accordo*, in *Riv. trim. dir. proc. civ.*, 1998, pp. 347-364

L.MANSANI, *Contenuti generati dagli utenti*, in *AIDA*, 2010, pp. 244-257

A. MANTELERO, *Big data: i rischi della concentrazione del potere informativo digitale e gli strumenti di controllo*, in *Dir. inf.*, 2012, pp. 135-144

A. MANTELERO, *Observatory on ICT law: the control of over digital information in the Big Data Era*, in *Contratto e impresa. Europa*, 2012, pp. 961-966

R. MORO DI VISCONTI, *Valutazione dei Big Data e impatto su innovazione e ditigal branding*, in *Dir. ind.*, 2016, pp. 46-53

R. MORO DI VISCONTI, *Internet delle cose, Networks e plusvalore della connettività*, in *Dir. ind.*, 2016, pp. 536-544

G. MUSCOLO, *Innovazione nella rete e diritti non titolati: il ruolo di know-how, copyright, banche dati e pratiche commerciali sleali*, in *Dir. ind.*, 2015, pp. 114-116

G. OPPO, *Disumanizzazione del contratto?*, in *Riv. trim. dir. proc. civ.*, 1998, pp. 525-533

P. PAGANINI, *Verso l'Internet delle cose*, in *Dir. ind.*, 2015, pp. 107-113

P. PASSAGLIA, *Diritto di accesso ad Internet e giustizia costituzionale comparata. Una (preliminare) indagine comparata*, available at http://www.giurcost.org/studi/passaglia.htm

S. PATTI, *La globalizzazione del diritto e il contratto*, in *Obbl. contr.*, 2009, pp. 495-499

P. PERLINGIERI, *L'informazione come bene giuridico*, in *Rass. dir. civ.*, 1990, pp. 326-347

P. PERLINGIERI, *Nuovi profili del contratto*, in *Riv. crit. dir. priv.*, 2001, pp. 223-246

P. PERLINGERI, *Mercato, solidarietà e diritti umani*, in *Rass. dir. civ.*, 1995, pp. 84-117

A.R. POPOLI, *Social network e concreta protezione dei dati sensibili: luci ed ombre di una difficile convivenza*, in *Dir. inf.*, 2014, pp. 981-1017

G. RESTA, *La disponibilità dei diritti fondamentali e i limiti della dignità (note a margine della Carta dei Diritti)*, in *Riv. dir. civ.*, 2002, pp. 801-848

G.M. RICCIO, *Social network e responsabilità civile*, in *Dir. informatica*, 2010, pp. 859-871

M. RICOLFI, *Questioni in tema di regime giuridico dello sfruttamento commerciale dell'immagine*, in *Nuova. giur. civ. comm.*, 1992, pp. 44-56

R. ROMANO, *Diritto d'autore, tecniche di tutela e risarcimento del danno*, in *Dir. aut.*, 2003, pp. 406-415

R. ROMANO, *Innovazione, rischio e "giusto equilibrio" nel divenire della proprietà intellettuale*, in *Riv. dir. civ.*, 2015, pp. 532-553

E. ROSATI, G. SARTOR, *Social networks e responsabilità dei provider*, in *European University Institute - EUI Working Papers*, available at www.cadmus.eui.eu

P. SAMMARCO, *Le clausole contrattuali di esonero e trasferimento della responsabilità inserite nei termini d'uso dei servizi del web 2.0*, in *Dir. inf.*, 2010, pp. 631-643

S. SCALZINI, *I servizi di online social network tra privacy, regole di utilizzo e violazione dei diritti dei terzi*, in *Giur. mer.*, 2012, pp. 2569-2590

L. SCHIUMA, *Il Software tra brevetto e diritto d'autore*, Relazione al Convegno "La tutela del software tra brevetto e diritto d'autore", Facoltà di Giurisprudenza dell'Università Lumsa di Roma, Roma, 24 giugno 2004, in *Riv. dir. civ.*, 2007, pp. 683-707

C. SCOGNAMIGLIO, *Il diritto all'utilizzazione del nome e dell'immagine delle persone celebri*, in Dir. informatica, 1988, pp. 1-40

S. SICA, G. GIANNONE CODIGLIONE, *Social network sites e il «labirinto» delle responsabilità*, in *Giur. mer.*, 2012, pp. 2715-2733

E. TOSI, *Prime osservazioni sull'applicazione della disciplina generale della tutela dei dati personali a internet e al commercio elettronico*, in *Dir. informatica*, 1999, p. 591 ss.

W. VIRGA, *Inadempimento di contratto e sanzioni private nei social network*, in *Ann. it. dir. ind.*, 2011, pp. 219-240

V. ZENO ZENCOVICH, *Profili negoziali degli attributi della personalità*, in *Dir informatica*, 1993, p. 545-597

V. ZENO ZENCOVICH, *Sull'informazione come «bene» (e sul metodo del dibattito giuridico)*, in *Rass. dir. civ.*, 1999,pp. 485-491

(*Bibliography follows on the next page*)

B.3 United States of America and Canada

J. ANGWIN, *The Web's New Gold Mine: Your Secrets*, the *Wall Street Journal*, 30 July 2010, available at http://online.wsj.com/article/SB10001424052748703940904575-395073512989404.html

E. ASANO, *How Much Time Do People Spend on Social Media?*, available at http://www.socialmediatoday.com/marketing/how-much-time-do-people-spend-social-media-infographic

T. BERNERS-LEE, J. HENDLER, O. LASSILA, *The Semantic Web - A new form of Web content that is meaningful to computers will unleash a revolution of new possibilities*, available at https://www.scientificamerican.com/article/the-semantic-web/

D.M. BOYD, N.B. ELLISON, *Social Network Sites: Definition, History, and Scolarship*, in *Journal of Computer-Mediated Communication*, 2007, pp. 210-230

D. BUTLER, *When Google Got Flu Wrong*, 494 Nature 155, 155-56 (2013), available at http://www.nature.com/news/when-google-got-flu-wrong-1.12413

K. CUKIER, V. MAYER-SCHOENBERGER, *The Rise of Big Data: How It's Changing the Way We Think About the World*, Foreign Affairs, 3, 2013, pp. 28-40

J. DE WATCHER, *Big Data and IP Business Strategy*, in *Trading Secrets – Web portal TradeSecretsLaw.com*, 2014

J. BICK, *Big Data rights protection found in Internet copyright law*, J. Bick, *Big Data rights protection found in Internet copyright law*, New Jersey Law Journal April 15, 2015

V. CHIAPPETTA, *Myth, Chameleon or Intellectual Property Olympian? A Normative Framework Supporting Trade Secret Law*, 8 GEO. MASON L. REV. 69, 76 (1999)

D. CROUCH, *A Comparison of the EU Trade Secrets Directive and the US Defend Trade Secrets Act*, 16 May 2016 post on the Patently-O blog, available at https://patentlyo.com/patent/2016/05/comparison-secrets-directive.html

M. FLYNN, R. HUNG, *Big Data: Legal Aspects under Canadian Law*, in *World Data Protection Report*, *Bloomberg BNA*, 2015, pp. 19-24

J. HOOFNAGLE, J. WHITTINGTON, *Free: Accounting for the Costs of the Internet's Most Popular Price*, *UCLA Law Review*, 2014, pp. 606-670

R. MERGES, *Symposium: Go ask Alice – what can you patent after Alice v CLS Bank*, 20 June 2014 post on the SCOTUS blog, available at http://www.-scotusblog.com/2014/06/symposium-go-ask-alice-what-can-you-patent-after-alice-v-cls-bank/

J.A. OBAR, S. WILDMAN, *Social media definition and the governance challenge: An introduction to the special issue. Telecommunications policy*, *Quello Center Working Paper no. 2647377*, pp. 745-750

D.A. PRANGE, *Navigating the protection of big data*, in *Intellectual Property Magazine*, 1/2017, pp. 54-55

D.A. PRANGE, *Big data and trade secrets: part 2*, in *Intellectual Property Magazine*, 2/2017, pp. 41-42

A. GOTHING, A. MUÑOZ-KAPHING, *Big data and Alice v CLS: predicting what's next*, in *Intellectual Property Magazine*, 7/2014, pp. 21-22

L.H. GUTTENPLAN, L. JOHNSTON, *Big brother, Big Data: how museums are collecting, using and storing ditigal data, ALI CLE Course of Study Materials*, 2014, pp. 1-16

J. HUNTLEY, N. MCKERREL, S. ASHGAR, *Universal Service, the Internet and the Access Deficit*, SCRIPTed Vol 1(2), 2004, available at https://script-ed.org/wp-content/uploads/2016/07/1-2-Huntley.pdf

J. KELLY, *Big Data Vendor Revenue and Market Forecast 2013-2017*, available at http://wikibon.org/wiki/v/Big_Data_Vendor_Revenue_and_Market_Forecast_2013-2017

E. MICHIKO MORRIS, *Alice, Artifice and Action - and Ultramercial*, 8 July 2014 post on the Patently-O blog, available at http://fstp-expert-system.typepad.com/files/94-e.-morris_alice-artifice-and-action-and-ultramercial_iu-i.n.-08.07.2014.pdf;

E. MICHIKO MORRIS, *What is Technology?, V 20 Boston University Journal of Science and Technology Law*, 2014.

A. NARAYANAN AND V. SHMATIKOV, *Robust De-anonymization of Large Datasets (How to Break Anonymity of the Netflix Prize Dataset*, the University of Texas, 2008 available at https://arxiv.org/PS_cache/cs/pdf/0610/0610105v2.pdf

D. NAVETTA, *Legal Implications of Big Data: A Primer, Issa Journal*, 2013, available at https://c.ymcdn.com/sites/www.issa.org/resource/resmgr/journalpdfs/feature0313.-pdf

T. O'REILLY, *What Is Web 2.0 - Design Patterns and Business Models for the Next Generation of Software*, available at www.oreilly.com

G. PRESS, *Big Data News: A Revolution Indeed*, available at http://www.forbes.com/sites/gilpress/2013/06/18/big-data-news-a-revolution-indeed/-#5496d3b7b9fb

G. PRESS, *12 Big Data Definitions: What's Yours?*, available at http://www.forbes.com/sites/gilpress/2014/09/03/12-big-data-definitions-whats-yours/#3aafa1e421a9

P. SAMUELSON, *Privacy as Intellectual Property?*, 52 Stanford Law Review, 2000, pp. 1125 ss.

J. TSOUKAS, N.J. VERMETTE, *Intellectual Property Protection for Big Data in Canada and the United States*, in *Information Law Journal*, 2016, 7(3), pp. 18-23

T. VARE, M. MATTIOLI, *Big business, big government and big legal questions*, in *Managing IP*, October 2014, pp. 46-48

W. WHITE, N.P. TATONETTI, N.H. SHAH, R.B ALTMAN, E. HORVITZ, *Web-Scale Pharmacovigilance: Listening to Signals from the Crowd*, 2013, available at http://jamia.bmj.com/content/20/3/404.full.pdf

C. Case law precedents

C.1 Australia

Federal Court of Australia's decision of 5 September 2005 in *Universal Music Australia Pty Ltd v Sharman License Holdings Ltd [2005] FCA 1242*

Federal Court of Australia's decision of 18 December 2006 in *Cooper v Universal Music Australia Pty Ltd [2006] FCAFC 187*

Federal Court of Australia's decision of 4 February 2010 in *Roadshow Films Pty Ltd v iNet Limited* (No. 3) [2010] FCA 24

C.2 Canada

CCH Canadian Ltd. v. Law Society of Upper Canada, 2004 SCC 13

Attorney General v. Amazon.com, Inc., 2011 FCA 328

C.3 Costa Rica

Sala Constitucional De La Corte Suprema De Justicia, Judgement of 30 July 2010, *sentencia:* 12790, *expediente:* 09-013141-0007-CO

C.4 EU

Commission Decision of 11 March 11 2008 in case No. COMP/M.4731 – *Google/ DoubleClick*;

Commission Decision of 3 October 3 2014 in case No. COMP/M.7217 – *Facebook/Whatsapp*;

Commission decision of 3 July 3 2001 in case No. COMP D3/38.044 – *IMS Health*;

Commission decision of 20 December 2012 in case No. case AT.39654 – *Reuters Instrument Codes*;

Commission Decision of 6 December 2016 in case No. COMP/M.8124 – *Microsoft/LinkedIn*

Court of Justice of the European Union, Judgment of 9 November 2004, (C-203/02) *British Horseracing Board v. William Hill*

Court of Justice of the European Union, Judgment of 9 October 2008, (C-304/07) *Directmedia v. Albert-Ludwigs-Universitaet*

Court of Justice of the European Union, Judgment of 1 March 2012, (C-604/10) *Football Dataco v. Yahoo!*

Court of Justice of the European Union, Judgment of 18 October 2012, (C-173/11) *Football Dataco v. Sportradar*

Court of Justice of the European Union, Judgment of 13 May 2014, (C-131/12) *Google Spain SL and Google Inc. v Agencia Española de Protección de Datos (AEPD) and Mario Costeja González*

Court of Justice of the European Union, Judgment of 19 December 2013, (C-202/12) *Innoweb v. Wegener*

Court of Justice of the European Union, Judgment of 3 July 2012, (C-128/11) *UsedSoft GmbH v. Oracle International Corp*

European Patent Office, Board of Appeal's decision of 1 July 1998, (T-1173/97) *Computer program product/IBM*

C.5 France

Conseil constitionnel, 10 June 2009 no. 2009-580 DC

Cour de cassation, March 2009, Precom *and Ouest France v. Direct Annonce*

Cour d'appel de Paris, 15 November 2013, *Pressimmo online v. Yakaz and Gloobot,*

C.6 Germany

Bundesgerichtshof, 2.06.2011, *Automobil-Onlinebörse* (I ZR 159/10)

LAG Sachsen, Urteil vom 17.01.2007 – 2 Sa 808/05, MMR 2008, 416

OLG Nürnberg 1. Strafsenat, Beschluss vom 23.01.2013 - 1 Ws 445/12, CR 2013

C.7 Italy

App. Milano Sez. spec. propr. industr. ed intell., 21 November 2011, no. 3206

Autorità Garante della Concorrenza e del Mercato, decision no. 10276 of 20 December 2001, in *Giust. civ.*, 2002, p. 1748 ss.

Corte Costituzionale, decision no. 247 of 21 October 2015

Corte Suprema di Cassazione, decision no. 2288 of 6 February 2004, in *Contratti*, 2004, p. 801

Corte Suprema di Cassazione, decision no. 6496 of *7 June 1991, in Fisco*, 1991, p. 5007

Corte Suprema di Cassazione, decision no. 10612 of 9 October 1991, in *Giust. Civ.*, 1991, p. 2895

Corte Suprema di Cassazione, decision no. 3142 of 13 May 1980, in *Mass. Giur. It.*, 1980

Garante per il trattamento dei dati personali, decision of 13 May 2008, docweb no. 1521775

Garante per il trattamento dei dati personali, decision of 12 March 2003, docweb no. 29844

Garante per il trattamento dei dati personali, decision dated 28 May 1997, docweb no. 40425

T.A.R. Lazio, sez. I, ruling dated 26 September 2016, no. 10016

Trib. Bologna Sez. spec. propr. industr. ed intell., 10 August 2011

Trib. Latina, 19 June 2006, in *Foro it.*, 2007, c. 324 ss.

Trib. Milano, 17 June 2016, *Google-Attrakt*

Trib. Roma, 26 July 2007, in *Dir. informatica*, 2007, p. 859 ss.

C.8 Ireland

High Court Of Justice, Chancery Division, 17 October 2014, *Cartier International AG, Montblanc Simplo GmbH and Richemont International SA vs British Sky Broadcasting Limited, British Telecommunications Plc, EE Limited, Talktalk Telecom Limited and Virgin Media Limited* [2014] EWHC 3354 (Ch)

Irish High Court's decision of 16 April 2010 *Emi Records (Ireland) Ltd et al. vs Eircom Ltd.* [2010] IEHC 106

C.9 UK

Aerotel Ltd v. Telco Holdings Ltd [2006] EWCA Civ 1371

Attheraces Ltd & Another v. The British Horse Racing Board [2005] WEHC 3015 (Ch)

Coco v. AN Clark (Engineers) Ltd [1968] F.S.R. 415

Oxford v Moss (1979) 68 Cr App Rep 183

Your Response Ltd v. Datateam Business Media Ltd [2014] EWCA Civ 281; [2014] 3 W.L.R. 887

C.10 United States of America

Alice Corp. v. CLS Bank International 573 U.S. __, 134 S. Ct. 2347 (2014)

American Guarantee and Liability Insurance Co v. Ingram Micro, Inc 2000 WL 726789 (D. Ariz, 2000)

AOL v. St Paul Mercury Insurance 207 F. Supp. 2d 459 (E.D. Va 2002)

Association for Molecular Pathology v. Myriad Genetics [1] No. 12-398 (569 U.S. ___ June 13, 2013)

Bilski v. Kappos 561 U.S. 593 (2010)

Bragg v. Linden Research, Inc. 487 F. Supp. 2d 593 (E.D.Pa. 2007)

Diamond v. Diehr 450 U.S. 175 (1981)

Feist Publications, Inc., v. Rural Telephone Service Co. 499 U.S. 340 (1991)

GC Technology Development, LLC v. Big Fish Games, Inc. 2:16-cv-00857-RCJ-VCF (D Nev 29 August 2016)

Gottschalk v. Benson 409 U.S. 63 (1972)

In re Yazoo Pipeline Co LP 459 B.R. 636 (Bankr. S.D. Tex. 2011)

Mayo Collaborative Services v. Prometheus Laboratories, Inc. 132 S. Ct. 1289 (2012)

National Cable & Telecommunications Association et al. v. Brand X Internet Services et al. (04-277) 545 U.S. 967 (2005)

Parker v. Flook 437 U.S. 584 (1978)

Reno v. American Civil Liberties Union 521 U.S. 844 (1997)

Retail Systems Inc v. CNA Insurance Cos 469 N.W. 2d 735 (Minn. App. 1991)

Sporn v. MCA Records 58 N.Y. 2d 482, 489 (1983)

Thyroff v. Nationwide Mutual Ins. Co N.Y. 3d 283 (2007)

United States v. Aleynikov, 10 Crim. 96 (S.D.N.Y. Feb 10, 2010), 2010 WL 4000356

Whyte v. Schlage Lock Co 125 Cal Rptr 2d 277 (4th Dist 2002)

WL Gore & Assoc, Inc v. Wu 2006 WL 2692584 (Del Ch 2006)

***