jipitec

1 | 2018

Volume 9 (2018)
Issue 1 ISSN 2190-3387

Editors:
Thomas Dreier
Axel Metzger
Gerald Spindler
Lucie Guibault
Miquel Peguera
Séverine Dusollier
Chris Reed
Karin Sein

Journal of Intellectual Property, Information Technology, and Electronic Commerce Law

www.jipitec.eu

jipitec

Journal of Intellectual Property, Information Technology and Electronic Commerce Law

Volume 9 Issue 1 May 2018
www.jipitec.eu
contact@jipitec.eu

A joint publication of:

Prof. Dr. Thomas Dreier, M. C. J.,
KIT - Karlsruher Institut für Technologie, Zentrum für Angewandte Rechtswissenschaft (ZAR), Vincenz-Prießnitz-Str. 3, 76131 Karlsruhe Germany

Prof. Dr. Axel Metzger, LL. M.,
Humboldt-Universität zu Berlin, Unter den Linden 6, 10099 Berlin

Prof. Dr. Gerald Spindler,
Dipl.-Ökonom, Georg-August-Universität Göttingen, Platz der Göttinger Sieben 6, 37073 Göttingen

Karlsruhe Institute of Technology, Humboldt-Universität zu Berlin and Georg-August-Universität Göttingen are corporations under public law, and represented by their respective presidents.

Editors:
Thomas Dreier
Axel Metzger
Gerald Spindler
Lucie Guibault
Miquel Peguera
Séverine Dusollier
Chris Reed
Karin Sein

Board of Correspondents:
Graeme Dinwoodie
Christophe Geiger
Ejan Mackaay
Rita Matulionyte
Giovanni M. Riccio
Cyrill P. Rigamonti
Olav Torvund
Mikko Välimäki
Rolf H. Weber
Andreas Wiebe
Raquel Xalabarder

Editors-in-charge for this issue:
Lucie Guibault and Karin Sein

Technical Editor:
Philipp Schmechel

ISSN 2190-3387

Funded by
DGRI Deutsche Gesellschaft für Recht und Informatik e.V.

Table Of Contents

Editorial

by **Lucie Guibault and Karin Sein**

© 2018 Lucie Guibault and Karin Sein

Everybody may disseminate this article by electronic means and make it available for download under the terms and conditions of the Digital Peer Publishing Licence (DPPL). A copy of the license text may be obtained at http://nbn-resolving.de/urn:nbn:de:0009-dppl-v3-en8.

Recommended citation: Lucie Guibault and Karin Sein, Editorial, 9 (2018) JIPITEC 1 para 1.

1 In the wake of the recent *Cambridge Analytica* scandal and in the midst of the controversy around the European copyright reform, our fresh Spring number will delight readers for its varied and in-depth coverage of many of the hot topics in the contemporary digital legal discourse. Variety is also present in relation to scientific methods, as several articles use inter-disciplinary approach, combining traditional legal analysis with the application of empirical methods, behavioral economics and psychology.

2 What the 2013 *Snowden* affair was to the covert massive online surveillance of citizens by secret government services, are the 2018 *Cambridge Analytica* revelations to the manipulation of social media profiles for political campaigning purposes. Both represent a major breakdown in the way governments and private entities ought to deal with personal data. The consequences of Cambridge Analytica's operations are enormous: Trump and Brexit! The fact that the data consulting firm was located on UK territory while engaging in dubious activities is all the more disconcerting, as one would think that the firm was bound by the European norms of protection of personal data. Or is it that these norms, including the newly implemented General Data Protection Regulation (GDPR), are incapable of preventing this type of malicious activities? Probably. In reality, the citizen's naivety and trust are blatantly misused on all sides. The *Snowden* and *Cambridge Analytica* affairs show once more that a regulatory system based on the notion of consent to the collection and processing of an individual's personal data leaves gaping holes in the protection.

3 The two first articles in this issue propose other ways to look at the problem of data protection, i.e. through an increase in transparency of the algorithmic decision-making process and through greater empowerment of data subjects before disclosure of personal data. On the first point, *Guido Noto La Diega* speaks against the exclusive automated decision-making and presents three legal routes – intellectual property law, data protection law and the access right under the freedom of information regime – that would help to 'open up' the algorithms. From these three routes the GDPR rules seem to be the most promising for the affected persons, although much is still depending on the national implementation measures. He concludes that only an integrated approach combining elements of all these three routes would be able to provide the affected person with an effective remedy.

4 On the second point, *Santiago Ramírez López* explores the possibilities to learn from the behavioral economics and Kahneman's theory on thinking fast and slow in order to empower the data subjects. He proceeds from the assumption that while Western traditions embrace the concept of control of the data subjects as the main guideline of data protection, the reality of the online world has shown that the informed content model has failed to provide such control. He analyzes alternative methods of providing user-friendly information online, mainly using the example of the Human Readable layer of the Creative Commons license and also considers it necessary to establish guidelines for such icon-based information model.

5 Data protection is not the only controversial topic these days in Europe. The *Proposal for a Directive on Copyright in the Digital Single Market* (DSM), published in September 2016 and expected to be adopted by 2019, is turning into an arm-wrestling match. The top three most disputed provisions in the Proposal are the Commission's push for the adoption of a new press publishers' right (article 11), an obligation on online platforms to put upload content filters (article 13), and a narrow exception for text and data mining (article 3). All three proposed provisions risk severely encroaching upon Europe's principles of open science and freedom of expression. The opposition to the press publisher's right and the content filter obligation is so strong and the perceived weakness of the remaining provisions in the Proposal so great that more than a hundred legal scholars and a plethora of organizations, including associations of European public institutions, companies and start-ups, journalists and libraries, news publishers and civil society organizations, have let their voices heard in different open letters to Member of Parliament Voss, to express their deep concerns about the DSM Proposal.[1]

6 With respect to research data in particular, open-research advocates argue that limiting the beneficiaries of the proposed text and data mining exception only to research organisations and only for purposes of scientific research would effectively undermine the European Union's commitment to the 3 O's: Open Innovation, Open Science and Open to the World.[2] As in the case of data protection, it might be useful to examine other possible legal avenues than copyright law for the use of publicly funded research. One such avenue could be the inclusion of public research and educational establishments within the scope of the Directive regulating the re-use of public sector information ('PSI Directive'), as presented in *Heiko Richter*'s article. The paper evaluates the legal consequences of such an inclusion. As the PSI Directive is characterized by considerable legal uncertainty, it is difficult to derive robust assumptions that can form the basis for predicting the effects of extending the PSI Directive's scope to research information. Richter concludes that a potential revision of the PSI Directive aiming to include research organizations and educational establishments should reduce this uncertainty.

7 In the context of the European copyright reform controversy, the strongest argument that can be made against the Commission's ill-conceived plans is by putting facts forward. *Kristofer Erickson* and *Martin Kretschmer* do just that in their article entitled '"This Video is Unavailable" Analyzing Copyright Takedown of User-Generated Content on Youtube'. Using empirical methods, their analysis of right holder behavior complements and offers a new perspective on recent empirical work assessing the appropriateness of notice-and-takedown procedures as a means of balancing the interests of right holders, innovative services and citizens. More specifically, they investigate the factors that motivate takedown of user-generated content by copyright owners. The main finding is that policy concerns frequently raised by right holders are not associated with statistically significant patterns of action. They suggest that evolving policy on intermediary liability - for example with respect to imposing filtering systems (automatically ensuring "stay-down" of potentially infringing content) - should be carefully evaluated against evidence of actual behavior, which this study shows may differ materially from stated policy positions. In other words, a measure such as that proposed in article 13 of the DSM Proposal would, in line with these findings, not necessarily address the true concerns of right holders, while bearing the risk of creating disproportionately high obstacles to user-generated content as to have a chilling effect on users' exercise of their freedom of expression. This should be the nail in article 13 DSM's coffin!

8 The last document in this issue takes the discussion full circle on the topic of data protection and citizen empowerment. The Weizenbaum Institute research group led by *Axel Metzger* discusses the Proposal of Digital Content Contracts Directive that is currently in the final stage of trilogues. The authors concentrate, inter alia, on the concept of data-as-counterperformance claiming that the notion should be explicitly kept in the operative text of the directive and that its scope of application should be opened irrespective of whether the consumer provides personal data actively or passively. They also encourage to regulate the multi-party scenarios in the context of supplying smart goods at the EU level – questions that under the current version of the proposal are left to the national law. Indeed, even if it is too late for the Digital Content Directive to take up new regulatory issues, problems arising from the so-called unbundling could still be dealt with during the ongoing discussions and 'digitisation' of the amended Proposal of Consumer Sales Directive.

Enjoy the reading!

Lucie Guibault and Karin Sein

1 Open Letter in Light of the 27 April 2018 COREPER I Meeting, Brussels, 26 April 2018, available at: <http://copybuzz.com/wp-content/uploads/2018/04/Open_Letter_on_Copyright_Reform_27_April_COREPER_Meeting.pdf>; Statement from EU Academics on Proposed Press Publishers' Right, 24 April 2018, available at: <https://www.ivir.nl/academics-against-press-publishers-right/>; Letter to MEP Axel Voss, Brussels, 24 April 2018, <https://www.communia-association.org/wp-content/uploads/2018/04/OpenLetter_AxelVoss_DeleteArticle11_English.pdf>.

2 European Commission, Research and Innovation, Brussels, <https://ec.europa.eu/research/openvision/index.cfm>.

Against the Dehumanisation of Decision-Making

Algorithmic Decisions at the Crossroads of Intellectual Property, Data Protection, and Freedom of Information

by **Guido Noto La Diega***

Abstract: This work presents ten arguments against algorithmic decision-making. These revolve around the concepts of ubiquitous discretionary interpretation, holistic intuition, algorithmic bias, the three black boxes, psychology of conformity, power of sanctions, civilising force of hypocrisy, pluralism, empathy, and technocracy. Nowadays algorithms can decide if one can get a loan, is allowed to cross a border, or must go to prison. Artificial intelligence techniques (natural language processing and machine learning in the first place) enable private and public decision-makers to analyse big data in order to build profiles, which are used to make decisions in an automated way. The lack of transparency of the algorithmic decision-making process does not stem merely from the characteristics of the relevant techniques used, which can make it impossible to access the rationale of the decision. It depends also on the abuse of and overlap between intellectual property rights (the "legal black box"). In the US, nearly half a million patented inventions concern algorithms; more than 67% of the algorithm-related patents were issued over the last ten years and the trend is increasing. To counter the increased monopolisation of algorithms by means of intellectual property rights (with trade secrets leading the way), this paper presents three legal routes that enable citizens to 'open' the algorithms. First, copyright and patent exceptions, as well as trade secrets are discussed. Second, the EU General Data Protection Regulation is critically assessed. In principle, data controllers are not allowed to use algorithms to take decisions that have legal effects on the data subject's life or similarly significantly affect them. However, when they are allowed to do so, the data subject still has the right to obtain human intervention, to express their point of view, as well as to contest the decision. Additionally, the data controller shall provide meaningful information about the logic involved in the algorithmic decision. Third, this paper critically analyses the first known case of a court using the access right under the freedom of information regime to grant an injunction to release the source code of the computer program that implements an algorithm. Only an integrated approach – which takes into account intellectual property, data protection, and freedom of information – may provide the citizen affected by an algorithmic decision of an effective remedy as required by the Charter of Fundamental Rights of the EU and the European Convention on Human Rights.

Keywords: Algorithmic decision-making; algorithmic bias; right not to be subject to an algorithmic decision; GDPR; software copyright exceptions; patent infringement defences; freedom of information request; algorithmic transparency; algorithmic accountability; algorithmic governance; Data Protection Act 2018

© 2018 Guido Noto La Diega

Everybody may disseminate this article by electronic means and make it available for download under the terms and conditions of the Digital Peer Publishing Licence (DPPL). A copy of the license text may be obtained at http://nbn-resolving.de/urn:nbn:de:0009-dppl-v3-en8.

Recommended citation: Guido Noto La Diega, Against the Dehumanisation of Decision-Making – Algorithmic Decisions at the Crossroads of Intellectual Property, Data Protection, and Freedom of Information, 9 (2018) JIPITEC 3 para 1.

A. Context and scope of the research

1 This work argues that algorithms cannot and should not replace human beings in decision-making, but it takes account of the increase of algorithmic decisions and, accordingly, it presents three European legal routes available to those affected by such decisions.

2 Algorithms have been used in the legal domain for decades, for instance in order to analyse legislation.[1] These processes or sets of rules followed in calculations or other problem-solving operations raised limited concerns when they merely made our lives easier by ensuring that search engines showed us only relevant results.[2] However, nowadays algorithms can decide if one can get a loan,[3] is hired,[4] is allowed to cross a border,[5] or must go to prison.[6] Particularly striking is the episode concerning a young man sentenced in Wisconsin to a six-year imprisonment for merely attempting to flee a traffic officer and operating a vehicle without its owner's consent. The reason for such a harsh sanction was that Compas, an algorithmic risk assessment system, concluded that he was a threat to the community. The proprietary nature of the algorithm did not allow the defendant to challenge the Compas report. The Supreme Court found no violation of the right to due process.[7]

3 Artificial intelligence techniques (natural language processing, machine learning, etc.) and predictive analytics enable private and public decision-makers to extract value from big data[8] and to build profiles, which are used to make decisions in an automated way. The accuracy of the profiles is further enhanced by the linking capabilities of the Internet of Things.[9] These decisions may profoundly affect people's lives in terms of, for instance, discrimination, de-individualisation, information asymmetries, and social segregation.[10]

4 In light of the confusion as to the actual role of algorithms, it is worrying that in "the models of game theory, decision theory, artificial intelligence, and military strategy, the algorithmic rules of rationality replaced the self-critical judgments of reason."[11]

5 One paper[12] concluded by asking whether and how algorithms should be regulated. This work aims to constitute an attempt to answer those questions with a focus on the existing rules on intellectual property, data protection, and freedom of information. In particular, it will be critically assessed whether "the tools currently available to policymakers, legislators, and courts (which) were developed to oversee human decision-makers (...) fail when applied to computers instead."[13]

6 First, the paper presents ten arguments why algorithms cannot and should not replace human decision-makers. After this, three legal routes are presented.[14] The General Data Protection Regulation

* Lecturer in Law (Northumbria University); Director (Ital-IoT Centre for Multidisciplinary Research on the Internet of Things); Fellow (Nexa Center for Internet & Society).

1 William Adam Wilson, 'The Complexity of Statutes' (1974) 37 Mod L Rev 497.

2 The algorithm used by Google to rank search results is covered by a trade secret.

3 More generally, on the use of algorithms to determine the parties' contractual obligations, see Lauren Henry Scholz, 'Algorithmic Contracts' (*SSRN*, 1 October 2016), <https://ssrn.com/abstract=2747701> accessed 1 March 2018.

4 On the negative spirals that automated scoring systems can create, to the point of making people unemployable, see Danielle Keats Citron and Frank Pasquale, 'The scored society: Due process for automated predictions' (2014) 89(1) Washington Law Review 1, 33.

5 Jose Sanchez del Rio et al., 'Automated border control e-gates and facial recognition systems' (2016) 62 Computers & Security 49.

6 As written by Frank Pasquale, 'Secret algorithms threaten the rule of law' (*MIT Technology Review*, 1 June 2017) <https://www.technologyreview.com/s/608011/secret-algorithms-threaten-the-rule-of-law/> accessed 1 March 2018, imprisoning people "because of the inexplicable, unchallengeable judgements of a secret computer program undermines our legal system". For a files $10 million lawsuit related to face-matching technology that allegedly ruined an American man's life see Allee Manning, 'A False Facial Recognition Match Cost This Man Everything' (*Vocativ*, 1 May 2017) <http://www.vocativ.com/418052/false-facial-recognition-cost-denver-steve-talley-everything/> accessed 1 March 2018.

7 *State v Loomis*, 881 N.W.2d 749 (Wis. 2016). Cf Adam Liptak, 'Sent to Prison by a Software Program's Secret Algorithms' (*New York Times*, 1 May 2017), <https://www.nytimes.com/2017/05/01/us/politics/sent-to-prison-by-a-software-programs-secret-algorithms.html?_r=0> accessed 1 March 2018.

8 In analysing the algorithms used by social networks, Yoan Hermstrüwer, 'Contracting Around Privacy: The (Behavioral) Law and Economics of Consent and Big Data' (2017) 8(1) JIPITEC 12, observes that for these "algorithms to allow good predictions about personal traits and behaviors, the network operator needs two things: sound knowledge about the social graph [describing the social ties between users] and large amounts of data."

9 Article 29 Working Party, 'Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679' (2017) 17/EN WP 251.

10 See Bart W. Schermer, 'The limits of privacy in automated profiling and data mining' (2011) 27 Computer law & security review 45, 52, and Article 29 Working Party (n 9) 5.

11 Lorraine Daston, 'How Reason Became Rationality' (*Max-Planck-Institut für Wissenschaftsgeschichte*, 2013) <https://www.mpiwg-berlin.mpg.de/en/research/projects/DeptII_Daston_Reason> accessed 1 March 2018.

12 Solon Barocas et al., 'Governing Algorithms: A Provocation Piece' (*SSRN*, 4 April 2013) 9 <https://ssrn.com/abstract=2245322> accessed 1 March 2018.

13 Joshua A. Kroll et al., 'Accountable Algorithms' (2017) 165 U Pa L Rev. 633.

14 Other routes may be explored. In the US, Keats Citron (n 4) 33 suggested that the principles of due process may constitute

(GDPR)[15] bans solely automated decisions having legal effects on the data subject's life "or similarly significantly affects him or her."[16] However, when such decisions are allowed, the data controller shall ensure the transparency of the decision, and give the data subject the rights to obtain human intervention, to express their point of view, as well as to contest the decision. Data protection is the most studied perspective but invoking it by itself is a strategy that "is no longer viable."[17] Therefore, this paper approaches this issue by integrating data protection, intellectual property, and freedom of information.

7 As to the intellectual property route, some copyright and patent exceptions may allow the access to a computer program implementing an algorithm, notwithstanding its proprietary nature.

8 In turn, when it comes to the freedom of information, an Italian court stated that an algorithm is a digital administrative act and therefore, under the freedom of information regime, the citizens have the right to access it.[18]

9 In terms of method, the main focus is a desk-based research of EU laws, and of the UK and Italian implementations. The paper is both positive and normative. Whilst advocating against algorithmic decision-making, this research adopts a pragmatic approach whereby one should take into account that the replacement of human decision-makers with algorithms is already happening. Therefore, it is important to understand how to solve the relevant legal issues using existing laws. If algorithms are becoming "weapons of math destruction,"[19] it is crucial that awareness is raised regarding the pervasivity of algorithmic decision-making and that light is shed on the existing legal tools, in anticipation of better regulations and more responsible modelers. Without clarity on the nature of the phenomenon and the relevant legal tools, it is unlikely that citizens will trust algorithms.

a sufficient answer against algorithmic decisions (in particular, against automated scoring systems). The authors recommend that the Federal Trade Commission interrogate scoring systems under their unfairness authority.

15 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 [2016] OJ 119/1.

16 GDPR, art 22.

17 Schermer (n 10) 52.

18 TAR Lazio, chamber III *bis*, 22 March 2017, No 3769.

19 Cathy O'Neil, *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy* (Crown 2016).

B. Positive and normative arguments against algorithms as a replacement for human decision-makers

10 The first part of this section is dedicated to presenting the main reasons why algorithms cannot replace human decision-makers. The second part discusses the reasons why such a replacement is not desirable. The analysis is carried out with the judge as the model of a decision-maker.

I. The unfeasibility of the replacement

11 The untenability of the replacement is mainly related to the role and characteristics of legal interpretation. Algorithms could replace human decision-makers if interpretation were a straightforward mechanical operation of textual analysis; where the meaning is easily found by putting together the facts and the norms. The said model of interpretation, which seems flawed, is accompanied by the conviction that there is a clear distinction, on the one hand, between interpretation and application and, on the other hand, between easy cases and hard cases. However, legal interpretation seems to have the opposite characteristics. Indeed, it is ubiquitous[20] and its extreme complexity relates to several factors,[21] such as the psychological (and not merely cognitive) nature of the process.[22] This highlights

20 Given the features of legal interpretation in practice, the brocard *in claris non fit interpretatio* should be replaced by *in claris fit interpretatio* (*cf* Francesco Galgano, *Tutto il rovescio del diritto* (Giuffrè 2007) 100, who points out how the attempts to rule out legal interpretation by means of clear statutes (Carlo Ludovico Muratori) or through proposals to expressly prevent judges from interpreting the statutes (Pietro Verri) today would be laughable. *cf* Vittorio Villa, *Una teoria pragmaticamente orientata dell'interpretazione giuridica* (Giappichelli 2012).

21 For instance, due to the intrinsic vagueness of the legal language and because of the importance of general principles, one of the main tasks of judicial interpretation is striking a balance between conflicting interest, which shall be done on a case-by-case basis. However, some scholars believe that "balancing works with mathematical rules" (Pier Luigi M. Capotuorto, 'Arithmetical Rules for the Judicial Balancing of Conflicts between Constitutional Principles: From the 'Weight Formula' to the Computer-Aided Judicial Decision' (2007) 3(2) Rivista di Diritto, Economia e Gestione delle Nuove Tecnologie 171.

22 Richard A Posner, 'The Role of the Judge in the Twenty-First Century' (2006) 86 B U L Rev 1049, 1060, believes that the psychological component is dominant when it comes to the sources of ideology, which plays a fundamental role in the decisions of all judges. Works on the prediction of judicial decisions usually focus on non-textual elements such as the nature and the gravity of the crime or the preferred policy position of each judge. See e.g. Benjamin E Lauderdale and

why it is currently impossible to develop an algorithm capable of interpreting the law as a human judge would do.[23] The high degree of discretion of the relevant process seems to be the main reason for the impossibility of the replacement. Dworkin's view whereby there is only one right answer to legal questions[24] has very few defenders indeed.[25] Hart[26] clearly proved his doctrine of strong discretion in judicial interpretation, as "a necessary byproduct of the inherent indeterminacy of social guidance."[27] A factor that increases the hermeneutical discretion is that interpreting and applying the law requires value judgements and choices, which are very hard to formalise and compute because of their indeterminacy.[28] One may object that AI may replace humans at least in the legal interpretation of easy cases (for instance, because there is a robust body of case law on the exact issue at hand). However, it has been shown that it is impossible to determine *ex ante* whether a case is easy or difficult: the complexity of the legal experience tells us that the factual and normative circumstances make a case easy or difficult. The similar suggestion to limit the use of algorithms to the application of the law is based, finally, on the wrong assumption that there is an interpretation-application dichotomy and that there is no room for interpretation when one applies the law. Conversely, application seems the last (and most important) phase of the interpretive process.[29]

12 Even leaving the philosophy of law aside, the actual development of statutory interpretation shows the increasing discretion of this activity. Indeed, it seems clear that nowadays the literal rule of interpretation plays a small and often rhetoric role, whereas a purposive approach to statutory interpretation has become commonplace,[30] in part as a consequence of the EU's influence. It has been noted that, whatever the philosophical view one adopts, the discretional power of courts is never expressed in a pure mechanical operation.[31] A good example of the new face of legal interpretation is provided by the case of the Psychoactive Substances Act 2016.[32] The parliamentary debate[33] clearly shows that the intention of the legislator was to ban the so-called poppers (of the class 'alkyl nitrites'), a recreational drug used traditionally by men who have sex with men due to its effects on the relaxation of muscles (including the sphincter). The broad definition of psychoactive substance seemed to allow the interpretation whereby poppers were banned[34] and some law enforcement agencies applied it consistently.[35] However, the final result is that

Tom S Clark, 'The Supreme Court's many median justices' (2012) 106(4) American Political Science Review 847.

23 There are several studies in the field of AI & Law that develop models to explain the legal reasoning, but this is an ex-post operation, as opposed to a genuinely predictive one. See, for instance, Latifa Al-Abdulkarim et al., 'A methodology for designing systems to reason with legal cases using Abstract Dialectical Frameworks' (2016) 24(1) Artif Intell Law 1.

24 See, for instance, Ronald Dworkin, 'No Right Answer?' (1978) 53 New York University Law Review 1; Ronald Dworkin, *Law, Morality, and Society* (Peter Hacker & Joseph Raz eds, Clarendon Press 1977).

25 For instance, an author like Thomas B. Colby, a strong assertor of the rule of law, recognises that the law is often ambiguous or open-ended and, therefore, "there is no objectively correct answer that can be discerned simply by calling balls and strikes." (Thomas B. Colby, 'In Defense of Judicial Empathy' (2012) 96 Minn. L. Rev. 1944, 2015) Even those who argue for an overcoming of the centrality of the Hart-Dworkin debate cannot "envision a jurisprudential future without Hart's masterful work at its center" (Brian Leiter, 'Beyond the Hart-Dworkin Debate: The methodology problem in jurisprudence' (2003) 48 Am. J. Juris. 17, 18). For a recent critique to Dworkin, see Aulis Aarnio, 'One right answer?' [2011] Essays on the doctrinal study of law 165. As suggested by Tony Ward in his comments on a previous draft of this paper, one should note that, even in the event that Dworkin were right, it is unclear how the Hercules algorithm would be programmed.

26 See H. L. A Hart, *The Concept of Law* (Penelope Bulloch & Joseph Raz eds, Clarendon Press 1994) 123.

27 Scott J. Shapiro, 'The "Hart-Dworkin" debate: A short guide for the perplexed' (2007) Public Law and Legal Theory Working Paper Series No. 77, 16.

28 It has been noted that "even if computers were technically able to mimic legal decision making in a mechanical fashion they would necessarily miss the subtle institutional, value-based, experiential, justice-oriented, and public policy dimensions that are the heart of lawyerly analysis" (Lisa A. Shay et al., 'Do robots dream of electric law? An experiment in the law as algorithm' in Ryan Calo, A. Michael Froomkin, and Ian Kerr (eds) *Robot Law* (Elgar 2016) 274, 277, citing Harry Surden, 'Computable Contracts' (2012) 46 U.C. Davis L. Rev. 629).

29 Francesco Viola, 'Interpretazione e indeterminatezza della regola giuridica' (2002) 7-8 Diritto privato 49, 51, explains in this way one of the differences between Hart and Kelsen (the *ex-ante* distinction between easy and difficult cases – and between interpretation and application – would be possible adopting a Kelsenian perspective).

30 Catherine Elliott & Frances Quinn, *English Legal System* (17th ed, Pearson 2016) 61. The shift is very significant, and it is suggested already from the use of 'approach' instead of 'rule', which hints at a more flexible strategy drawing from several sources and taking into account many factors, rather the mechanical operation of subsuming a fact under a rule.

31 Pier Luigi M. Lucatuorto, 'Modelli computazionali della discrezionalita del giudice: uno studio preliminare' (2006) 7(3) Ciberspazio e diritto 1, 2.

32 I am thankful to Chris Ashford for the insight provided in his "The UK Poppers 'Ban' and the Psychoactive Substances Act 2016: New Legal Frontiers in the Homonormative Imagination" (Northumbria University Gender, Sexuality and Law Research Seminar, Newcastle upon Tyne, 14 June 2017).

33 The Burnham amendment and the Scottish National Party (SNP) one, aimed to allow poppers, were rejected.

34 Psychoactive Substances Act 2016, s 2.

35 See Steven Hopkins, 'Crawley Police Forced To Apologise After Wrongly Seising Poppers After Legal High Ban Came Into Effect' (*Huffington Post UK*, 26 May 2016) <http://www.huffingtonpost.co.uk/entry/crawley-police-forced-to-

poppers *are not* banned, because the UK Advisory Council on the Misuse of Drugs explained that since poppers have a merely indirect effect on the nervous system, they do not technically qualify as psychoactive substances and, therefore, fall outside the scope of the Act.[36] Finally, sectoral empirical studies[37] are showing that algorithms cannot cope with legal interpretation in a satisfactory way. For instance, it has been shown[38] that algorithms often reflect a wrong interpretation of the law they enforce,[39] in particular with regards to the fair use analysis in online infringement cases.[40] These are just a couple of examples of how interpretation is discretionary, ubiquitous, complex, and unpredictable.[41] Therefore, it seems that it is currently impossible to design an interpretive algorithm.

13 This study itself confirms this view, in as much as from an apparently simple provision, such as Article 22 of the GDPR, stem a number of complicated interpretative problems for which there is no easy answer. The relevant difficulties will be explained in section 4 below. Here suffice to say that there is a meta-problem. Even if algorithms could perfectly replace human decision-makers, arguably it would not be fair to let them interpret a provision – Article 22 – which has the aim of protecting citizens from algorithmic decisions.

14 The above considerations regard the current progress in algorithms-related technologies. However, AI's growth is exponential, therefore the considerations above may prove to be wrong soon, especially in fields where the issues arising are often similar and there is a lot of precedent. Less so where there is no established case law, and/or the field is fast evolving.[42] For example, predicting the outcome of succession cases involving only land may prove easier than cyber law cases with cross-border elements. That said, alongside the technologies, the scholarship is evolving. Recently, the first systematic study on predicting the outcome of cases tried by the European Court of Human Rights based solely on textual content was presented.[43] The model is quite accurate, being able to predict the outcome in 79% of cases. However, there are some considerations to be made. Especially in matters as important as human rights, reaching a wrong decision in 21% of the cases would be utterly unacceptable. Secondly, the reasons for this margin of error should be better analysed; they might stem from the fact that interpretation is not a mere mechanical operation of text analysis. Thirdly, the authors themselves point out that the model would not be a substitute for the human decision-maker, because its role would rather be an "assisting tool"[44] to identify cases and extract patterns. Lastly, the study did not predict the outcome using the documents filed by the applicants, but only analysing the published rulings. This means that a human judge had already selected the materials and interpreted them, which affects the results of the study.[45] More generally, it still holds true that "[j]ustification, persuasion and discretion are the main limits of the Artificial Intelligence application in Law."[46]

15 Second, human learning is much more complex than machine learning. According to the seminal *Mind over Machine*,[47] the characteristics of human learning would explain why prophecies about real machine intelligence have all been proven wrong,[48] and why

apologise-after-wrongly-seising-poppers-after-legal-high-ban-came-into-effect_uk_57472a41e4b0ebf6a3297cac> accessed 1 March 2018.

36 Advisory Council on the Misuse of Drugs, 'CMD review of alkyl nitrites (poppers)' (*The UK Government,* 16 March 2016) <https://www.gov.uk/government/publications/acmd-review-of-alkyl-nitrites-poppers> accessed 1 March 2018.

37 Along with the studies cited in the following footnotes, see, for instance, Joe Karaganis & Jennifer Urban, 'The rise of the robo notice' (2015) 58(9) Communications of the ACM 28.

38 Maayan Perel & Niva Elkin-Koren, 'Black Box Tinkering: Beyond Disclosure in Algorithmic Enforcement' (2017) 69 Fla. L. Rev. 181, 210.

39 Kenneth A. Bamberger, 'Technologies of Compliance: Risk and Regulation in a Digital Age' (2010) 88 Tex. L. Rev. 669, 675-6.

40 Specifically, when the researchers tried to upload a 48 seconds homemade video of a child dancing a protected song by Justin Bieber, 25% of video-sharing platforms removed the video, notwithstanding the fact that it clearly constituted a fair use.

41 Reed C. Lawlor, 'What Computers Can Do: Analysis and Prediction of Judicial Decisions' (1963) 49(4) American Bar Association Journal 337, conjectured that in the future machines would be able to predict the outcomes of judicial decisions. The prophecy was not fulfilled (yet), but Nikolaos Aletras et al., 'Predicting judicial decisions of the European Court of Human Rights: a Natural Language Processing perspective' (2016) 2:e93 PeerJ Comput. Sci., constitute a progress.

42 However, as said above, it is not possible to assess *ex ante facto* whether a case is easy or hard (and even *ex post facto*, the lines are blurred and interpretation is needed for the hard cases as well as for the easy cases).

43 Aletras (n 41).

44 ibid 3.

45 ibid 2, assume that "the text extracted from published judgments of the Court bears a sufficient number of similarities with, and can therefore stand as a (crude) proxy for, applications lodged with the Court as well as for briefs submitted by parties in pending cases". They accept, however, that "full acceptance of that reasonable assumption necessitates more empirical corroboration".

46 Pier Luigi M. Lucatuorto, 'Computer-Aided Sentencing: Computer Science and Legal Aspects: The Chinese Case' (2006) 2(4) Rivista di Diritto, Economia e Gestione delle Nuove Tecnologie 388.

47 Hubert L Dreyfus and Stuart E Dreyfus, *Mind over Machine: The Power of Human Intuition and Expertise in the Era of the Computer* (Free Press 1988).

48 For instance, Herbert A Simon, *The Shape of Automation for*

small scale successful experiments conducted in laboratories were not as successful once extended to larger systems and the real world. In particular, machines will not be able to replace human beings when cognitive tasks require intuition and holistic thinking.[49] By presenting a five-stage model of acquisition of expertise (novice, advanced beginner, competent, proficient, and expert), these authors show that there is more to human intelligence than the computer's calculative rationality. Only the human brain, at least currently, is capable to properly learn and understand through holistic intuition a world that is – unlike the laboratory – incomplete, imprecise, and unreliable. It seems, indeed, unlikely that training a machine with millions of legal provisions and case law can lead to the same results to the learning of a judge, who is immersed in the real world and learns in ways, which cannot be coded.

II. Eight arguments against the desirability of algorithms replacing human decision-makers

16 Let us assume that the thesis of this paper is wrong. Let us say, for the sake of argument, that either interpretation is not ubiquitous, or it is not discretionary (or that algorithms can cope well with strongly discretionary processes). Let us posit, then, that algorithms can learn in the same way as the humans. Nonetheless, there are at least eight reasons why they *should not* replace human decision-makers. Two reasons refer to why one should not trust algorithms. Six arguments are, in turn, presented to show why we should trust humans.

1. The replacement is undesirable because there are good reasons not to trust the algorithms

17 Let us start with what is not to like in algorithms. One of the strong arguments in favour of the algorithms is that they are more reliable than human beings are. However, there is evidence that algorithms can make mistakes and, when they do so, the effects are on a larger scale than an error made by a human judge in a ruling.[50] More importantly, algorithms are not more reliable than human beings, because of the emerging problem of algorithmic (or machine) bias.[51] The founder of the Algorithmic Justice League, for instance, stated that a facial recognition machine could not see her because she is black and, probably, the machine learning algorithm was trained only using white faces.[52] Contrary to popular belief, algorithms do not eliminate bias, because the relevant models are opaque, unregulated, and incontestable.[53] Even those who believe that AI should be used (in combination with law and self-regulation) for the governance of the Internet, admit that the "[l]ack of transparency on how algorithms operate is a real issue, as well as the problem that artificial intelligence tends to share the biases of the humans it learns from."[54]

18 In the context of the UK inquiry on algorithms in decision-making,[55] six reasons why algorithmic systems can produce biased outcomes have been presented.[56] First, design choices make the decision-making process or the factors it considers too opaque; these choices may also limit the control of the designer.[57] Second, the output of the system may

Men and Management (Harper & Row1965) 38, foresaw that in 1985 machines would have been capable of doing any work that a man could do. In hindsight, that prediction was not entirely accurate.

49 Computer "reasoning" is deemed to be ontologically different to human know-how: "a far superior holistic, intuitive way of approaching problems that cannot be imitated by rule-following computers" (Dreyfus (n 47) 193). For some recent developments in intuition modelling, see Ulrich Hoffrage and Julian N Marewski, 'Unveiling the Lady in Black: Modeling and aiding intuition' (2015) 4 Journal of Applied Research in Memory and Cognition 145.

50 One need only think of the wrong calculations that affected 20.000 divorced couples due to a software glitch (see, e.g. Will Grice, 'Divorce error on form caused by UK Government software glitch could affect 20,000 people' (*The Independent*, 18 December 2015) <http://www.independent.co.uk/news/uk/home-news/ministry-of-justice-software-glitch-could-see-thousands-revisiting-painful-divorce-settlements-a6777851.html> accessed 1 March 2018.

51 Cf. Megan Garcia, 'Racist in the Machine: The Disturbing Implications of Algorithmic Bias' (2016) 33(4) World Policy Journal 111, and, more generally, Kroll (n 13) 633.

52 <http://www.ajlunited.org/> accessed 1 March 2018. These kind of problems had already been evidenced by Brandan F. Klare et al., 'Face Recognition Performance: Role of Demographic Information' (2012) 7(6) IEEE Transactions on Information Forensics and Security 1789.

53 This is one of main ideas of O'Neil (n 19).

54 Andrés Guadamuz, 'Whatever happened to our dream of an empowering Internet (and how to get it back)' (*TechnoLlama*, 5 June 2017), <http://www.technollama.co.uk/whatever-happened-to-our-dream-of-an-empowering-internet-and-how-to-get-it-back> accessed 1 March 2018. On the phenomenon of machine bias (or algorithmic bias) see below.

55 <https://www.parliament.uk/business/committees/committees-a-z/commons-select/science-and-technology-committee/inquiries/parliament-2015/inquiry9/> accessed 1 March 2018.

56 Science and Technology Committee, 'Written evidence submitted by Dr Alison Powell (ALG0067)' (*UK Parliament*, 2017) <http://data.parliament.uk/writtenevidence/committeeevidence.svc/evidencedocument/science-and-technology-committee/algorithms-in-decisionmaking/written/69121.html> accessed 1 March 2018.

57 Alan Dix, 'Human issues in the use of pattern recognition techniques', in R Beale and J Finlay (eds), *Neural Networks*

be affected by the biases in data collection.[58] Third, unlike human beings, algorithms cannot balance biases in interpretation of data by a conscious attention to the redress of the bias.[59] Fourth, there are biases in the ways that learning algorithms are tuned based on the testing users' behaviour.[60] Fifth, algorithms may be designed for a purpose, but then inserted into systems designed for other purposes.[61] Lastly, as already said with regard to the Algorithmic Justice League, another factor is the biases in the data used to train the decision-making systems.[62]

19 Algorithmic bias is the main problem regarding automated decision-making with legal effects.[63] It has been submitted that "while persistent inequities stem from a complex set of factors, digitally automated systems may be adding to these problems in new ways."[64] It is arguable that even if the automated decision (e.g. a ruling) is biased, the move to algorithms "may at least have the salutary effect of making bias more evident."[65] Algorithmic bias is dealt with in a recital of the GDPR,[66] in a way which is not entirely satisfactory. Indeed, the GDPR calls on the data controller to "use appropriate mathematical or statistical procedures for the profiling, implement technical and organisational measures appropriate to ensure, in particular, that factors which result in inaccuracies in personal data are corrected and the risk of errors is minimised, secure personal data in a manner that takes account of the potential risks involved for the interests and rights of the data subject and that prevents, inter alia, discriminatory effects"[67] on the basis of sensitive data. Now, it would seem that the GDPR's focus is misplaced. The point with discrimination is not only that the data are inaccurate or that they are not secure. The main problem is that these data should never be used to discriminate in the first place,[68] regardless of their being accurate or not, or that independency constraints should be put in place.[69]

20 The second argument revolves around transparency. Indeed, making bias evident would mean ensuring transparency, which seems a chimera for a number of reasons, including the fact that the more accurate an algorithm is, the less transparent.[70] The trade-off accuracy vs. transparency is easily explained. On the one hand, modelers tend to develop more accurate models "with increasingly complex, data-mining-based black-box models."[71] On the other hand, model users tend to favour "transparent, interpretable models not only for predictive decision-making but also for after-the-fact auditing and forensic purposes."[72] Against the dominant idea that transparency will solve all the problems, some scholars point out that "[d]isclosure of source code is often neither necessary (because of alternative techniques from computer science) nor sufficient (because of the issues analysing code) to demonstrate the fairness of a process."[73] Arguably, however, such disclosure would be necessary to comply with the right to an effective remedy and to a fair trial under the EU Charter of Fundamental Rights and the European Convention on Human Rights.

21 The lack of transparency is related to the so-called black box (better said, black *boxes*). Arguably, three different black boxes may be distinguished: the organisational; the technical; and the legal one. The organisational black box will not be the subject of specific analysis. Suffice to say that algorithms are mostly implemented by "private, profit-maximising entities, operating under minimal

and Pattern Recognition in Human Computer Interaction (Ellis Horwood 1992) 429.

58 Stella Lowry & Gordon Macpherson, 'A blot on the profession' (1988) 296 British Medical Journal 657.

59 Aylin Caliskan et al., 'Semantics derived automatically from language corpora contain human-like biases' (2017) 356(6334) Science 183.

60 Dix (n 57) 57.

61 Louise Amoore, *The politics of possibility: risk and security beyond probability* (Duke University Press 2013).

62 Klare (n 52).

63 Algorithmic bias has many potential consequences. For instance, in the context of social media, it may lead to the so-called filter bubble. See, e.g., William H. Dutton et al., 'Search and Politics: The Uses and Impacts of Search in Britain, France, Germany, Italy, Poland, Spain, and the United States' (*Quello Center Working Paper No. 5-1-17*, 2017).

64 Seeta Peña Gangadharan et al., *Data and Discrimination: Collected Essays* (Open Technology Institute 2014).

65 Barocas (n 12) 9.

66 GDPR, recital 71.

67 ibid

68 Rakesh Agrawal and Ramakrishnan Srikant, 'Privacy-Preserving Data Mining' (2000) 29(2) ACM SIGMOD Record 439 (2000).

69 Toon Calders et al., 'Building Classifiers with Independency Constraints' (2009) IEEE ICDM Workshop on Domain Driven Data Mining 13. Therefore, for instance, sensitive attributes such as sex shall be included, but the program would be instructed to make predictions independently of the said attributes. This second approach seems preferable for accountability reasons.

70 Barocas (n 12) 9 accept that "algorithms may involve rules of such complexity that they defy attempts to trace their reasoning".

71 Innocent Kamwa et al., 'On the accuracy versus transparency trade-off of data-mining models for fast-response PMU-based catastrophe predictors' (2012) 3(1) IEEE Transactions on smart grid 152.

72 ibid 152. They conclude that "for catastrophe anticipation purposes, we would favor fuzzy logic-based transparent solutions over black box solutions for implementation ease and robustness, as well as for their suitability in the auditing process, even while sacrificing some predictive accuracy" (ibid 160).

73 Kroll (n 13) 633.

transparency obligations."[74] As to the technical black box, artificial intelligence makes the rationale of decisions intrinsically difficult to access. This is particularly evident with the so-called neural networks that, being modelled on the brain, are at least as opaque. One need only imagine a deep-learning neural network which is trained using old mammograms that have been labelled according to which women went on to develop breast cancer.[75] It could help us to make predictions on which breasts are likely to develop cancer, but without knowing the risk factors (the rationale), it is unlikely that the patient would undergo therapy and, more generally, the development of cancer research would not be substantive. The legal black box relates to intellectual property and will be presented in the following section.

22 The lack of transparency has obvious repercussions on the accountability issue. For instance, ensuring fair, lawful, and transparent processing may be difficult "due to the way in which machine learning works and / or the way machine learning is integrated into a broader workflow that might involve the use of data of different origins and reliability, specific interventions by human operators, and the deployment of machine learning products and services."[76] Some technical tools to ensure accountability in algorithmic scenarios have been presented,[77] but they do not seem sufficient to offset the inherent problems in algorithmic decision-making.

2. The replacement is undesirable because there are good reasons to trust the human beings

23 This subsection is dedicated to the reasons why one should trust humans over algorithms and, more generally, over non-human agents.

24 First, human beings tend to emulate the behaviour of the majority of fellow human beings. This should ensure consistency and predictability in societal behaviours. This phenomenon was observed with particular clarity by Solomon Asch, who developed the so-called psychology of conformity.[78] Needless to say that non-human agents do not have a consciousness[79] and, therefore, psychology does not apply to them. One could object, however, that conforming to the majority does not equal pursuing the common good, because it could lead to the oppression of the minorities. However, humans have some built-in safeguards.

25 The argument can be put forward that, despite the different characteristics of human beings, humans tend to act consistently towards the common good. This may be explained with the power of sanctions.[80] Human beings comply with the law not for a natural disposition, but because they do not wish to be sanctioned. However, it is hardly arguable that non-human agents share this fear. Indeed, neither can they be imprisoned (criminal sanctions), nor do they own assets that can be used to execute civil and administrative sanctions.

26 The third argument refers, like the previous one, to the effects of group pressure, but in a different setting. It can be summed up by saying that hypocrisy has a civilising force.[81] Indeed, with regards to the relationship between deliberation and publicity, it has been observed that "the effect of an audience is to replace the language of interest by the language of reason and to replace impartial motives by passionate ones."[82] These considerations, rooted in human psychology, do not apply to non-human agents. Therefore, hypocrisy cannot civilise algorithmic decision-makers.

27 Let us say that it is possible for an algorithm to learn and decide like a human judge. At this point, one

74 Perel & Elkin-Koren (n 38) 181.

75 The scenario, imagined by Andrea Vedaldi (University of Oxford) is referred to by Davide Castelvecchi, 'Can we open the black box of AI?' (*Nature*, 15 October 2016) <http://www.nature.com/news/can-we-open-the-black-box-of-ai-1.20731> accessed 1 March 2018. *cf* Krzysztof J Geras, Stacey Wolfson, Yiqiu Shen, S. Gene Kim, Linda Moy, and Kyunghyun Cho, 'High-Resolution Breast Cancer Screening with Multi-View Deep Convolutional Neural Networks' (2017) arXiv:1703.07047 [cs.CV].

76 Dimitra Kamarinou et al., 'Machine Learning with Personal Data' (*Queen Mary School of Law Legal Studies Research Paper No. 247*, 2016) 22. Christopher Kuner et al., 'Machine learning with personal data: is data protection law smart enough to meet the challenge?' (2017) 7(1) International Data Privacy Law 1, observe that "[m]achine learning is data driven, typically involving both existing data sets and live data streams in complex training and deployment workflow [therefore it] may be difficult to reconcile such dynamic processes with purposes that are specified narrowly in advance."

77 Kroll (n 13) 633.

78 See Solomon E Asch, 'Effects of group pressure upon the modification and distortion of judgment', in H Guetzkow (ed), *Groups, leadership and men* (Carnegie Press 1951) 222.

79 However, there is a significant debate about artificial consciousness, whose functions have been described as including awareness of self, will, instinct, and emotion (Igor Aleksander, 'Machine consciousness' (2008) 3(2) Scholarpedia 4162). It seems that the prevalent position in the literature is against the existence of a proper artificial consciousness.

80 *cf* Antonio Pagliaro, 'Sanzione. Sanzione penale' [1992] 28 Enciclopedia giuridica 3; David R Carp, 'The Judicial and Judicious Use of Shame Penalties' (1998) 44(2) Crime & Delinquency 277.

81 Jon Elster (ed), *Deliberative Democracy* (Cambridge University Press 1998).

82 ibid 111.

may argue, it would be sufficient to find the best judge in the world and create a large number of non-human clones that will gradually replace all human judges. However, this scenario raises some issues. Pluralism seems to be the main one.[83] Indeed, if pluralism is rooted in the respect for the minorities and in the belief that a multiplicity of viewpoints enriches the understanding of the world, then erasing this by cloning the perfect judge would at least be problematic. Even before that, how does one find the perfect judge to clone? What does it mean to be the best judge? Is it possible to entirely eliminate human bias?[84]

28 A fifth reason why this paper takes a humanist stance is empathy, which is the "cognitive ability to understand a situation from the perspective of other people, combined with the emotional capacity to comprehend and feel those people's emotions in that situation."[85] This could come as a surprise, since usually empathy is seen as a bias[86] and, therefore, as an argument in favour of non-human agents. Conversely, empathy is "a requirement of judicial neutrality."[87] It has been shown that arguments in favour of judicial empathy are rooted, perhaps unexpectedly, in "a firm commitment to the rule of law and a deep-seated appreciation of—rather than rejection of— legal doctrine."[88] A recent study shed light on the shortcomings of the anti-empathic consensus; indeed, it descends of XIX century formalism, but it has "drifted from its source such that it would almost certainly be condemned by the very formalist scholars from whom it is descended."[89] Not only is empathy not a defect in human decision-making, it serves a positive function. This is required by the paramount function of concepts such as reasonableness and balancing tests.[90] More generally, it can be argued that empathy is the way justice (as opposed to law) enters the decision. When Cicero wrote *"summum ius summa iniuria"*[91] he meant that the mechanical application of the law leads to unjust results. Empathy tempers legalistic excesses and algorithms are not capable of it.

29 Lastly, one needs to choose between democracy and technocracy. In a democratic context, laws are the product of a debate between politicians. This debate is public, and the politicians are democratically elected and accountable both politically and legally. Human judges are either democratically elected or receive specific legal training. Conversely, algorithmic law (as in Lessig's "code is law"[92]) is more problematic. Indeed, "software development, even open source, is opaque, and concentrated in a small programming community, many of whom are employed by few oligopolistic corporations directly accountable to no external party."[93] Algorithms could be suitable to apply algorithmic laws, but given the said characteristics, it is hoped that their role and scope remains limited.

30 For the reasons above, the replacement of algorithms to human beings seems both unfeasible and undesirable.

C. Intellectual property rights: more a problem, than a solution

31 Even though there are good reasons to believe that algorithms cannot and should not replace human decision-makers, it is becoming obvious that the replacement is already taking place, regardless of the relevant pitfalls. Therefore, a lawyer should be able to provide a sufficiently clear answer to a client subject to an algorithmic decision.

32 There are at least three routes that can be taken, should the relevant requirements be met. In this section, the focus will be on intellectual property and the relevant exceptions that may enable access to a computer program implementing an algorithm, or the relevant invention, notwithstanding its proprietary nature. The features of the analysed exceptions made scholars talk of "the advent of a more active approach to copyright exceptions,"[94] which creates quasi-rights, "legal hybrids between exceptions and rights."[95] This must be taken into account when interpreting the relevant provisions

83 When asked about this argument during the conferences cited in the acknowledgments, the audience also mentioned other negative repercussions. The most relevant one seems to be the lack of legal innovation deriving from a single approach to decision-making.

84 As to the last question, it is submitted that if one eliminates ideologies in the attempt of eliminating bias, the output would be a useless algorithm, incapable of deciding. Indeed, ideologies guide human judges in deciding, for instance, whether intellectual property rights should prevail over access to knowledge, whether the reasons of privacy should take precedence over those of free speech, etc.

85 Colby (n 25) 1945.

86 See, e.g., Adam N Glynn and Maya Sen, 'Identifying Judicial Empathy: Does Having Daughters Cause Judges to Rule for Women's Issues?' (2015) 59(1) American Journal of Political Science 37.

87 Colby (n 25) 2015.

88 ibid 1946.

89 Brenner Fissel, 'Modern critiques of judicial empathy: A revised intellectual history' (2016) Mich. St. L. Rev 817.

90 Colby (n 25) 1946.

91 Cicero, *De officiis*, I, 10, 33.

92 Lawrence Lessig, *Code* (v.2.0, Basic Books 2006).

93 Kieron O'Hara, 'Smart contracts – Dumb idea' (2017) The Digital Citizen 2, 5.

94 Tatiana-Eleni Synodinou, 'The lawful user and a balancing of interests in European copyright law' (2010) 41 IIC 819, 826.

95 ibid 826.

and striking a balance with the restricted acts. Equally, defences to patent infringement will be dealt with, although there is not enough evidence to claim their nature as quasi-rights.

33 A major issue is understanding the rationale of algorithmic decisions. This is made difficult by the so-called black boxes. The organisational black box and the technical one have been presented above. The legal black box remains to be analysed. This depends primarily on the (ab)use of intellectual property rights (trade secrets, database rights, etc.) and the kindred rights that companies are collecting on the users' data, that do not fit easily in the traditional intellectual property categories and are leading to the datafication of the digital economy. Along the same lines, it has been noted that "data, originating from users, from devices, sent through the 4G and 5G networks to the client servers and the Cloud are heavily boxed in by intellectual property rights."[96]

34 Even though there are many open-source machine learning frameworks (e.g. Apache Singa, Shogun, and TensorFlow), most AI algorithms are proprietary (Google search and Facebook news feed are the classical examples) i.e. covered primarily by trade secrets,[97] which is the "most common form of protection used by business."[98] Under the new Trade Secrets Directive,[99] algorithms can be covered by trade secrets because they are not generally known or easily accessible and they have commercial value.[100] This is true as long as the person who has control of the algorithm takes steps to keep it secret.[101] The general rule is that the unauthorised acquisition, use, or disclosure of algorithms covered by trade secrets is unlawful.[102] However, the acquisition shall be lawful in a limited number of circumstances, the most relevant of which seems to be the "observation, study, disassembly or testing of a product or object that has been made available to the public or that is lawfully in the possession of the acquirer of the information who is free from any legally valid duty to limit the acquisition of the trade secret."[103] This appears to be a reference to one of the permitted uses of computer programs under the Software Directive.[104] There is a potential contrast between the two regimes. To say that the acquisition is legal only if "free from any legally valid duty to limit [it],"[105] may be construed as meaning that if the owner of the algorithm contractually restricts the said exception, then no observation, study, disassembling, or testing of the algorithm would be allowed. However, under the Software Directive, there is a right to "observe, study or test the functioning of the program in order to determine the ideas and principles which underlie any element of the program."[106] This Directive goes on pointing out that any contractual provisions contrary to said exception "shall be null and void."[107] In the UK, the Copyright, Designs and Patents Act 1988 is clear where it provides that "it is irrelevant whether or not there exists any term or condition in an agreement which purports to prohibit or restrict the act (such terms being [...] void)."[108] The leading case on the matter is *SAS Institute v World Programming*, where it was found that copyright owners cannot restrict the purposes for which the analysed permitted acts are carried out. Additionally, even though only lawful users can avail themselves of the defence, these are not limited to those who click through the licence.[109]

96 Bjorn Lundqvist, 'Big Data, Open Data, Privacy Regulations, Intellectual Property and Competition Law in an Internet of Things World' (Faculty of Law, University of Stockholm Research Paper No. 1/2016) 10.

97 This has the potential to impact many fundamental rights, such as the one of access to public information. For instance, crashes such as the one that, on 6 May 2010, caused the Dow Jones Industrial Average to drop by 9% thus burning millions of dollars, cannot be explained "not least because many of the algorithms involved are proprietary" (Scholz (n 3) 103).

98 James Pooley, 'Trade secrets: The other IP right' (2013) 3 WIPO Magazine.

99 Directive (EU) 2016/943 of the European Parliament and of the Council of 8 June 2016 on the protection of undisclosed know-how and business information (trade secrets) against their unlawful acquisition, use and disclosure (Trade Secrets Directive). Member States shall transpose it by 9 June 2018. By that date, the UK will still be part of the EU, but prima facie there are not significant differences between the rules on the breach of confidentiality and the new EU regime.

100 Trade Secrets Directive, art 2(1)(a)-(b).

101 Trade Secrets Directive, art 2(1)(c).

102 Trade Secrets Directive, art 4.

103 Trade Secrets Directive, art 3(1)(b).

104 Directive 2009/24/EC of the European Parliament and of the Council of 23 April 2009 on the legal protection of computer programs (Software Directive).

105 Trade Secrets Directive, art 3(1)(b).

106 Software Directive, art 5(3).

107 Software Directive, art 8(2). See also recital 16.

108 Copyright, Designs and Patents Act 1988, ss 50BA and 296A. The Copyright (Computer Programs) Regulations 1992, SI 1992/3233 inserted s 50B in the Copyrights, Designs and Patents Act 1998, allowing lawful users of computer programs to decompile programs to achieve interoperability. In turn, the Copyright and Related Rights Regulations 2003 introduced s 50BA, regarding the exception on observing, studying and testing computer programs.

109 *SAS Institute Inc v World Programming Ltd III* [2013] EWHC 69 (Ch), [60]-[61]. In that case, employees who did not click through the licence had observed and studied the computer programme without infringing copyright because the colleague who acquired the programme was operating on behalf of the employer, which was a legal person. It was deemed immaterial that the licence openly restricted the use to the person who clicked through the licence. At a closer look, the distinction between licensed employee and unlicensed ones. Indeed, art 9(1) of the Software Directive renders null and void any contractual restrictions to the exceptions and "this includes a contractual restriction on the employees by whom a legal person in the position of

35 Moreover, the Trade Secrets Directive itself recognises the legality of the acquisition, use or disclosure of trade secrets for purposes of freedom of expression and information.[110] Arguably, there is not an actual conflict here. As an example, let us imagine one buys an Amazon Echo. Under one of the several contracts that one has to accept, one agrees that "all Confidential Information will remain [Amazon's] exclusive property"[111] and one may not "reverse engineer, decompile, or disassemble"[112] the Alexa[113] Service or the Alexa Materials.[114] Under the Trade Secrets Directive this section would be enforceable; however, since the Software Directive, being a *lex specialis,* will prevail the section would be unenforceable.[115] Indeed, the conflict is merely ostensible.

36 In the event that trade secrets were deemed to prevail over the exceptions provided by the Software Directive, it may be worth it to take account of the relevant defences. The most relevant and flexible defence seems the public interest one. It has been stated that "the right of confidentiality, whether or not founded in contract, is not absolute. That right must give way where it is in the public interest that the confidential information shall be made public."[116] It is noteworthy that the disclosure may be seen as in the public interest if there has been non-compliance with a legal obligation.[117] One may argue that the circumvention of the Software Directive consisting in secreting an algorithm in an absolute way falls within this scenario. However, the defendant in the relevant infringement proceedings would need to prove that the disclosure be in the public interest and not merely interesting to the public, which may be difficult.[118] Unfortunately, the Trade Secrets Directive does not leave much room for the public interest or other defences. However, it recognises that the Directive shall not affect "the application of [EU] or national rules requiring trade secret holders to disclose, for reasons of public interest, information, including trade secrets, to the public or to administrative or judicial authorities for the performance of the duties of those authorities."[119] The European provision regarding the exceptions does not introduce a stand-alone public interest defence. Indeed, a defence is available if the acquisition, use, or disclosure was "for revealing misconduct, wrongdoing or illegal activity, provided that the respondent acted for the purpose of protecting the general public interest".[120] Unlike the EU, in the UK the public interest is a defence in its own right.[121] Since the transposition deadline is in June 2018, one needs to wait and see how this provision will be interpreted.[122]

37 Additionally, one should remember that copyright protection covers both source code and object code[123] of the computer program implementing the algorithm. However, it leaves out some aspects, such as functionalities, data file formats, programming language, and graphic user interface. They are treated as "ideas" and therefore not copyrightable due to the idea-expression dichotomy.[124] The dichotomy is

WPL can exercise the right under Article 5(3)" (ibid [61]). The decision was upheld in appeal, although for different reasons (*SAS Institute v Worlds Programming Ltd IV* [2013] EWCA Civ 1482 [61], [109] per Tomlinson LJ).

110 Trade Secrets Directive, art 5(a). However, this defence risks weaknesses because the courts may be tempted to interpret it narrowly given that the underlying debate was about the protection of whistle-blowers and journalists, as one can also infer from the express reference to the freedom and pluralism of media.

111 Alexa Voice Service Agreement (last updated 30 January 2017), s 8. The Agreement was updated on 15 February 2018 in order to add Amazon Seller Services Pvt Ltd as the Amazon Party to the Agreement for developers who reside in India. Since the update is minor, this paper keeps referring to the previous version.

112 Alexa Voice Service Agreement, s 9.

113 Amazon's AI virtual assistant.

114 These include "images, audio, logos, specifications, code, documents, data, software, software development kits, libraries, application programming interfaces, applications, services and other information, technology, and related materials" (Alexa Voice Service Agreement, s 2).

115 In *SAS Institute Inc v Worlds Programming Ltd I (WPL)* [2010] EWHC 1829 (Ch), a similar program had been developed studying the competitor's one in breach of the license terms, because the purpose of the permitted act was not learning to use the SAS system (the sole purpose allowed by the license). After the reference to the Court of Justice and *SAS Institute Inc v World Programming Ltd II* [2012] ECR, the national court stated that if an act (e.g. studying) is permitted by the license, the purpose thereof is immaterial, and the exception operates (*SAS Institute III* (n 109); *SAS Institute IV* (n 109) [101]. On the EU case, see Guido Noto La Diega, 'Le idee e il muro del suono: I programmi per elaboratore nella più recente giurisprudenza europea' (2013) 2 Europa e diritto private 543.

116 *Campbell v Frisbee* [2002] EMLR 31, [23].

117 Lionel Bently and Brad Sherman, *Intellectual Property Law* (4th ed, OUP 2014) 1181.

118 *Lion Laboratories Ltd v Evans* [1985] QB 526, 537.

119 Trade Secrets Directive, art 1(2)(b).

120 Trade Secrets Directive, art 5(b).

121 *Initial Services v Putterill* [1968] 1 QB 396.

122 The only Member State that transposed the Trade Secrets Directive is Croatia, with *Zakon o zaštiti neobjavljenih informacija s tržišnom vrijednosti* ('Law on the Protection of Unpublished Information with Market Value') of 30 March 2018. Regrettably, not only the Croatian statute does not introduce a stand-alone public interest defence: it introduces an exception which is narrower than the Directive, being seemingly reduced to a defence for journalists (see art 8(1) and its reference to reporting, media, and pluralism).

123 Agreement on the trade-related aspects of intellectual property rights (TRIPs), art 10(1).

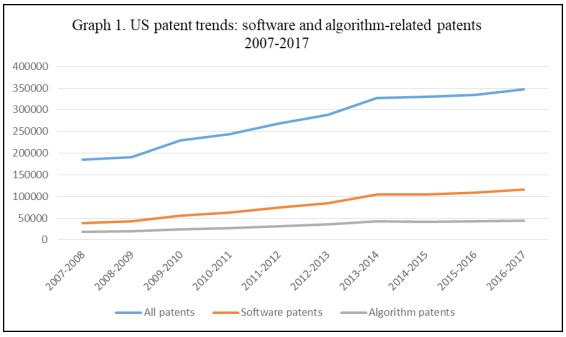
124 In the field of computer programs, on the idea whereby copyright covers the expression of the ideas and not the ideas in themselves, see *SAS Institute II* (n 117); *Navitaire Inc*

also one of the alleged reasons of the patentability of computer-implemented inventions. It has been noted, indeed, "copyright is not a sufficient form of protection where it is the *idea* behind the program which is its commercially valuable element."[125] Computer-related inventions are growing significantly also in connection to the Internet of Things,[126] despite the fact that the relevant patents can stifle innovation.[127] In the US,[128] in September 2017, there were 481,608 patent specifications referring to algorithms.[129] More than 67% of the algorithm-related patents (325,805) were issued over the last ten years with a growing trend reflecting the general increase in patents as shown by Table 1 and Graph 1.[130] Nearly 13% of all patents granted over the last 12 months concern algorithms (ten years ago only 9% of patents were algorithm-related).

38 Table 1. Software and algorithm patent trends in the US (2007-2017).

Period	All patents granted	Software patents	Algorithm patents
2016-2017*	346,543	115,896	44,110
2015-2016	333,767	108,305	42,481
2014-2015	329,722	104,212	41,125
2013-2014	327,729	103,918	42,215
2012-2013	288,989	84,891	35,427
2011-2012	268,157	74,689	32,070
2010-2011	244,338	63,638	27,377
2009-2010	229,694	56,367	24,920
2008-2009	190,285	42,947	19,426
2007-2008	185,340	38,225	17,871

** The period analysed is from 9 September 2016 to 9 September 2017. The same applies to the following rows.*



39 In theory, in the countries that signed the European Patent Convention (and in the others which adopted a hybrid system,[131] such as India), computer programs are not patentable "as such".[132] Features of the computer program,[133] as well as the presence of a device defined in the claim[134] may lend technical character. Moreover, a computer program by itself can be patented if it brings about a further technical effect going beyond the normal physical interactions between the said program and the computer.[135] In the UK, after *Symbian v Comptroller-General of Patents,*[136] the focus is not on the question whether the contribution falls within the excluded subject matter,[137] but on whether the invention makes a technical contribution to the known art, even if the computer program does not bring any novel effect outside of a computer.[138]

v EasyJet Airline Co Ltd (III) [2004] EWHC 1725 (Ch). On the graphic user interface (GUI), see *Bezpečnostní softwarová asociace - Svaz softwarové ochrany v Ministerstvo kultury* [2010] ECR I-13971.

125 Daniel J.M. Attridge, 'Challenging claims! Patenting computer programs in Europe and the USA' (2001) 1 Intellectual Property Quarterly 22.

126 Guido Noto La Diega, 'Software patents and the Internet of Things in Europe, the United States, and India' (2017) 39(3) European Intellectual Property Review 173.

127 As shown by Daehwan Koo, 'Patent and copyright protection of computer programs' (2002) Intellectual Property Quarterly 172, 173, "[p]atent and copyright do not provide optimum protection for software innovations, because they are based on exclusive property rights which impede follow-on small-scale innovations such as software innovations".

128 The search engine of the European Patent Office does not allow to retrieve data of a similar granularity.

129 These data and the following data, including those used in Table 1, were retrieved (and partly calculated) using the US Patent and Trademark Office (USPTO) Patent Full-Text and Image Database, accessed on 9 September 2017. The method is the same used by Allen Clark Zoracki, 'When Is an Algorithm Invented: The Need for a New Paradigm for Evaluating an Algorithm for Intellectual Property Protection' (2005) 15 Alb. L.J. Sci. & Tech. 579, 585, with regards to the patents granted between 1994 and 2003.

130 The faster growth of software patents may be related also to the fact that algorithms can be perceived as mere abstract ideas non-eligible for patent protection. Therefore, some applicants may purposely shy away from using the word 'algorithm' in the specifications.

131 There are mainly three systems for the protection of computer programs. First, one may refer to the double binary of copyright and patent protection, as exemplified by the US approach. Second, there is the hybrid system where alongside copyright, there is a rule excluding computer programs from patentability, but only if claimed 'as such.' This is the system that one finds in Europe. Finally, there is single binary (only copyright) protection. This system is the least common, see the Philippines, which are moving towards the hybrid system. There is a convergence between the double binary and the hybrid systems, with a trend towards a *de facto* generalised double binary.

132 European Patent Convention, art 52(2).

133 T 1173/97 (Computer program product) of 1 July 1998.

134 T 0424/03 (Clipboard formats I/MICROSOFT) of 23 February 2006; T 0258/03 (Auction method/HITACHI) of 21 April 2004.

135 G 3/08 (Referral by the President of the EPO in relation to a point of law ... of 16 October 2009; T 1173/97 (Computerprogrammprodukt) of 1 July 1998.

136 [2008] EWCA Civ 1066, [16] [49] [51 [59].

137 This was the law under *Aerotel v Telco Holdings* [2007] 1 All ER 225, [40].

138 *Shopalotto.com Ltd, Re patent application GB 0017772.5*: PATC 7 November 2005 [2005] EWHC 2416 (Pat), found that in

40 Even though some courts or examiners may consider algorithms as computer programs, they should probably be more precisely seen as mathematical methods. The European Patent Office's Board of Appeals stated that algorithms are mathematical methods, as such deemed to be non-inventions; therefore, a technical character of the algorithm can be recognised only if it serves a technical purpose.[139] The fact that a computer-implemented invention includes an algorithm can make the latter patentable. Indeed, it has been recognised that mathematical algorithms may contribute to the technical character of an invention, inasmuch as they serve a technical purpose.[140] For example, text classification does not qualify as technical purpose.[141] A technical effect may arise either from the provision of data about a technical process, or from the provision of data that is applied directly in a technical process.[142] However, the inclusion of an algorithm in a patent application for a computer-implemented invention does not, in itself, ensure patentability. Indeed, not all efficiency aspects of an algorithm are by definition without relevance for the question of whether the algorithm provides a technical contribution. However, such technical considerations must go beyond merely finding a computer algorithm to carry out some procedure.[143] In the US, legal scholars[144] have focused on how to evidence an improvement in algorithmic technique. It has been suggested to run the algorithm on test problems with known solutions and compare the results with those of algorithms in the prior art, with particular regards to speed, performance, memory usage, and ease of implementation.[145]

41 Unlike copyright, most uses of a computer-implemented invention are prohibited if not authorised and maybe that is why scholars tend to overlook patent exceptions.[146] However, in proceedings for infringement, defendants may avail themselves of the private non-commercial use[147] and experimental use[148] defences. One can qualify for the first immunity even when the resulting information has a commercial benefit, or the subjective intention was not commercial.[149] This is particularly interesting because in the UK there is no private copy exception to copyright.[150] As to the second defence, activities to discover something unknown, to test a hypothesis or to assess whether an invention works are considered as experiments and non-infringing.[151] However, this defence may be of limited use in the context of accessing algorithms, because it cannot be invoked to show that a product works in the way claimed by the maker.[152] Yet, arguably, when accessing the algorithm, the affected individual would have an interest to show that the algorithm-related invention does *not* work in the way claimed by the maker. Thus, this defence could be usefully invoked when an algorithm-related invention is used to take decisions whose rationale one wants to contest.

42 Intellectual property seems to create more problems than solutions to the issue at hand. The route above is weak for at least four reasons. First, the overlap between, if not abuse of, intellectual property rights[153] create a legal black box which is very difficult to open. Second, the application of the study and observation exception presupposes the lawful use of a copy of the software,[154] which is rarely the case in the event of algorithmic decisions. Third, even though the analysed copyright exceptions have been qualified as quasi-rights, there is no precedent

a claim for a lottery game played on the internet, the technical effect did not go beyond the mere loading of a program into a computer.

139 T 1784/06 (Classification method/COMPTEL) of 21 September 2012. In the UK, in *Gale's Application* [1991] RPC 305, it was held that an algorithm used to calculate square roots could not be patented because it lacked any technical character.

140 Ibid. 3.1.1. See also T 2249/13 (Mobile device/TRADE CAPTURE) of 17 October 2014.

141 T 1358/09 (Classification/BDGB ENTERPRISE SOFTWARE) of 21 November 2014; T 1316/09 of 18 December 2012.

142 T 1670/07 (Shopping with mobile device/NOKIA) of 11 July 2013.

143 T 1358/09 (n 142); see G 3/08 (n 136). *HTC Europe Co Ltd v Apple Inc* [2013] EWCA Civ 451, provides a good guidance to understand if computer programs and algorithms are patentable because the invention produces a technical effect that goes beyond the excluded subject matter.

144 Zoracki (n 129) 579.

145 ibid 605.

146 See David Gilat, *Experimental use and patents* (Wiley 1995); Alan J Devlin, 'Restricting experimental use' (2009) 32(2) Harvard Journal of Law and Public Policy 599; Jessica C Lai, 'A right to adequate remuneration for the experimental use exception in patent law: collectively managing our way through the thickets and stacks in research?' (2016) 1 Intellectual Property Quarterly 63.

147 Patents Act 1977, s 60(5)(a).

148 Patents Act 1977, s 60(5)(b).

149 *SKF Laboratories Ltd v Evans Medical Ltd* [1989] FSR 513.

150 See Guido Noto La Diega, 'In Light of the Ends. Copyright Hysteresis and Private Copy Exception after the British Academy of Songwriters, Composers and Authors (BASCA) and Others v Secretary of State for Business, Innovation and Skills Case', in *Studi giuridici europei 2014* (C Franchini ed, Giappichelli 2016) 39.

151 *Monsanto Co. v Stauffer Chemicals Co. and another* [1985] RPC 515 (CA); *Micro-Chemicals et al. v Smith Kline and French Inter-American Corporation* (1971) 25 DLR 78, 89.

152 *Monsanto* (n 152) 542; *Auchinloss v Agricultural and Veterinary Supplies Ltd* [1999] RPC 397, 405.

153 *cf*, more generally, Neil Wilkof and Shamnad Basheer (eds) *Overlapping Intellectual Property Rights* (OUP 2012).

154 Only the 'person having a right to use a copy of a computer program' can avail themselves of the exception (Software Directive, art 5(3).

interpreting said exceptions to open the algorithmic black box. Lastly, it requires considerable skills to open an algorithm by observing and studying the software that implements it. In most cases, there would be the need to ask an expert third party to carry out such activities on behalf of the lawful user of the software. However, applying *SAS Institute*,[155] it is unclear whether said third parties would qualify as lawful users. In the negative, this exception would be of little use in the majority of cases.

43 To add to the complexity, intellectual property will always be balanced with competing interests, such as data protection. As correctly pointed out, for instance, "trade secrecy (...) may make it difficult for data controllers to comply with their obligation of transparent processing."[156] Let us have a look, therefore, at the relevant data protection regime.

D. Algorithmic decision-making and EU data protection

44 The use of algorithms is under the lens of the data protection authorities, especially with regards to profiling. The European Data Protection Supervisor[157] has pointed out that the problem is not profiling as such, but "the lack of meaningful information about the algorithmic logic which develops these profiles and has an effect on the data subject."[158]

45 Under the Data Protection Directive,[159] there is a right not to be subject to a decision which produces legal effects or significantly affects the data subjects, if the decision is based solely on automated processing of data aimed at evaluating certain personal aspects concerning them (e.g. creditworthiness). Moreover, there is a right to know the logic involved in any automated processing of data.[160] Nonetheless, one may be subject to an algorithmic decision in two scenarios.[161] Firstly, in the course of the entering into a contract (or of the performance thereof), provided the request for the entering into the contract (or the performance thereof), lodged by the data subject, has been satisfied or that there are suitable measures to safeguard his legitimate interests (e.g. the data subject could express their viewpoint). For instance, some law firms[162] are using AI-enabled computer programs to assess the merits of personal injury cases and decide, therefore, whether to accept the case or to draft contingency fee agreements. Secondly, and more generally, algorithmic decision-making may be authorised by a law, if there are measures to safeguard the data subject's legitimate interests.[163] Fraud and tax evasion prevention are the typical examples.[164]

46 The rules on algorithmic decision-making have been amended by the GDPR,[165] which is set to come into effect on 25 May 2018, also in the UK, regardless of Brexit.[166] The general principle is that data subjects should not be subject to algorithmic decisions. However, when non-human agents take a decision that has legal effects on the data subject's life "or similarly significantly affects him or her,"[167] the data subject has the rights to obtain human intervention, to express their point of view, as well as to contest the decision.[168] Correspondingly, the data controller

155 *SAS Institute IV* (n 108).

156 Kamarinou (n 76) 23.

157 The European Data Protection Supervisor is the EU data protection authority. They inter alia ensure the protection of personal data and privacy when EU institutions and bodies process personal data. See Regulation (EC) No 45/2001 of the European Parliament and of the Council of 18 December 2000 on the protection of individuals with regard to the processing of personal data by the Community institutions and bodies and on the free movement of such data [2001] OJ L 8/1.

158 European Data Protection Supervisor, 'Recommendations on the EU's options for data protection reform' (2015/C 301/01), para 3.1.

159 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 (Data Protection Directive), art 15.

160 Data Protection Directive, art 12(a); recital 41. For a national implementation, see the UK Data Protection Act 1998, s 7(1)(d).

161 Data Protection Directive, art 15(2).

162 See Jane Croft, 'Legal firms unleash office automatons' (*The Financial Times*, 16 May 2016), https://www.ft.com/content/19807d3e-1765-11e6-9d98-00386a18e39d.

163 In the UK, the Secretary of State may prescribe in which circumstances (apart from a contract) an algorithmic decision may be exempt from the said rules (Data Protection Act 1998, s 12(5)(b)).

164 Information Commissioner's Office, *Overview of the General Data Protection Regulation (GDPR)* (ICO 2017) 27. The same example can be found in the GDPR, recital 71.

165 The underlying principle is the same, that is that "fully automated assessments of a person's character should not form the sole basis of decisions that significantly impinge upon the person's interests" (Lee Bygrave, 'Automated Profiling, Minding the Machine: Article 15 of the EC Data Protection Directive and Automated Profiling' (2001) 17(1) Computer Law & Security Review 17, 21) as suggested by Kamarinou (n 76) 8.

166 In the time between 25 May 2018 and 29 March 2019 the rules about algorithmic decision-making will be those resulting from a combination of GDPR and the Data Protection Act, after 29 March 2019 it is likely that only the Data Protection Act as amended will be in force. *cf* The Rt Hon Karen Bradley MP, Culture, Media and Sport Committee, *Oral evidence: Responsibilities of the Secretary of State for Culture, Media and Sport*, HC 764 (24 October 2016).

167 GDPR, art 22.

168 As to the latter, it would seem to us that this right as enshrined in art 22 of the GDPR is the same as the right to "challenge the decision" under recital 71. *Contra*, see Kuner (n 76) 2, who observe that even though the recital is not

shall provide "meaningful information about the logic involved"[169] in the algorithmic decision. It is likely that the national implementing measures of the Data Protection Directive will be amended or replaced to recognise a stronger protection to data subjects against algorithmic decisions.[170]

I. The general prohibition on solely automated decisions with a significant effect

47 Let us start with the provisions directly dealing with algorithmic decision-making;[171] it is open to debate whether they constitute a considerable step forward. The main right available to the data subject is the right not to be subject to a solely automated decision with legal effects or similarly significantly affecting them.[172] This can be interpreted as a general prohibition to make algorithmic decisions using personal data, or as a mere right to be oppose (after being informed about) the algorithmic decision.[173] In the UK, data subjects can require that no solely algorithmic decision be taken against them. However, if no such notice has effect and the decision is taken, the data controller has 21 days to give a written notice explaining the steps that they will take to comply with the data subject request.[174] Positively, in issuing some guidelines on algorithmic decision-making, the Article 29 Working Party[175] recommends treating this right as a general prohibition.[176] Regrettably, however, the only amendment introduced by the Data Protection Bill with regards to algorithmic decisions concerns the safeguarding measures that controllers should take when availing themselves of the consent-based exception.[177] Arguably, by refusing the "general prohibition" approach, the UK will not comply with the GDPR, with practical consequences for instance in terms of the legality of the EU-UK data transfers. If this provision expresses a core data protection principle,[178] a partial compliance may cause the EU to deem the UK protection of personal data inadequate, hence hindering cross-border data flows.[179]

48 Looking at the core of art 22, there are two main differences between the Data Protection Directive and the GDPR.

49 First, in the new provision there is an express reference to profiling as an example of automated processing. This brings clarity in a field currently perceived as particularly relevant, but it risks

binding, it "may embolden regulators and courts to try to compel data controllers to provide explanations of specific outcomes in particular cases, and not merely 'meaningful information' about 'logic'".

169 GDPR, arts 13(2)(f) and 14(2)(g).

170 In August 2017, the UK government announced a new Data Protection Bill, where "individuals will have greater say in decisions that are made about them based on automated processing. Where decisions are based on solely automated processing individuals can request that processing is reviewed by a person rather than a machine." (UK Department for Digital, Culture Media & Sport, A new data protection bill: Our planned reforms, The UK Government (7 August 2017), <https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/635900/2017-08-07_DP_Bill_-_Statement_of_Intent.pdf> accessed 3 March 2018). The Bill was introduced in the House of Lords on 13 September 2017 and it passed second reading at the House of Commons on 5 March 2018. Members of Parliament are considering the Bill in a Public Bill Committee, which is set to finish by 27 March 2018. See below for the analysis of s 14 of the Bill (as brought from the Lords), regarding algorithmic decision-making authorised by the law. It must be said that the fact that the only relevant provision in the Bill regards the limited issue of the algorithmic decisions authorised by law may be seen as a missed opportunity to thoroughly review the regime laid out in s 12 of the Data Protection Act 1998.

171 The focus of this section is on the rules regarding algorithmic decision-making. These do not apply to profiling *per se* if it is not followed by an algorithmic decision producing legal effects or similarly significantly affecting the data subject. For more information on the rules about profiling, regardless of whether or not it is followed by an algorithmic decision, please see ibid 17-25.

172 GDPR, art 22(1).

173 For the first interpretation, in favour of a general prohibition of algorithmic decisions, see the French *Loi n° 78-17* of 6 January 1978 *relative à l'informatique, aux fichiers et aux libertés*, art 10.

174 Data Protection Act 1998, s 12(3).

175 The Article 29 Working Party is an advisory body set up under the Data Protection Directive, art 29. It is composed by representatives from the Member States' data protection authorities, the European Data Protection Supervisor and the European Commission. The GDPR will replace it with the European Data Protection Board.

176 Article 29 Working Party (n 9) 12.

177 Other Member States are adopting implementing measures that are overlooking algorithmic decision-making. For instance, on 21 March 2018, the Italian Cabinet (*Consiglio dei Ministri*) adopted the draft decree implementing the GDPR (*Disposizioni per l'adeguamento della normativa nazionale alle disposizioni del Regolamento (UE) 2016/679 del Parlamento europeo e del Consiglio, del 27 aprile 2016, relativo alla protezione delle persone fisiche con riguardo al trattamento dei dati personali, nonché alla libera circolazione di tali dati, e che abroga la direttiva 95/46/CE (Regolamento generale sulla protezione dei dati*), hereinafter 'Draft GDPR implementing decree'). The Draft GDPR implementing decree does not provide anything on the matter. Unlike the UK, however, the GDPR will be directly applicable and, therefore, Italian data controllers will be bound directly by art 22 of the GDPR.

178 This is the view expressed by Bygrave (n 165) 22 about the similar provision in the Data Protection Directive. His observation is all the more true with regards to the GDPR for at least two reasons. First, because algorithmic decisions have become more common and more intrusive. Second, because the GDPR strengthens the relevant regime, thus confirming the importance of the provision.

179 GDPR, art 45.

narrowing the interpretation of the provision thus excluding forms of algorithmic decision-making which do not include profiling. Therefore, it is positive that the Article 29 Working Party has observed that "(a)utomated decisions can be made with or without profiling; profiling can take place without making automated decisions."[180]

50 Second, and most importantly, one has the said right only if the decision produces legal effects concerning one "or *similarly* significantly affects him or her."[181] This addition goes in the opposite direction to the one taken when the draft GDPR was first published and it had been suggested that art 22 should cover not only decisions producing legal effects or which significantly affect data subjects, but also the "collection of data for the purpose of profiling and the creation of profiles as such."[182]

51 Now, "legal effect" is quite straightforward, including all the scenarios where a decision affects a person's rights based on laws or contracts.[183] In turn, "similarly" may narrow the scope of the provision, if compared with the previous wording, where no reference to this adverb was made. Indeed, it may be seen as meaning that one does not have the right to object to algorithmic decision-making if the effect is not similar to a legal effect[184] (e.g. significant distress or missed professional opportunities as a consequence of being permanently banned from a popular social network).[185] If this interpretation were followed, broader national implementations may need to be reviewed accordingly. For instance, the UK refers to decisions take for "the purpose of evaluating matters relating to him such as, for example, his performance at work, his creditworthiness, his reliability or his conduct"[186]. The Information Commissioner's Office accepts that it is hard to explain what "significant effect" means, but it suggests that it refers to "some consequence that is more than trivial and potentially has an unfavourable outcome."[187] Businesses have been asking for more detailed guidance[188] and this has partly arrived with the Article 29 Working Party's guidelines that indicated that "similarly" means that "the threshold for *significance* must be similar."[189] Therefore, in order for a decision to fall within the scope of art 22, it must not necessarily be a quasi-legal effect in terms of content, being sufficient a decision which profoundly affects the individual as much as a decision affecting her or his rights would. Adding details to the UK attempt of definition, the EU advisory body point out that a similarly significant effect must be "more than trivial and must be sufficiently great or important to be worthy of attention."[190] The concept is broad enough to encompass a vast number of scenarios, from e-recruiting to online behavioural advertising, especially if intrusive and targeted to vulnerable groups,[191] as well as consumer manipulation.[192]

52 Even before understanding what 'legal' means, one should clarify what a 'decision' is. It has been suggested that this could include "an interim or individual step taken during the automated processing."[193] It would seem, however, that only rarely interim measures and individual steps will qualify for the application of art 22 of the GDPR, because the provision requires a decision with legal effect or "similarly significant."

53 Some aspects of this regime are not clear yet. For instance, it is open to debate what *solely* automated means. In the past, it was relatively easy to understand what 'solely' meant. There was a limited number of organisations taking significant algorithmic decisions and the technologies used were

180 ibid 8.

181 GDPR, art 22(1).

182 Article 29 Working Party, 'Advice paper on essential elements of a definition and a provision on profiling within the EU General Data Protection Regulation' (13 May 2013), para 2(a).

183 *cf* ibid 10.

184 This would constitute a weakening of many national implementing regimes. For instance, in the UK the Data Protection Act 1998 refers generally to decisions which significantly affect the individual (s 12(1)).

185 *cf* Jilian York, 'Getting banned from Facebook can have unexpected and professionally devastating consequences' (Quartz, 31 March 2016) <https://qz.com/651001/getting-banned-from-facebook-can-have-unexpected-and-professionally-devastating-consequences/> accessed 1 March 2018.

186 Data Protection Act 1998, s 12(1).

187 Information Commissioner's Office, *Feedback request – profiling and automated decision-making* (ICO 2017) 19.

188 The digital technology industry in Europe would welcome such guidance. For instance, 'Input on Automated Individual Decision Making & Data Breach Notification' (DigitalEurope, 5 April 2017) 3 <http://www.digitaleurope.org/DesktopModules/Bring2mind/DMX/Download.aspx?Command=Core_Download&EntryId=2390&language=en-US&PortalId=0&TabId=353> accessed 1 March 2018, "would appreciate (...) clarification in the future guidance on how companies should interpret these two cumulative conditions as well as examples of such effects in different sectors".

189 Article 29 Working Party (n 9) 10.

190 ibid 10.

191 *cf* Guido Noto La Diega, 'Some considerations on intelligent online behavioural advertising' (2017) 66-67 Revue du droit des technologies de l'information 53.

192 Artificial intelligence is increasingly used to predict consumers' behaviour in order to lock them in by means of addiction. Evidence has been recently uncovered about such manipulating practices in the gambling industry. See Mattha Busby, 'Revealed: how bookies use AI to keep gamblers hooked' (*The Guardian*, 30 April 2014) <https://www.theguardian.com/technology/2018/apr/30/bookies-using-ai-to-keep-gamblers-hooked-insiders-say> accessed 2 May 2018.

193 ibid 12.

quite rudimental; therefore, reviewing the machine-generated data was relatively straightforward and once a human being reviewed the data, the decision was no longer solely automated.[194] In light of increasingly complex (and accordingly opaque) algorithmic techniques and of the ubiquitous nature of the phenomenon of algorithmic decisions, that approach should be abandoned. To what extent is the human intervention meaningful vis-á-vis black-box decisions?

54 The UK Information Commissioner's Office recently requested feedback on some points of the GDPR,[195] and they have suggested that 'solely' should "cover those automated decision-making processes where a human exercises no real influence on the outcome of the decision, for example where the result of the profiling or process is not assessed by a person before being formalised as a decision."[196] The risk of this interpretation is that it is not always easy - especially from the data subject's perspective - which role the human being played in the decision (was the human being a passive operator? Which discretion did they have while assessing the result?). Moreover, "it may not be feasible for a human to conduct a meaningful review of a process that may have involved third-party data and algorithms (which may contain trade secrets), prelearned models, or inherently opaque machine learning technique."[197] Therefore, it would seem more appropriate to recognise the right not to be subject to an algorithmic decision every time that there is not a human being clearly taking the final decision.[198] It would seem that the Article 29 Working Party hold similar views when they state that a decision is not wholly automated when alongside an automated profile, there is "additional meaningful intervention carried out by humans before any decision is applied to an individual."[199] However, there is still a lack of clarity. Indeed, in order to clarify when art 22 GDPR applies or not, the Article 29 Working Party makes the following examples. If a human decides whether to agree the loan based on a profile produced by purely automated means, then art 22 will not apply. In turn, if an algorithm decides whether the loan is agreed and the decision is automatically delivered to the individual, without any meaningful human input, then art 22 will apply. The point is that there is a substantial grey area here. For instance, it is unclear whether art 22 applies when the algorithmic system takes the decision, but a human being reviews it. Arguably, the human review could qualify as "meaningful human input", but this will have to be assessed on a case-by-case basis.

55 Even more importantly, controllers should refrain from "fabricating human involvement"[200] with the purpose of sidestepping art 22; this provision will apply every time that there is not meaningful and genuine human intervention, for instance in the form of actual oversight by a person with "authority and competence to change the decision."[201] It is important to stress that the GDPR applies to every automated profiling carried out on personal data to evaluate a natural person's personal aspects, not only to the 'solely' automated one, which means that the general GDPR rules and standards will apply to profiling even when a human being plays a substantial role in the creation of the relevant profile.[202]

II. Three exceptions: contract, consent, law

56 Even though "as a rule, there is a prohibition on fully automated individual decision-making (...) that has a legal or similarly significant effect,"[203] this rule has some exceptions. The GDPR has innovated the systems of the exceptions not only by adding a consent-based exception, but also by clarifying the scope of the pre-existing ones. It is unfortunate that the UK Data Protection Bill[204] is missing out on this opportunity. Indeed, the only innovation that it is being introduced regards algorithmic decisions authorised by law. The UK will keep allowing such decisions in circumstances prescribed by the Secretary of State, in relation to a contract, when authorised or required by or under any enactment, effect of the decision is to grant a request of the data subject, or when steps have been taken to safeguard the legitimate interests of the data subject.

194 This is still the approach that one can find in Information Commissioner's Office, 'Guide to Data Protection' (*ICO*, 11 May 2016) <https://www.legislation.gov.uk/ukpga/1998/29/pdfs/ukpgaod_19980029_en.pdf> accessed 18 March 2018.

195 Information Commissioner's Office (n 186).

196 ibid 19.

197 Kuner (n 76) 2.

198 Along the same lines, with regards to the Data Protection Directive, it has been noted that the regime will operate every time that there is not a human being exercising "real influence on the outcome of a particular decision-making process" (Bygrave (n 165) 20).

199 Article 29 Working Party (n 9) 8.

200 ibid 10. As an example, they observe that "if someone routinely applies automatically generated profiles to individuals without any actual influence on the result, this would still be a decision based solely on automated processing." (ibid 10).

201 ibid 10.

202 ibid 6. For instance, profiling is rarely transparent. However, the controller must provide data subjects with concise, transparent, intelligible and easily accessible information about the processing of their personal data (GDPR, art 12(1)).

203 ibid 9.

204 Data Protection Bill [HL] 2017-19, s 14.

No consent-based exception is provided. Unlike the interpretation of the right not to be subject to an algorithmic decision as a general prohibition, the lack of implementation of the consent-based exception is unlikely to endanger the cross-border data transfers with the EU. Indeed, a lack thereof might ensure a stronger protection of personal data. In turn, the broad wording of the contractual exception may be more problematic.[205]

57 Art 22 brings clarity to the scenario regarding the entering and performance of the contract by simplifying the language and restricting the contractual exception to the instances when the algorithmic decision-making is *necessary* to enter into a contract or for its performance.[206] One may argue, going back to the example of the contingency fee agreements, that in that scenario the algorithmic decision would not be necessary and, thus, it would not fall within the scope of this exception. Following the European Data Protection Supervisor's approach, if a less privacy-intrusive method is available, then the algorithmic decision is not necessary and, therefore, it is not allowed.[207]

58 In turn, the new exception based on the data subject's explicit consent[208] is problematic. Consent is explicit when there is "an express statement rather than some other affirmative action."[209] Indeed, given the imbalance of bargaining power that characterises many transactions, one should not be surprised if, for instance, a bank could force a potential client requesting a loan to consent to a decision taken by an algorithm. The exception based a law authorising the decision while laying down measures to safeguard the data subject's legitimate interest[210] now includes a reference to the data subject's rights and freedoms and to both EU and national law. These changes are nugatory. Firstly, based on an *a minore ad maius* argument, it is obvious that if the decision shall respect the legitimate interests of the data subject, it shall do so also with regards to the more relevant rights and freedoms. Secondly, while the reference to national laws is a truism, the one to EU law cannot be interpreted as a power to legislate beyond what already provided by the treaties. However, the growth of artificial intelligence (AI) may have an impact on the analysed regime. Not only because, generally, AI does not always make it feasible to access the rationale of algorithmic decisions. With specific regards to the consent-based exception, it is fair to wonder, "how can informed consent be obtained in relation to a process that may be inherently non-transparent (a 'black box')."[211]

59 The third exception regards national and EU laws authorising algorithmic decisions.[212] Regrettably, the Article 29 Working Party do not provide any guidance on the matter. Whereas recital 71 refers only to fraud, tax evasion, and reliability of the service, it would seem that EU and national authorities may allow algorithmic decisions for a potentially infinite number of purposes. Indeed, recital 73 provides that EU and national laws can impose restrictions concerning "decisions based on profiling" in inter alia order to prevent or react to breaches of ethics for regulated professions or for the keeping of public registers kept for reasons of public interest. Therefore, for instance, a Member State could allow algorithmic decisions to disbar a barrister who behaved unethically. Nor are there limits to which kind of public registers a state may keep, for instance for surveillance purposes.[213] One should not think, however, that if a law authorises the algorithmic-decision making in a specific field, say fraud, data protection legislation can be eluded altogether. Alongside the rights to access, the information rights and right to a human judge, data controllers will still have to comply with all the other data protection principles, including accountability.[214] The Data Protection Directive required the laws authorising algorithmic decisions to safeguard only the data subjects' legitimate interests and not also their rights

205 The Data Protection Act 1998 enables data controllers to make algorithmic decisions in the course of steps taken for the purpose of considering whether to enter into a contract, with a view to entering into such a contract, or in the course of performing such a contract, or if the decision is authorised or required by or under any enactment (Data Protection Act 1998, s 12(6)).

206 GDPR, art 22(2)(a).

207 *cf* European Data Protection Supervisor, 'Assessing the necessity of measures that limit the fundamental right to the protection of personal data. A Toolkit' (European Data Protection Supervisor, 11 April 2017), <https://edps.europa.eu/sites/edp/files/publication/17-04-11_necessity_toolkit_en_0.pdf> accessed 9 March 2018.

208 GDPR, art 22(2)(c).

209 Article 29 Working Party (n 9) 13. These guidelines do not provide sufficient clarity as to how to ensure explicit consent. The matter will be addressed in the forthcoming consent guidelines.

210 GDPR, art 22(2)(b).

211 Kuner (n 76) 1.

212 GDPR, art 22(2)(b).

213 Nonetheless, the restrictions should be necessary and proportionate in a democratic society to safeguard public security and in compliance with Charter and in the European Convention for the Protection of Human Rights and Fundamental Freedoms.

214 The GDPR provides that the algorithmic decision-making for purposes authorised by EU or national law should be "conducted in accordance with the regulations, standards and recommendations of Union institutions or national oversight bodies." (recital 71). Even though it may be interpreted as referring only to fraud and tax evasion, it would be absurd to exclude other purposes specifically authorised by the law (the reference is illustrative, not exhaustive). It is submitted that the data protection authorities should be deemed as oversight bodies and the data protection laws should still apply even when an algorithmic decision is allowed.

and freedoms. Moreover, it did not specify which laws could authorise algorithmic decisions. The GDPR, in turn, now includes a reference to the data subject's rights and freedoms and it clarifies that both EU and national laws can authorise algorithmic decisions. Arguably, these changes are nugatory. Firstly, based on an *a minore ad maius* argument, it is obvious that if the decision should respect the legitimate interests, all the more it should do so with rights and freedoms. Secondly, the clarification that national law can be a legal basis is a truism. So is the one about EU law, which should not be interpreted as a power to legislate beyond what already provided by the treaties.

60 The UK Data Protection Bill[215] provides more detail as to the procedure to follow when an algorithmic decision falls under the third exception. Indeed, the controller must, as soon as reasonably practicable, notify the data subject in writing that a decision has been taken based solely on automated processing. Correspondingly, the data subject may, before the end of the period of 21 days, beginning with receipt of the notification, request the controller to reconsider the decision, or take a new decision that is not based solely on automated processing. The provision goes on to point out what the controller must do if such request is made. The procedure is the same that the Data Protection Act currently provides for non-exempt decisions, but interestingly the new regime is more protective of the data subject if compared to the previous one. Indeed, currently the data controller's notice must only indicate the steps the controller intends to take to comply with the request. This information must be notified before the end of the period of 21 days beginning with receipt of the request. On top of this, the Data Protection Bill provides that when the law authorises an algorithmic decision, the data controller shall consider the request, comply with it, and inform the data subject of the steps taken to comply, and of the outcome of complying with the request. The wording suggests that data controllers have some discretion in complying. However, the discretion regards how to comply, not whether to comply. The only reason why a denial could be allowed would be if the algorithmic decision was not taken solely on the basis of automated processing, if the decision does not significantly affect the data subject, or if it is impossible to identify the data subject.[216] If the data controller violated the limits of its discretion, the data subject may appeal the decision judicially.

61 Interpreters will need to avoid a visible inconsistency in the new UK regime on algorithmic decision-making. Namely, it is not rational to give the data subject a weaker protection when a non-exempt decision is at issue, if compared to a decision authorised by the law.

62 One may observe a departure of UK data protection law from the GDPR. In the UK, there is a three-layered system. As a rule, data subjects must be informed of non-exempt algorithmic decisions and can request that no such decision be taken. If no request has effect, they still have a right to be informed and to request a reconsideration or a human decision. Reconsideration and the right to a human judge, after the Data Protection Bill is enacted, will apply also to the algorithmic decisions authorised by law. Obviously, no right to pre-empt such a decision would apply. Thirdly, data subjects have no rights regarding the other exempt decisions.[217] This may raise concerns in terms of adequacy of the protection of personal data in the UK in the context of cross-border data transfer with the EU. Since consent is not one of the exceptions, the rights of the first layer will apply. In the EU, in turn, there is a clearer and stronger model. The rule is the general prohibition to take solely algorithmic decisions. There are only three justifications that can be used to make some decisions, but all of them are accompanied by strong safeguards for the data subject.

63 Lastly, it is not entirely clear if the list of exceptions (contract, consent, law) is exhaustive. A recital[218] refers to algorithmic decision-making for the purpose of ensuring the security and reliability of a service provided by the controller. However, this should not be interpreted as a fourth exception or as proof of the non-exhaustive character of the list of exceptions. It is plausible, indeed, that this is only an example of a purpose for which national and EU laws can authorise the said decision-making.[219]

III. Measures to safeguard the data subjects' rights, freedoms, and legitimate interests

64 The main commendable innovation in the GDPR regards the measures to safeguard the data subject's rights, freedoms, and legitimate interests affected by

215 Data Protection Bill [HL] 2017-19, s 14.

216 The GDPR is very clear in stressing that the data controller "shall not refuse to act on the request of the data subject for exercising his or her rights under Articles 15 to 22, unless the controller demonstrates that it is not in a position to identify the data subject" (art 12(2).

217 Under the Data Protection Act, s 12(4), the data subject's request not to take solely algorithmic decisions does not have effect in relation to an exempt decision; and nothing in s 12(2), regarding the data controller's notice, applies to an exempt decision.

218 GDPR, recital 71.

219 See Article 29 Working Party (n 9) 12.

an algorithmic decision.[220]

65 First, now these measures refer also to the contractual and consent-based exceptions. Second, they are no longer limited to the right to express one's viewpoint. The provision shall be interpreted as the right to obtain human intervention on the part of the controller and the right to contest the decision. Therefore, if there is a law authorising algorithmic decision making,[221] if this is necessary for a contract, or if there is the data subject's explicit consent, a data controller may use algorithms to take decisions having legal effects or similarly affecting the data subject. However, data controllers shall put in place a procedure to appeal the decision with meaningful oversight by a human being that shall ensure an effective right of defence to the data subject.[222]

66 This is a major victory for those who think that human decision-making is still better than the automated one.[223] However, it is unclear which steps the data controller should take once the data subjects avail themselves of the analysed remedy. The Article 29 Working Party further clarify that the review must be carried out by a human being with appropriate authority and capability to change the decision and who shall thoroughly assess "all the relevant data, including any additional information provided by the data subject."[224]

220 GDPR, art 22(3).

221 The wording of the provision is not crystal clear. Indeed, art 22(2)(b) applies the said measures to the algorithmic decision-making authorised by the law. Then, the following paragraph extends these measures to the other two exceptions and it specifies that they include "at least" the right to human intervention, to express the viewpoint, and to contest the decision. It may be argued, therefore, that when the law authorises algorithmic decision-making, the mere right to express one's viewpoint (as provided under the old regime) would be sufficient. However, this would seem to go against the overall purpose of the GDPR and of art 22. Moreover, the express reference to "at least" is likely to mean that those three rights are the minimum core of the measures that safeguard the data subject's rights, freedoms, and legitimate interests. Furthermore, recital 71 suggests that these measures should be put in place "[i]n any case".

222 Obviously, if such a system is not in place or if the data subject is not satisfied, the usual judicial remedies will be available.

223 A slightly different perspective is taken by Kamarinou (n 76) 22, who observe that "it may already in some contexts make sense to replace the current model, whereby individuals can appeal to a human against a machine decision, with the reverse model whereby individuals would have a right to appeal to a machine against a decision made by a human".

224 Article 29 Working Party (n 9) 15.

IV. Transparency obligations: a right to explanation?

67 Moving onto the transparency obligations, these are nearly entirely new,[225] given that under the Data Protection Directive there was only the right to access, which included the logic involved in the algorithmic decision.[226] Innovatively, the processing is not deemed fair and transparent, if the controller does not - at the time when personal data is obtained from the data subject - provide specific information on three matters.[227] First, controllers must disclose the existence of algorithmic decision-making. Second, they need to inform the data subject about the logic involved. Third, the algorithm must be opened in order to provide "meaningful information about [...] the significance and the envisaged consequences of such processing for the data subject."[228] The same right applies when the data was not obtained from the data subject, who has the right to be informed within a reasonable timeframe[229] (at the latest within one month),[230] at the time of the first communication with the data subject,[231] or when the data is fist disclosed to a third party.[232] Data controllers who merely make the information available, without actively bringing it to the data subject's attention, do not meet their transparency obligations. On top of the obligation to inform, there is the right of access, which again regards the existence of the algorithmic decision-making itself and meaningful information about the logic, the significance, and the consequences.[233]

68 One should welcome positively the obligation to provide (and the right to access) meaningful information and the reference to the envisaged consequences and significance of the decision. While

225 They are new at an EU level, but not necessarily at the national one. For instance, the Data Protection Act 1998 provides the controller's obligation to notify the data subject that the decision was algorithmic (s 12(2)(a)), unless the data subject already required that the decision is not taken based solely on automated processing (s 12(1)-(2)).

226 Data Protection Directive, art 12(a).

227 Information rights exist under the GDPR also when there is no algorithmic decision significantly affecting a data subject. See the principles of fair and transparent processing and arts 13 and 14 of the GDPR. According to Article 29 Working Party (n 9) 13 considers as "good practice to provide the above information whether or not the processing falls within the narrow Article 22(1) definition."

228 GDPR, art 13(2)(f).

229 This is similar to the UK provision, which refers to "as soon as reasonably practicable" (Data Protection Act 1998, s 12(2)(a)).

230 GDPR, art 14(3)(a).

231 GDPR, art 14(3)(b).

232 GDPR, art 14(3)(c).

233 GDPR, art 15(1)(h).

"envisaged" suggests that information must be provided "about intended or future processing,"[234] it would seem that "significance" requires real, tangible examples of how the decision may affect the data subject.[235]

69 Generally speaking, such meaningful information is what the data subject, who normally will not be a computer scientist, is likely to be interested in. Therefore, for instance, a technical document which includes the algorithm used and the mere explanation of the logic in mathematical terms will not in itself meet the legal requirement. Arguably, this should be interpreted as the disclosure of the algorithm with an explanation in non-technical terms of the rationale of the decision and criteria relied upon.[236] Regrettably, the Article 29 Working Party[237] do not consider the disclosure of the algorithm as necessary under the said transparency obligations. However, in order to have a full picture, the data subject has a legitimate interest in asking an expert to analyse the algorithm in order to better challenge the decision. A different interpretation would not comply with right to an effective remedy[238] and to a fair trial[239] under the Charter of Fundamental Rights of the EU and the European Convention of Human Rights.

70 Obviously, it may be the case that, due to the characteristics of artificial intelligence alone, it could be impossible to explain an algorithmic process "in a way that is intelligible to a data subject."[240] However, the data controller should make any reasonable effort to adequately inform the data subject.

71 Scholars have recently criticised the provision because it would entail a right to be informed, but no right to explanation.[241] Others,[242] conversely, have pointed out that Articles 15 and 22 should have a wide interpretation that might prove adequate to cope with the transparency challenge; they propose a legibility stress test for the data controller.

72 To overcome this issue, those who exclude that a right to explanation is provided by the GDPR make a number of recommendations to improve transparency and accountability of algorithmic decision-making, including a trusted third-party regulatory or supervisory body that can investigate algorithmic decisions if one feels that they have been discriminated against. Whereas the idea of an AI watchdog can be a positive one, this paper argues that the information rights provided with regards to algorithmic decision-making – which include a reference to the significance and consequences of the decision – can be interpreted as meaning a right to explanation.[243] Denying it would mean playing down the great potential of legal interpretation. A counterargument could be that the wording 'right to obtain information' can be found in recital 71, but not in art 22; this placement in a non-binding part of the Regulation (a recital) has been seen as "a purposeful change deliberated in trilogue."[244] However, the pivotal role of recitals in interpreting the provisions of an EU act has been expressly recognised by the Commission.[245] The reference to the right of explanation in the recital shall be, therefore, used to properly construe art 22 to reflect the context of the provision and the overall purpose of the GDPR, that is increasing the protection of the data subjects' rights. Hence, even though applying the literal rule, art 22 would not contain a right to explanation, a purposive approach and a correct valorisation of the role of recitals make it clear that data subjects are entitled to such a right. In addition, the data controller is expressly required to provide "concise, transparent, intelligible and easily accessible form, using clear and plain language."[246]

234 Article 29 Working Party (n 9) 14.

235 ibid 14.

236 ibid.

237 This is the interpretation given to recital 60 of the GDPR by Article 29 Working Party (n 9) 13.

238 Charter of Fundamental Rights of the European Union, art 47(1); European Convention on Human Rights, art 13.

239 Charter of Fundamental Rights of the European Union, art 47(2); European Convention on Human Rights, art 6.

240 Kuner (n 76) 1, who suggest that a "high-level, non-technical, description of the decision-making process is more likely to be meaningful" (ibid 2).

241 Sandra Wachter et al., 'Why a Right to Explanation of Automated Decision-Making Does Not Exist in the General Data Protection Regulation' (2017) 7(2) International Data Privacy Law 76. For a similar somehow pessimistic take, see Lilian Edwards and Michael Veale, 'Slave to the algorithm? Why a 'right to an explanation' is probably not the remedy you are looking for' (2017) 16(1) Duke Law & Technology Review 18.

242 Gianclaudio Malgieri and Giovanni Comandé, 'Why a Right to Legibility of Automated Decision-Making Exists in the General Data Protection Regulation' (2017) 7(4) International Data Privacy Law 243. On the optimistic front, see also Julia Powles and Hal Hodson, 'Google DeepMind and healthcare in an age of algorithms' (2017) 7 Health Technol 351. Between the two poles, see e.g. Tal Zarsky, 'The Trouble with Algorithmic Decisions: An Analytic Road Map to Examine Efficiency and Fairness in Automated and Opaque Decision Making' (2016) 41(1) Science, Technology, & Human Values 118; Mireille Hildebrandt, 'The New Imbroglio - Living with Machine Algorithms', in Liisa Janssens (ed) *The Art of Ethics in the Information Society* (Amsterdam University Press 2016).

243 There is the risk, however, that the courts will interpret the analysed provisions in a narrow way, focusing on the weaknesses of the new regime.

244 Wachter (n 238) 96.

245 Roberto Baratta, 'Complexity of EU law in the domestic implementing process' (2014) 19th Quality of legislation seminar "EU legislative drafting: Views from those applying EU law in the Member States" 4 <http://ec.europa.eu/dgs/legal_service/seminars/20140703_baratta_speech.pdf> accessed 1 March 2018.

246 GDPR, art 12(1).

73 Lastly and commendably, the GDPR details the timescale and procedure to provide information.[247] In particular, the information should be provided without undue delay and in any event[248] within one month of receipt of the request. The information must be in electronic form to reflect the form of the request, unless the data subject requests otherwise.

74 Obviously, the problems with the black boxes remain, no matter how broad the interpretation given to the transparency obligations is. Therefore, the transparency obligations may not be fully effective "in cases where a machine learning process involves multiple data sources, dynamic development, and elements that are opaque, whether for technological or proprietary reasons."[249]

V. Algorithmic decisions with sensitive personal data

75 Another positive new provision regards sensitive personal data (e.g. data on health or sexuality). Artificial intelligence increasingly relies on this kind of data. One need only think that deep neural networks have been recently used to infer the sexual orientation of people from their faces.[250] Indeed, in principle, algorithmic decisions shall not be based on sensitive personal data.[251] For instance, an employer may not let an algorithm decide whether to fire an employee using health data. However, this data may be used with the data subject's explicit consent or in the interest of public health, provided that measures to safeguard the data subject's rights, freedoms, and legitimate interests are in place. Even though ideally it would have been preferable not to have another consent-based exception, unlike the homologous exception regarding non-sensitive personal data, here it is provided that EU or national laws can decide that the prohibition to process sensitive data "may not be lifted by the data subject."[252]

247 GDPR, art 12(3).

248 If the data controller proves that more time is necessary to respond because the request is very complex and there is a high number of requests, there may be an extension by two further months. See GDPR, art 12(3).

249 Kuner (n 76) 2.

250 Yilun Wang and Michal Kosinski, 'Deep neural networks are more accurate than humans at detecting sexual orientation from facial images' (OSFHome, 15 February 2017) <https://osf.io/zn79k/> accessed 1 March 2018 (forthcoming in Journal of Personality and Social Psychology).

251 GDPR, art 22(4).

252 GDPR, art 9(2)(a).

VI. Data Protection Impact Assessments for algorithmic decisions

76 Lastly, one of the main innovations of the GDPR is the data protection impact assessment (DPIA).[253] These impact assessments are tools for organisations to manage data protection hazards, a form of a form of 'meta-regulation' whereby "state efforts to make corporations responsible and accountable for their own efforts to self-regulate."[254] In this field, DPIAs are "a way of showing that suitable measures have been put in place to address those risks (associated to algorithmic decision-making) and demonstrate compliance with the GDPR."[255] It is commendable that DPIAs are mandatory when a systematic and extensive evaluation of personal aspects is based on automated processing, and on which decisions are based that produce legal effects or similarly significantly affect a natural person.[256] Commendably, DPIAs are required both when the decision is wholly automated and when there is human intervention, not only when it is solely based on automated processing.[257]

VII. Can children be subject to algorithmic decisions?

77 An example of poor drafting regards the algorithmic decision-making concerning children. Hidden in a long recital, one finds the obscure sentence "[s]uch measure should not concern a child."[258]

78 Naturally, one would think that children cannot be subject to algorithmic decisions. However, the sentence follows the one that regards the measures to safeguard the data subject's rights, freedoms, and legitimate interests. Therefore, it may be interpreted as meaning that these safeguarding measures do not apply to children, who could nonetheless be

253 See Article 29 Working Party, 'Guidelines on Data Protection Impact Assessment (DPIA) and determining whether processing is "likely to result in a high risk" for the purposes of Regulation 2016/679' (European Commission, 4 April 2017), <http://ec.europa.eu/newsroom/document.cfm?doc_id=44137> accessed 7 March 2018.

254 This is the theory of Reuben Binns, 'Data protection impact assessments: a meta-regulatory approach' (2017) 7(1) International Data Privacy Law 22, 23. The notion of meta-regulation was developed by C Parker, 'Meta-regulation: Legal Accountability for Corporate Social Responsibility' in D McBarnet, A Voiculescu and T Campbell (eds), *The New Corporate Accountability: Corporate Social Responsibility and the Law* (CUP 2007) 29.

255 Article 29 Working Party (n 9) 27.

256 GDPR, art 35(3)(a).

257 Article 29 Working Party (n 9) 27.

258 GDPR, recital 71.

subject to algorithmic decisions. This is obviously against the purpose of the GDPR, which provides an advanced protection to children. The doctrine of *noscitur a sociis* would lead to absurd consequences; therefore, a purposive approach should prevail. Thus, children should never be subject to algorithmic decision-making.

79 Regrettably, the Article 29 Working Party does not see this provision as an absolute prohibition, since the wording of the recital is not reflected in art 22. However, they recommend that "wherever possible, controllers should not rely upon the exceptions in art 22(2) to justify"[259] algorithmic decision-making affecting children. Nonetheless, such decisions may be necessary for instance to protect the children's welfare, in which case data controllers may resort to the exceptions. Positively, in turn, it is suggested that 'legal effect' and 'similarly significant effect' be interpreted broadly, because "solely automated decision making which influences a child's choices and behaviour could potentially have a legal or similarly significant effect on them, depending upon the nature of the choices and behaviours in question."[260] Similarly, organisations must put in place safeguards tailored to the specific needs and features of the child.[261]

VIII. Collective algorithmic decisions

80 It is unclear, then, what happens to collective algorithmic decisions (e.g. to charge a higher rate of car insurance to the citizens associated to a particular postcode). Indeed, it has been questioned "whether data subjects are protected against decisions that have significant effects on them but are based on group profiling."[262] In general, the stress on the shift from individual to collective privacy should be welcomed.[263] With regards to collective algorithmic decisions, it would seem that art 22 "does not limit 'profiling' as such to individual profiling but only requires that the decision based on such profiling is addressed to an individual, in a way that has legal or significant effects for him/her as an individual."[264] Therefore, collective profiling is covered by the GDPR when used for individual decisions.

259 Article 29 Working Party (n 9) 26.

260 ibid 26.

261 ibid 26.

262 Kamarinou (n 76) 10.

263 Alessandro Mantelero, 'Personal data for decisional purposes in the age of analytics: From an individual to a collective dimension of Data Protection' (2016) 32(2) Computer Law & Security Review 238.

264 Kamarinou (n 76) 10.

IX. Data portability, accountability, and data minimisation

81 Although the focus is on the provisions specifically dedicated to algorithmic decision-making, other rules and principles may affect it. One need only mention data portability, accountability, and data minimisation.

82 The right to data portability could be used to obtain not only information about the logic, significance, and consequences of the algorithmic decision, but also all "the personal data concerning him or her, which he or she has provided to a controller, in a structured, commonly used and machine-readable format."[265] One could use this right to export the profiles used for the algorithmic decision.

83 The principle of accountability, then, may play a positive role. Indeed, in order "to mitigate the risks of automated profiling we must look towards mechanisms that increase the accountability (both through ex ante screening of data mining applications for possible risks and ex post checking of results) and transparency of automated profiling."[266] In particular when relying on the consent-based exception, data controllers will have to document it carefully to prove that consent was explicit.

84 Certain rules should be interpreted broadly, taking into account the characteristics of the phenomenon at hand. For instance, data minimisation and data exclusion, if interpreted narrowly, "may reduce the accuracy of data mining and may deny us the data necessary to detect discrimination in automated profiling."[267] However, the principle of data minimisation means that data should be *adequate*, relevant, and limited to what is *necessary in relation to the purposes* for which they are processed.[268] Arguably, this does not mean that data controllers shall always collect as little data as possible. It means that the quantity must be related to the purpose, provided that the data are adequate. Arguably, the application of artificial intelligence to take decisions that have legal effects can justify the processing of large amounts of data, for at least two interwoven reasons. First, the more data are used to train the algorithm, the more accurate the output may be (big data are 'necessary' for the functioning of artificial intelligence). Second, the processing of a low quantity of data, leading to an inaccurate output, would be 'inadequate' if one has to take a decision with legal consequences (or which similarly significantly affects the individual).

265 GDPR, art 20.

266 Schermer (n 10) 52.

267 ibid 52.

268 GDPR, art 5(1)(c).

X. Algorithmic decisions taken by EU institutions and bodies

85 A brief note, finally, on the algorithmic decision-making carried out by the EU and its institutions and bodies (e.g. e-procurement and e-recruiting). The current rules[269] are more or less the same as the ones laid out in the Data Protection Directive, with the right to be informed about the logic involved in the decision, the right not to be subject to it, and the data controller's obligation to put in place measures to protect the data subject's legitimate interests. The only exception recognised is the express authorisation by national law, EU law, or the European Data Protection Supervisor. In January 2017, the Commission adopted a proposal for a new regulation on the processing of personal data by the EU institutions, bodies, offices, and agencies.[270] The draft provides the same rules as the GDPR as to the information rights (existence, logic, significance, consequences),[271] right to access,[272] right to not to be subject,[273] and mandatory data protection impact assessment.[274]

XI. An overall assessment of the new data protection rules on algorithmic decisions

86 In conclusion, overall the GDPR strengthens the rules on algorithmic decision-making timidly and with some significant flaws, though some positive elements have to be acknowledged. It may well be the case that, as it has been suggested, this regime will act as "legal incentives for technology producers to build accountability mechanisms into the technology."[275] It still holds true that even if Article 15 of the Data Protection Directive and Article 22 of the GDPR show that the promise in terms of providing a counterweight to algorithmic decision-making is tarnished by complexities and ambiguities, they nonetheless shall be regarded as expression of a core data protection principle to be embodied in all data protection instruments.[276]

87 Now, before moving on to the third legal route, one needs to take account of the relation between intellectual property and data protection. It has been shown above that the Software Directive can prevail on the Trade Secrets Directive. It remains to be assessed what happens if there is a clash between trade secrets (and, more generally, intellectual property rights) and the data subject's rights. Under the GDPR, the right of access cannot 'adversely affect the rights and freedoms of others,'[277] which include 'trade secrets or intellectual property and in particular the copyright protecting the software.'[278] However, this provision has been interpreted narrowly by the Article 29 Working Party that observe that intellectual property rights cannot be invoked to deny access or refuse to provide information to the data subject.[279] In allowing the disclosure of an algorithm covered by a trade secret, however, courts shall dictate measures that safeguard the commercial value of the trade secret, for instance by preventing its further disclosure. It is important to note that intellectual property must be balanced with data protection only when it comes to the right of access. Conversely, it is submitted that, in principle, when it comes to the other data subject's rights and data controller's obligations, intellectual property will not be a valid legal basis for exceptions or limitations.

88 Another regime to take into account – and whose interplay with intellectual property and data protection remains partly unsolved – is freedom of information.

269 Regulation (EC) No 45/2001 of the European Parliament and of the Council of 18 December 2000 on the protection of individuals with regard to the processing of personal data by the Community institutions and bodies and on the free movement of such data, arts 13 and 19, recital 29.

270 Proposal for a Regulation of the European Parliament and of the Council on the protection of individuals with regard to the processing of personal data by the Union institutions, bodies, offices and agencies and on the free movement of such data, and repealing Regulation (EC) No 45/2001 and Decision No 1247/2002/EC (hereinafter 'draft regulation on the protection of individuals with regard to the processing of personal data by the Union institutions'). For the Commission proposal, the first reading Position of the European Parliament and the General Approach of the Council, see Council of the EU 13436/17 of 30 October 2017.

271 Draft regulation on the protection of individuals with regard to the processing of personal data by the Union institutions, arts 15(2)(f) and 16(2)(f).

272 Draft regulation on the protection of individuals with regard to the processing of personal data by the Union institutions, art 17(1)(h).

273 Draft regulation on the protection of individuals with regard to the processing of personal data by the Union institutions, art 23.

274 Draft regulation on the protection of individuals with regard to the processing of personal data by the Union institutions, art 39.

275 Chris Reed et al., 'Responsibility, Autonomy and Accountability: Legal Liability for Machine Learning' (Queen Mary School of Law Legal Studies Research Paper No. 243/2016) 29 <https://ssrn.com/abstract=2853462> accessed 1 March 2018.

276 Bygrave (n 165) 22.

277 GDPR, art 15(4).

278 GDPR, recital 63.

279 Article 29 Working Party (n 9) 17.

E. Freedom of information and access to the algorithm. The Italian panorama

89 In 2015, the French *Commission d'accès aux documents administratifs* obliged the *Direction générale des finances publiques* to release the source code of the computer program used to estimate the income tax of natural persons.[280]More recently, the *TAR Lazio*,[281] administrative court[282] in Italy, stated that an algorithm is a digital administrative act and therefore, under the freedom of information regime, the citizens have the right to access it. This section critically analyses this ruling as a prism to understand the application of the freedom-of-information regime to algorithmic decision-making.

90 Under the Italian Administrative Procedure Act,[283] citizens have the right to view administrative documents and extract a copy thereof, if they have a "direct, specific, and actual interest, corresponding to a legally-protected situation and linked to the document one intends to access."[284] The typical example would be an individual unhappy with the outcome of a public competition (e.g. to become notary public) and, therefore, demands to access the documents relevant to the competition. An important limitation of freedom of information regimes is that they can be actioned only against the State or other public bodies and with regards to administrative documents.[285] The Government and the public bodies can lay out which documents cannot be accessed for a number of purposes listed in the Administrative Procedure Act, including privacy.[286] However, there is case law clarifying that in principle, if the right to access and privacy clash, the former shall prevail, at least in the sense that an access request will not be denied for privacy reasons, but the document may be anonymised.[287] More recently and generally, it has been stressed that freedom of information is a fundamental right and, therefore, the denial to access requests are allowed only in exceptional instances.[288] This approach can also be found in the Privacy Code[289], in which there is a right to access administrative documents even though they contain personal or even sensitive data, because the freedom of information regime "is deemed to be of relevant public interest."[290] The balance is struck slightly differently when it comes to data on health or sexual life. Indeed, the access request will only be accepted if the interest underlying the request is a personality right[291] or other fundamental right or freedom.[292] One may infer that normally the right to access prevails over opposite interests and rights, even in the event the opposite rights were fundamental, unless the computer program implementing the algorithm processes health data or data about the sexual life of the individual. Thus, it is submitted that also the potential clash between freedom of information and intellectual property should normally be resolved in favour of the former. The GDPR will not affect the balance between privacy and freedom of information, since the recently presented draft implementing decree clarified that access to administrative documents and civic access fall outside the scope of the GDPR, at least in the context of its Italian implementation.[293]

91 Only individuals who have a specific, direct, and actual interest in the access to the administrative document can exercise the right of access under the Administrative Procedure Act. However, in 2016, Italy introduced a more general freedom of information regime. Under the Citizen Access Act,[294] the individual has two rights. First, the right

280 Commission d'accès aux documents administratifs, avis 20144578 - 8 January 2015, <http://www.cada.fr/avis-20144578,20144578.html> accessed 1 March 2018.

281 TAR Lazio, chamber III bis, 22 March 2017, No 3769.

282 These courts administer justice mainly when a citizen claim the violation of their legitimate interest by a public body.

283 *Legge* 7 August 1990, No 241 *Nuove norme in materia di procedimento amministrativo e di diritto di accesso ai documenti amministrativi* (hereinafter 'Administrative Procedure Act'), Articles 22-28. *Decreto del Presidente della Repubblica* 12 April 2006, No 184 *Regolamento recante disciplina in materia di accesso ai documenti amministrativi*.

284 Administrative Procedure Act, art 22(1)(a).

285 Administrative documents are defined in a very broad way, that is "every graphical, photographical, electromagnetic representation (or any other kind of representation) of the content of documents - be they even internal or not related to a specific administrative procedure – which are in the possession of a public body and concern public interest activities, being immaterial the public or private nature of the relevant regime" (Administrative Procedure Act, art 22(1)(d)). For an even broader definition see *Decreto del Presidente della Repubblica* 28 December 2000, No 445 *Testo unico delle disposizioni legislative e regolamentari in materia di documentazione amministrativa*, art 1(1)(b).

286 Administrative Procedure Act, art 24(6)(d). See, for instance, *Regolamento del Comune di Salerno sull'accesso agli atti e sulla tutela della riservatezza dei dati contenuti in archivi e banche dati comunali*, art 5(2)(m).

287 Consiglio di Stato, chamber V, 28 September 2007, No 4999; Consiglio di Stato, chamber VI, 20 April 2006, No 2223; Consiglio di Stato, plenary session, 18 April 2006, No 6.

288 TAR Toscana, chamber I, 10 February 2017, No 200.

289 Decreto legislativo 30 June 2003, No 196, *Codice in materia di protezione dei dati personali* (Privacy Code).

290 Privacy Code, art 59.

291 By personality rights, it is meant rights, such as life and honour, that are absolute and refer to fundamental aspects of the human being. This is a civil law notion, which should not be confused with the common law one, where personality rights are the rights to control the commercial use of one's own name or other aspects of one's identity (name, likeness, etc.).

292 Privacy Code, art 60.

293 Draft GDPR implementing decree, art 55.

294 *Decreto legislativo* 14 March 2013, No 33 *Riordino della*

to access all documents, information, and data (not only administrative documents), if there were an obligation to publish them and the relevant public body infringed it by not publishing.[295] This right (so-called citizen simple access) is absolute and an access request under this provision cannot be denied.[296] Second, a right to access documents that the State or other public bodies are not obliged to publish, justified with the purpose to "favor a generalised control over the pursuit of the institutional functions and over the use of public resources, as well as to promote the participation to the public debate."[297] This citizen generalised access is a limited right.[298] Indeed, the relevant request can be denied for a number of reasons,[299] including data protection[300] and intellectual property.[301]

92 There is another regime that may be used to access algorithms used by the State and other public bodies, even though its scope is very narrow. As of 14 September 2016, under the Public Administration Code,[302] legal and physical persons have the right to reuse computer programs and other "solutions" in order to "adapt them to their needs".[303] Therefore, the State or other requested public body have an obligation to make the relevant source code publicly available "alongside the documentation"[304] under a free and open-source license. However, the requested body can deny access in three scenarios if the computer program or the solution owned by the State or public body were not developed "based on the specific indications by the public customer."[305] The denial may be justified also by *ordre public*, national security, defence, and elections.[306]

93 Let us focus on the recent case that applied the Administrative Procedure Act in order to recognise the right to access the source code of the computer program implementing the algorithm used by the Ministry of Education, University and Research with regards to the mobility of the teaching staff; the algorithm had been commissioned to a private company (HPE Services s.r.l.). The teachers' trade union claimed that they could not defend their members' right with regards to the mobility procedures if they were not allowed to access the algorithm. The computer program was used to manage the transfer of the teaching staff between provinces and the outcome of the procedure was solely determined by the algorithm. This means that, should the requirements be met (personal data, decision with legal effect, etc.), the applicant may exercise the rights recognised by the GDPR with regards to algorithmic-decision making.[307]

94 In the case at hand, the applicant sought to exercise the right to access under the freedom of information regime. However, this was denied by the Ministry of Education for a number of reasons. Firstly, the source code was not an administrative document (and the right to access under freedom of information can be exercised only with administrative documents).[308] Secondly, the computer program was covered by copyright. The court, however, dismissed both arguments.

disciplina riguardante il diritto di accesso civico e gli obblighi di pubblicità, trasparenza e diffusione di informazioni da parte delle pubbliche amministrazioni (Citizen Access Act), as amended by the *Decreto legislativo* 25 May 2016, No 97 *Revisione e semplificazione delle disposizioni in materia di prevenzione della corruzione, pubblicità e trasparenza, correttivo della legge 6 novembre 2012, n. 190 e del decreto legislativo 14 marzo 2013, n. 33, ai sensi dell'articolo 7 della legge 7 agosto 2015, n. 124, in materia di riorganizzazione delle amministrazioni pubbliche* (Prevention of Corruption Act).

295 Citizen Access Act, art 5(1).

296 Unless the public body proves that there was no obligation to publish or that the document, information or data is already published.

297 Citizen Access Act, art 5(2).

298 Most European jurisdictions have similar provisions. In the UK, the Freedom of Information Act 2000, that covers all recorded information held by a public authority (Information Commissioner's Office, *Freedom of Information Act Awareness Guidance No. 12*). However, an access may be denied for a number of reasons, including trade secrets and other commercial interests (Freedom of Information Act 2000, Section 43). It is notable that, unlike other commercial interests, if the algorithm is covered by a trade secret, the access request may be denied without considering whether or not the release may cause harm (Information Commissioner's Office, *Freedom of Information Act Awareness Guidance No. 5*).

299 Citizen Access Act, art 5 *bis.*

300 Citizen Access Act, art 5 *bis* (2)(a).

301 Citizen Access Act, art 5 *bis* (2)(c).

302 *Decreto legislativo* 7 March 2005, No 82 *Codice dell'amministrazione digitale* (Digital Administration Code), as amended by *Decreto legislativo* 26 August 2016, No 179 *recante "Modifiche e integrazioni al. Codice dell'amministrazione digitale, di cui al decreto legislativo 7 marzo 2005, n. 82, ai sensi dell'articolo 1 della legge 7 agosto 2015, n. 124, in materia di riorganizzazione delle amministrazioni pubbliche".*

303 Digital Administration Code, art 69(1).

304 Digital Administration Code, art 69(1). The wording is very vague; it is likely to refer primarily to all the documentation necessary to adapt the computer program to the applicant's needs.

305 Digital Administration Code, art 69(1).

306 In the UK, there is the right to access datasets for reuse and it is broader than the Italian regime, because it regards all copyright works (Freedom of Information Act 2000, s 11A).

307 Currently, in Italy, the Privacy Code does not regulate algorithmic decision-making. The GDPR, being a regulation as opposed to a directive, will play an important role in strongly harmonising the relevant national regimes, in some instances by innovatively recognising the right to be informed about and object to algorithmic decision-making (e.g. in Italy), in others by updating the existing regime (e.g. in the UK). As seen above, the implementing measures of said countries seem to partly or completely overlook the matter.

308 Administrative Procedure Act, art 22(1)(d).

95 Given that, with the current development of AI and kindred technologies, public bodies can increasingly replace human procedures with algorithmic ones, the court held that the use of the algorithm cannot act as justification for restricting the scope of application of the freedom of information regime. Let us imagine what would happen if all procedures were handled by algorithms and the freedom-of-information requests were not applicable to algorithmic documents: the said regime would still exist in the books, but no longer in practice.

96 The conceptual first step is recognising the existence of the concept of a digital administrative document. In a digital administrative document, an algorithm replaces a human agent acting on behalf of a public body; this is allowed only with regards to the non-discretionary administrative activities.[309] Indeed, non-discretionary power is compatible with the way computer programs work, because the latter can translate facts and legal data into code, thus bringing to an immutable conclusion through formalised reasoning.[310] This passage of the ruling reinforces this paper's argument that algorithms cannot replace human judges (and other decision-makers) because interpretation is ubiquitous and it is an intrinsically discretionary process.

97 This said, the court needed to qualify the computer program itself as a digital administrative document, otherwise no access to the source code could be granted (at least under this regime). The computer program qualifies as a digital administrative document because it materialises the ultimate will of the public body in a way that is able to create, modify, or extinguish the individual's legal positions. Consistently with the technology neutrality principle, the relevant statutory provision describes the 'administrative document' in a very broad way by encompassing also the electromagnetic representation of a document and any other form of representation.[311] Therefore, there is no problem in considering a computer program implementing an algorithm as an administrative document (if the other legal requirements are also met).[312] It may be conceded that, strictly speaking, a computer program is not a document in itself. However, recognising the right to access only to the final document resulting from the algorithmic procedure would equal denying the access request, because without the source code it may prove hard to understand the rationale of the final decision. The right to access often serve the purpose of lodging a complaint against a public body if the final decision affected the individual's rights or legitimate interests. However, it is unlikely that such a claim would be successful, if the individual does not have access to the rationale of the final decision (which means also accessing the source code, if the decision is algorithmic). Indeed, it is believed that a narrow interpretation of an 'administrative document' would not comply with the right to an effective remedy and to a fair trial as enshrined in the Charter of Fundamental Rights of the EU[313] and in the European Convention of Human Rights.[314]

98 One may object that granting the access in this case would be tantamount to granting access to the source code of the computer program (e.g. Microsoft Word) used to write an administrative document. Such an argument would be based on a wrong understanding of what is a digital administrative document. Indeed, the court distinguishes between documents drafted with the aid of a computer and electronically programmed documents, where the software finds and links data and norms. The latter is a digital administrative document (the source code of which is accessible) because it constitutes the final decision; it is not a mere aid to draft it.[315] This paper joins those who underline that "the electronic processing *is* the document, it represents it, it makes it known externally, it becomes the form of the document, thus being legally relevant *in its electronic form*, regardless of its paper transcription."[316] The

309 In Italy, the discretional power of the public administration is a fundamental principle, whereas only in a limited number of instances the State or other public body take a non-discretionary decision (with the content as well as the requirements rigidly predetermined by the law), for instance when an authorisation shall be released as a necessary consequence of the positive assessment of the existence of certain requirements. Some authors affirm that administrative power is always discretionary (e.g. Fabio Massimo Nicosia, *Potere ed eccesso di potere nell'attività amministrativa non discrezionale* (Jovene 1991), but this theory is not widely accepted (e.g. Paola Rossi, *Il riesame degli atti di accertamento* (Giuffrè 2008)).

310 The Italian Court of Cassation defined the digital administrative document in a narrow way by including only those documents which are directly and automatically processed from the computer, in as much as they do not require discretional assessments and argumentations linked to the specificities of the case at hand (Corte di Cassazione, chamber I, 28 December 2000, No 16204).

311 Administrative Procedure Act, art 22(1)(d).

312 In particular, the document must be in a public body's possession and it must regard public interest activities (Administrative Procedure Act, art 22(1)(d)). It is immaterial if the algorithm was developed as a consequence of contract (a private law tool), as long as the relevant activity is of public interest, which is the case here, given that the purpose of the program is to improve the management of a public service (education).

313 Art 47.

314 Art 6, art 13.

315 Contrary to what was held by the court, some scholars affirm that only the administrative document drafted with the aid of a computer is a digital administrative act. See Alfonso Contaldo and Luigi Marotta, 'L'informatizzazione dell'atto amministrativo: cenni sulle problematiche in campo' (2002) 18(3) Diritto dell'informazione e dell'informatica 576.

316 Massimiliano Minerva, 'L'attività amministrativa in forma elettronica' (1997) 4 Foro amministrativo 1300, italics

very broad definition of administrative document is seen by the court and by legal scholars as a shift from a focus on the pedigree of the document, to its function:[317] if the function is administrative (as in concerning the public interest), then it is immaterial how the document was formed and access shall be granted in any event, if the general requirements are met. This said, it is important to stress that the court stated that electronically programmed documents are not allowed when it comes to the exercise of discretionary power,[318] due to the difficulty "which is scientific as opposed to legal, to map the reasoning underlying the document,"[319] if this is the outcome of an algorithmic procedure (and not simply drafted by a human being with the aid of a word processor). Again, there is no place for algorithmic decisions where the relevant process is discretionary.[320]

99 After recognising the right to access the computer program, the court went on to state that providing the applicant with the mere description of the algorithm and of its functioning is not a sufficient response.[321] Only the access to the source code is. Indeed, the Ministry of Education had responded to the access request by describing the algorithmic procedure (collection of input data, appointment to a certain school, distribution of the results), as well as reporting some case studies. The court, however, states, "the assessment of the functionality of the algorithm or of programming errors can be carried out exclusively in light of the knowledge"[322] of the source code. This should be accompanied by a thorough explanation of the rationale and of the consequences of the decisions, especially if personal data is involved.

100 Finally, as to the clash with the copyright on the computer program, the steps to follow are: i. Assessment of copyright subsistence; ii. Authorship and ownership; iii. Infringement; iv. Exceptions.

101 The subsistence, authorship, and ownership of the copyright do not seem to be problematic.[323] Even though there is no evidence on the point, the court assumes that the Ministry of Education owns the program under a license with HPE Services s.r.l., which retains authorship and the moral rights.[324]

102 The court goes on to observe that the purpose of the access does not conflict with the economic interest protected by copyright.[325] On this point, the court is not clear as to whether it is dealing with the assessment of infringement or with the exceptions. In the latter event, this would be a peculiar ruling, because it would take a flexible "fair use"[326]-like approach to copyright exceptions,

added.

317 Carmelo Giurdanella and Elio Guarnaccia, *Elementi di diritto amministrativo elettronico* (Halley 2005) 24.

318 Most scholars agree, see e.g. Contaldo (n 312) 580.

319 TAR Lazio, chamber III bis, 22 March 2017, No 3769. This idea was first expressed by A Ravalli, 'Atti amministrativi emanati mediante sistemi informatici: problematiche relative alla tutela giurisdizionale' (1989) 2 Trib. Amm. Reg. 261. The traditional theory that presents a dichotomy discretionary-non-discretionary and allows algorithmic decisions (or electronically processed administrative documents) only with regards to the latter is open to criticism. However, this is not because, as Ravalli thinks, even discretionary administrative activities are rational logical processes based on predetermined criteria (which is debatable). The point is that interpretation is always discretionary and even non-discretionary power is exercised through interpretation (given that the dichotomy interpretation-application is untenable, as shown by Hart; see Viola (n 29) 50).

320 This passage may be interpreted as the court espousing that line of thought whereby the admissibility of algorithmic decisions (or electronically processed administrative documents) depends not on the nature of the power, but to the scientific possibility to map the reasoning underlying the document (Giurdanella (n 314) 32; Michele Corradino, 'Inquadramento generale dell'atto amministrativo elettronico' (Convegno DAE 2004). However, before referring to the importance of the said scientific possibility (or the lack thereof), the court is adamant in reaffirming the old contraposition. Indeed, the court states that "we can easily agree that administrative documents which are the output of an algorithmic procedure are admissible with regards to the non-discretionary activity of the public bodies" (TAR Lazio, chamber III bis, 22 March 2017, No 3769). It then goes on to observe that "it is evident that different considerations apply to the discretionary activities" (ibid.). The reference to the fact that the admissibility of this kind of digital administrative act does not depend on the qualification of the activity as discretionary (but it would depend on the possibility of mapping the underlying reasoning) is introduced by a dubitative form ("it may be possible to assume that") and it seems an obiter dictum. One may infer this by the observation that "we believe that we can disregard the exam of this very interesting legal question" (ibid.).

321 Some believe that the description of the algorithm could solve the problem of making the citizen understand the software used by the public body. See Daniele Marongiu, 'Gli atti amministrativi ad elaborazione elettronica: la compilazione di un "pre-software" in lingua italiana' (Quaderni del DAE 2003) <http://www.cesda.it/quadernidae/pdf/MARONGIU_DAE2003.pdf> accessed 1 March 2018.

322 TAR Lazio, chamber III bis, 22 March 2017, No 3769.

323 The court accepts the Ministry of Education's allegations on the point, because there are no elements that may suggest that there is no copyright on the computer program at hand.

324 Transactions regarding moral rights (e.g. paternity waivers) are not enforceable under Italian copyright law (*Legge* 22 April 1941, No 633 *Legge a protezione del diritto d'autore e di altri diritti connessi al suo esercizio* (Copyright Act), art 22).

325 In Italy, the author of a copyright work has the exclusive right to use the work for economic purposes (Copyright Act, art 12(2)).

326 This is the doctrine of copyright exceptions in the US. It does not revolve around a list of permitted uses, but it is a flexible principle that enables the judge to assess all the

usually interpreted by applying the so-called three-step test, revolving around an exhaustive list of permitted uses.[327] There is currently no copyright exception for non-commercial use or for purposes of freedom of information. The access to the source code for this purpose may not conflict with the normal exploitation of the work and may not prejudice the interests of the author. However, the third step requires that the exception be expressly provided by the law, which currently does include a general exception for non-commercial acts. Conversely, the point should be better construed as meaning that there can be no infringement because the restricted act is not the distribution of the copyright work, but its distribution for commercial purposes. Indeed, the heading of the chapter of the Copyright Act on the restricted acts is "Protection of the economic use of the work"[328] and the first relevant provision recognises the "exclusive right to economically use the work within the limits of the Act."[329] From this perspective, the clash between freedom of information and copyright is merely ostensible, because the right to access administrative documents does not interfere with the uses of computer programs that are restricted by the law. Additionally, a different conclusion would have led to an unacceptable difference of treatment depending on the technological solution adopted. It is obvious that, in principle, public bodies own copyright on the documents they produce. However, it would be absurd to claim that a freedom of information request can be denied because the public body owns the relevant copyright. This would equal sterilising the right to access. Accordingly, the discretional adoption of a more modern technology cannot justify different considerations. Therefore, just like copyright could never be the basis of an access denial under the analysed regime, it will never justify the access denial with regards to computer programs.

103 An argument of the Ministry of Education was, then, that the so-called citizen generalised access request can be denied if necessary to avoid an actual prejudice to intellectual property.[330] However, the right to access under the Administrative Procedure Act (which is the one relevant here) and the citizen access are entirely different things. Their purposes are discrete. The former does not encompass a right to a generalised control over the public bodies:[331] it serves the purpose to enable the individuals to defend their rights and interests which may be affected by an administrative document. This generalised control, conversely, is the purpose of the citizen access rights under the Citizen Access Act. The requirements of the right to access administrative documents and the citizen access rights (both simple and generalised) are different; therefore, all the remedies can operate in parallel. The balance will have to be struck differently. On the one hand, the former requires access to more detailed information, because it serves the purpose of preparing a claim. On the other hand, under a citizen access regime, even less granular information will be sufficient (e.g., the description of the algorithm may suffice under this regime). The court states that, therefore, it may be that whereas a citizen access is denied, it may be accepted with regards to the same document if the same individual exercises the right to access administrative documents.

104 It is submitted that the court may have brought into play three more considerations. First and foremost, *ubi lex voluit dixit, ubi noluit tacuit*. The lawmaker expressly accepts that an access request can be denied for intellectual property purposes under the citizen access regime. However, the fact that the legislator does not provide a similar exception with regards to the right to access administrative documents constitutes evidence of the untenability of an intellectual property exception to the said right. Second, intellectual property is mentioned in the citizen access regime as an example of "economic and commercial interests."[332] Therefore, since it has already been proven that the access to the source code would not conflict with the use of the program for commercial purposes, even if the exception were extended to the right to access administrative documents, it would not apply in the case at hand. Third, the exceptions to the citizen access are allowed only if "necessary to avoid an actual prejudice" to the listed interests (including intellectual property). Arguably, denying access to the source code may not always be necessary to avoid such prejudice (for instance, if the applicant agrees to make a non-commercial use of it). Given that there is no intellectual property exception to the right to access administrative documents, one should bear in mind that also trade secrets and patents might not be used to prevent the said access. This is particularly important from our perspective, given the pivotal

circumstances of the case to ascertain whether the use of a copyright work was fair.

327 The exception must fall within the exhaustive list of the Copyright Directive (Directive 2001/29/EC of the European Parliament and of the Council of 22 May 2001 on the harmonisation of certain aspects of copyright and related rights in the information society), not conflict with the normal exploitation of the work and does not unreasonably prejudice the legitimate interests of the author (art 5(5)).

328 Copyright Act, *Titolo* I, *Capo* III, *Sezione* I.

329 Copyright Act, art 12(2). See arts 64 *bis* – 64 *quater* for the specific provisions on computer programs (the general rule on the economic use, however, applies also to computer programs).

330 Citizen Access Act, Article 5 *bis* (2)(c).

331 Administrative Procedure Act, Article 24(3). See, for instance, Consiglio di Stato, chamber V, 25 September 2006, No 5636.

332 Citizen Access Act, art 5 *bis* (2)(c).

role of trade secrets in keeping algorithms opaque.

105 As a consequence of the lack of the elements of infringement, of the inexistence of an intellectual property exception to the right to access administrative documents, as well as of the general assertion whereby "the nature of copyright work does not represent a justification for access denial,"[333] the court recognises the right to access the source code, provided that the applicant uses the information exclusively for the purposes that legitimised the claim (the right of the teachers' trade union to defend its members' rights).

106 For all the reasons analysed above, the court found in favour of the teachers' trade union and, therefore, annulled the access denial and ordered the Ministry of Education the release of a copy of the source code of the computer program implementing the algorithm used by the Ministry in handling the teachers' mobility.

107 The right to access administrative documents may be seen as a weak tool when it comes to the transparency of the algorithmic decisions taken by the State and other public bodies. Indeed, especially in AI / black box contexts, accessing the source code of the computer program implementing an algorithm does not provide the applicant with valuable and / or intelligible information.[334] However, denying such access would conflict with the fundamental right to an effective remedy, because an individual could hardly be successful in a claim against a public body, if they cannot access the rationale of an algorithmic decision affecting their rights and legitimate interests.

108 Some scholars suggest that, in the future, artificial intelligence will be used to adopt algorithmic administrative documents even when it comes to discretionary activity, with the possibility of leaving room for the human intervention in the most difficult cases.[335] They maintain that this is only a prediction but given the current developments of natural language processing and machine learning, arguably the relevant tools are already available. Even though it cannot be said that artificial intelligence should be banned altogether when it comes to discretionary power, it is believed that some room for *ex-ante* human intervention should always be left for a number of reasons, including the fact that all administrative activities (like all interpretive operations) are to some extent discretionary. This does not mean, however, that citizens cannot exercise the right to access under the freedom of information regime if the relevant administrative activity is non-discretionary. It means that public bodies are not allowed to use AI when they are exercising a discretionary power.

109 The question remains as to what citizens can do if public bodies start taking decisions against them even in the discretionary realm. The remedy described in this section operates *ex post*, once the decision has already been taken. Similarly, the copyright and patent exceptions may constitute a useful *ex-post* tool, but their scope is quite limited. From an *ex-ante* perspective, however, it may be argued that a potentially affected individual could obtain an injunction to prevent a public body from taking an algorithmic decision by using the data protection tool under Article 22 of the GDPR. Therefore, an integrated approach to the remedies against algorithmic decisions should be taken.

F. Conclusions

110 This study presented ten arguments against algorithmic decision-making, as well as three routes available to those affected by algorithms. As pointed out by some scholars,[336] the most important thing is providing individuals with the means to challenge adverse algorithmic decisions. To do so, intellectual property, data protection, and freedom of information provide adequate protections, particularly if one takes an integrated approach. National implementations of the GDPR should be a precious opportunity to detail the procedures to challenge algorithmic decisions, even though it does not seem that this is the direction that is being taken.

111 Intellectual property enables the legitimate user of a software implementing an algorithm or of an algorithm-related patent to carry out certain acts (study, observation, etc.) without the intellectual property owner's consent. Whilst these quasi-rights allow the user to try and understand the algorithm by themselves, they do not give them a positive right to demand the intellectual property owner's cooperation.

333 TAR Lazio, chamber III bis, 22 March 2017, No 3769.

334 It is not a coincidence that the applicant is not an individual, but a trade union, which is likely to have the resources to make sense of a source code. The fact that a lay person could hardly understand a source code has been used as an argument against the recognition of computer programs as digital administrative documents. However, the court points out that the choice of an innovative tool cannot deprive the citizens of the right to access administrative documents and that, anyway, the applicant may avail themselves of the collaboration of an IT person to decipher the code.

335 Giurdanella (n 314) 33, referring to Giovanni Sartor, *Le applicazioni giuridiche dell'intelligenza artificiale* (Giuffrè 1990) and Giovanni Sartor, 'Gli agenti software: nuovi soggetti del cyberdiritto?' (2002) Contratto e impresa 465.

336 Keats (n 4) 1.

112 Conversely, a freedom of information request allows all citizens to impose upon public bodies, under certain circumstances, an obligation to release the source code of computer programs that implement algorithms, while explaining the logic involved in the relevant decision. The main shortcoming of this regime is the limitation to public defendants. Much will depend on how courts will strike a balance between freedom of information and intellectual property. In Italy, the former prevails. In turn, arguably, the UK tend to favour the interests of the intellectual property owners.

113 The only *ad-hoc* regime against algorithmic decisions is provided by art 22 of the GDPR. One may criticise some aspects of this provision. For instance, it applies only to decisions "*solely* based on automated processing" means. This paper's suggestion is to recognise the right not to be subject to an algorithmic decision every time that there is not a human being taking the final decision substantially, as opposed to formally. In spite of its shortcomings, art 22 is clear and detailed in laying out the general principle that businesses, governments, judges, and other data controllers should not make decisions based solely on algorithmic processes. Under certain circumstances (e.g. explicit consent), such decisions can be made, but informing the data subject and allowing him or her to access to the logic involved in the decision, its significance, and the envisaged consequences. Much will depend on the national implementing measures. The UK Data Protection Bill risks not ensuring compliance with the GDPR, thus exposing the UK to the possibility of being considered as 'inadequate' in the context of cross-border EU-UK data transfers.

114 It is submitted that only a document which includes *both* the algorithm used and an explanation of the logic and consequences in non-technical terms would comply with the GDPR as interpreted in light of the Charter of Fundamental Rights of the EU and the European Convention on Human Rights. Then, the right to a human judge is paramount, because the right to access and to be informed may prove useless. Indeed, when artificial intelligence is used, it is sometimes unfeasible to access the relevant rationale. To the legal black box created by intellectual property rights, one needs to add the technical black box and the organisational one.

115 Practically, if the algorithmic decision is based on personal data, this latter route is preferable. If not and the decision-maker is a public body, one should opt for a freedom of information request. If a private decision-maker (e.g. a bank) makes an algorithmic decision based on non-personal data, then the route will be that of intellectual property exceptions. The freedom-of-information remedies operate *ex post*, once the decision has already been taken. In turn, the copyright and patent exceptions may be used before any decision is made, but only to access the algorithm, not to prevent the decision-maker from proceeding algorithmically. The only regime that prevents algorithmic decisions is the one provided by the GDPR.

116 The trust in artificial intelligence and algorithms derives from the belief that non-human agents are unbiased, and their decisions are not affected by passions and ideologies. In fact, algorithms are as biased as the people who trained them, but in a less transparent and accountable way. The more important algorithms will become, the more we will want them to embed our values (and, therefore, our ideologies and biases).[337] Further research should be carried out by diverse (also in terms of gender, ethnicity, etc.) multidisciplinary teams in order to find solutions to open the technical, organisation, and legal black boxes and to ensure fair algorithmic decision-making. Indeed, only a strong humanist stance will be able to reduce algorithmic bias.

117 This paper is a humanist manifesto. It is, indeed, permeated with the belief that we should trust our fellow human beings over the algorithms, despite developments in artificial intelligence allowing the deployment of increasingly refined legal applications. This does not mean that we should reject the use of algorithms altogether. For instance, judges shall use them to improve the quality and consistency of their decisions. However, they shall not let algorithms decide in their stead. In order to better understand how to make the human-algorithm cooperation work best, it has become crucial to shift the focus from the definition of algorithms, artificial intelligence etc. to the understanding of what makes us human.[338]

Acknowledgements

The author is grateful to Sue Farran, Paul Dargue, and Tony Ward for comments on previous drafts of this article. This work has greatly benefited also from the feedback received at the «XXVIII World Congress of Philosophy of Law and Social Philosophy» (Lisbon, 17 July 2017), at the «Café & Chat: Quem Governa os Algoritmos?» (IRIS - Instituto de Referência em Internet e Sociedade

337 The trends of data protection by design, ethics by design, etc. may be explained as a tendency to anthropomorphise non-human agents. See, for example, British Standard Institution, *Robots and robotic devices. Guide to the ethical design and application of robots and robotic systems* (BSI 2016); *cf* Guido Noto La Diega, 'The European strategy on robotics and artificial intelligence: too much ethics, too little security' (2017) 3(2) European Cybersecurity Journal 6.

338 For a positive step in this direction, see Bett M Frischmann and Evan Selinger, *Re-engineering humanity* (CUP 2018).

and GNET – UFMG, Faculdade de Direito da Universidade Federal de Minas Gerais, Belo Horizonte, 18 August 2017), at the research seminar organised by NINSO The Northumbria Internet & Society Research Interest Group (Newcastle upon Tyne, 8 December 2017), and at a guest lecture given at the Schmalkalden University of Applied Sciences (Schmalkalden, 21 December 2017). Thanks to Katie Atkinson and Giulia Caffarelli for the insight into the AI debate. The author's appreciation goes to the anonymous referees for the helpful comments and to Philipp Schmechel for the patient editing. Views and errors are solely the author's responsibility.

Informing Consent

Giving Control Back to the Data Subject from a Behavioral Economics Perspective

by **Santiago Ramírez López***

Abstract: The development of data privacy legislation in Europe and America has been highly influenced by the idea that individuals must maintain the autonomy to take decisions regarding the general purpose and uses of their personal data; an idea that has been generally instrumentalized with the mechanism of informed consent. Recently, both companies and researchers in the field have criticized this idea, arguing that with the new advances and technological progress, consent has lost importance due to the ubiquity of the data processing and the absence of real participation of the data subjects. This article seeks to take into account both points of view, by recognizing the importance of the autonomy of individuals to determine the destination of their personal data, but also by understanding the practical implications and the impossibilities derived from obtaining an informed consent from data subjects that are generally unfamiliar with the topic. Based on the analyses regarding the difficulties of obtaining an effective and informed consent, this contribution will examine how some of the bias and impasses studied through the discipline of behavioral economics may help us to understand the current problems in relation to the way in which consent is requested and provided by the data subjects. This contribution concludes by proposing alternatives that seek to overcome these biases and impasses with an easier provision of information of the data processing and the implementation of a data management and a value-oriented model, which would benefit the data subjects.

Keywords: Behavioral economics; Data Privacy; Data Protection; General Data Protection Regulation; Informational self-determination; Informed consent

© 2018 Santiago Ramírez López

Everybody may disseminate this article by electronic means and make it available for download under the terms and conditions of the Digital Peer Publishing Licence (DPPL). A copy of the license text may be obtained at http://nbn-resolving.de/urn:nbn:de:0009-dppl-v3-en8.

Recommended citation: Santiago Ramírez López, Informing Consent: Giving Control Back to the Data Subject from a Behavioral Economics Perspective, (2018) JIPITEC 35 para 1.

A. Introduction

1 It is safe to say that the notion of controlling the destination of one's personal information has been strongly involved in the development of the data privacy/data protection discipline.[1] The right to informational self-determination, developed in Germany during the 1980's, entails a value that is still applicable in recent history; that individuals should be able to limit the information that can be used from them.[2] The American tradition has long

* LLB (Del Rosario University - Colombia); LLM (University of Hannover / University of Oslo). IT Law Associate Professor (El Bosque University – Colombia).

1 Professor Lee A. Bygrave has a thoughtful definition of the field of data privacy law, and the meeting points and dissimilarities between different terms that compose the field, such as "data protection", "data privacy" and "data security". For conceptual purposes, this work will indistinctively use the terms "data protection" and "data privacy" to address, in the words of professor Bygrave, the regulation of "(...) all or most stages in the processing of certain kinds of data" as well as "(...) the ways in which the data is gathered, registered, stored, exploited and disseminated". See: Bygrave, L. A. (2014). *Data Privacy Law: An International Perspective*. London: Oxford University Press, pp. 1-5.

2 Schwartz, P. (1989). The Computer in German and American Constitutional Law: Towards an American Right of

recognized equivalent values, identifying the right "to be let alone" as a way to promote a non-intrusive information gathering against the media and the emerging technologies,[3] and as a right grounded in the control of the individual to determine what information can be openly communicated.[4]

2 On paper, the European legislation has included several provisions which seem to provide data subjects more control over their information. Indeed, the former Data Protection Directive[5] (from now on the "DPD") and the recently adopted General Data Protection Regulation[6] (from now on the "GDPR"), contain the basis to prevent practices that may constitute an illegitimate processing of data, and dedicate several provisions to the possibility of control of the data subject grounded on informed consent. Nevertheless, the reality of the online scenario has shown the inability of this model to provide control and to protect the data subject's right to privacy.

3 Among experts, there is a debate if whether providing more control to the data subject can be a solution applicable in the real world for the protection of the right to privacy. The critics of a control-oriented approach base their arguments on the practical, conceptual, and moral difficulties of the model,[7] but mainly on the fact that the consent, as the main mechanism of control of the data subject, has so far proved to be impractical and inefficient.[8]

4 A concise reason for the failure of a consent-oriented model is still subject to debate. Some, especially in the private sector, believe that modern society is currently suffering a transition, where the traditional concept of privacy, or privacy as a "social norm," is being dismissed with the excuse of an interconnected world.[9] Although this explanation is convenient for businesses, it also disregards the fact that society has become increasingly more suspicious of how companies are using the data and information provided by users.[10]

5 This article aims to demonstrate that, although the legal concepts of privacy between the predominant Western traditions contain discrepancies mainly in their formation, they are not so different in their outcome, as Western traditions embrace the concept of control of the data subjects as a capital guideline of data protection.

6 This paper will analyze the implications of the field of behavioral economics in the data privacy scenario. Supporting the position of other authors,[11] it will be argued that some of the bias and impasses studied in the field of behavioral economics may help to explain the issues and problems of consent as a way to provide control to the data subject based on a conscientious decision-making scenario.

7 The objective of this analysis is to restore the position of the concept of informed consent as the primary means of control for the data subject, while recognizing that to achieve such informed consent, the data subjects must be provided with more suitable conditions that allow them to overcome the biases and impasses.

8 As a conclusion, this contribution will analyze a proposal for the creation of such a suitable scenario, by implementing alternative ways to provide and manage information and by giving a tangible value to the data from the user's perspective. This proposal is composed of the following components: (i) alternative and user-friendly ways of providing the information required by Article 13 of the GDPR, resorting to existing models, such as Creative Commons; (ii) a data management system that contains unified information of the personal data circulating online of the data subject; and (iii) a model based on the value of the data in benefit of the

Informational Self-Determination. *The American Journal of Comparative Law*, 37, pp. 678-689.

3 Warren, S., & Brandeis, L. (1890, December 15). The Right to Privacy. *The Harvard Law Review*, IV(5).

4 *Ibid.*

5 The European Parliament and the Council of the European Union. Directive 95/46/EC of the European Parliament and 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.

6 The European Parliament and the Council of the European Union. Regulation (EU) 2016/679 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).

7 Allen, A. L. (2000). Privacy-as-Data Control: Conceptual, Practical, and Moral Limits of the Paradigm. *Connecticut Law Review*, 32, 861-875.

8 Koops, B.-J. (2014). The Trouble with European Data Protection Law. *Tilburg Law School Legal Studies Research Paper Series*, 4, p. 3.

9 In 2010 with the rapid increase in the use of social media, Mark Zuckerberg, founder of Facebook, stated that privacy was no longer a "social norm", as social media sharing reflected a change in attitude. See: Johnson, B. (2010, 1 11). *Privacy no Longer a Social Norm, says Facebook Founder.* Retrieved 7 1, 2016, from <https://www.theguardian.com/technology/2010/jan/11/facebook-privacy>.

10 Several polls have shown the rejection of the public in relation to surveillance and data gathering. These polls will be discussed in Section B.II. See: Jourová, V. (2015). *Data Protection Eurobarometer-Factsheet*. European Commission. See also Madden, M., & Rainie, L. (2015, 5 20). *Americans' Attitudes About Privacy, Security and Surveillance*. Retrieved 7 2, 2016, from <http://www.pewinternet.org/2015/05/20/americans-attitudes-about-privacy-security-and-surveillance/>.

11 See footnotes 54 to 56.

data subject, that recognizes the need for awareness of the users in relation to the costs and rewards of the data exchange.

9 The aim of proposing these components is not to anticipate their actual adoption - although their compatibility with the GDPR will be briefly examined - but to explore alternative ways of providing and managing information that, while not a novelty, may have useful implications in the assessment of the behavior of the data subject and in the analysis of future measures that seek to ensure a conscientious decision-making scenario in the data protection field.

10 It will be argued that this model may create awareness and responsibility in overcoming bias and impasses studied in the field of behavioral economics but, at the same time, recognizes the paramount economic and social importance of data processing in the current state of development of the technology industry.

B. Privacy in Western traditions: A story about finding and losing control

I. Privacy as control

11 An exposition of the right to privacy should start recognizing that, as Professor James Q. Whitman states, "the concept of privacy is embarrassingly difficult to define."[12] One of the probable causes for this statement is that the notion of privacy raises different connotations depending on social and legal traditions, mainly the Western traditions of Europe and North America.

12 According to Professor Whitman, the concept of privacy in the European tradition is seen as a right strongly attached to human dignity, which implies the control of information that can be disclosed about an individual.[13] In this context, the enemy of privacy is broadly understood as any person, natural or legal, that in some way acquires information and aims to disclose it. More importantly, the European concept of privacy deeply embraces the ability to control the information.

13 On the other hand, the American conception of privacy entails freedom from the intrusion of states and contains a deeper distrust of public agents.[14] Moreover, the American recognition of privacy, due to European influence, also adopted the control of the information as an important value. Arguably, the main and most influential basis for the modern conception of privacy in the United States originates from Samuel Warren and Louis Brandeis' article *The Right to Privacy.*[15] In this article, Warren and Brandeis embraced the right "to be let alone," as an extension of the inviolability of personality. But with the recognition of the right "to be let alone", Warren and Brandeis consequently embraced the need of control of the subject that creates the information:

> *"The common law secures to each individual the right of determining, ordinarily, to what extent his thoughts, sentiments, and emotions shall be communicated to others (...) the protection afforded to thoughts, sentiments, and emotions, expressed through the medium of writing or of the arts, so far as it consists in preventing publication, is merely an instance of the enforcement of the more general right of the individual to be let alone."*[16]

14 Professor Whitman notices the big influence that the European tradition and the concept of human dignity had in Warren and Brandeis, by proclaiming the dangers of losing the capacity of control over the personal information.[17]

15 In the course of the twentieth century, the dire consequences of the categorization and profiling performed by the Nazis certainly influenced the social perception of data processing in the years to come. It is considered that the strong protection of privacy in Germany, with the creation of measures such as the right to informational self-determination, has been a reaction to the Nazi and Communist eras.[18]

16 A parallel control-oriented development occurred in the United States in the second half of the twentieth century. As an example, in 1969, the famous Nader Report elaborated to examine the functioning of the Federal Trade Commission in the United States, raised several privacy concerns in relation to data mining. This report already foresees that the increase of mass data processing and the use of social-psychological analysis of potential markets affected the privacy and autonomy of the

12 Whitman, J. Q. (2004). The Two Western Cultures of Privacy: Dignity versus Liberty. *The Yale Law Journal*, 113, p. 1153.

13 *Ibid,* p. 1161.

14 *Ibid,* p. 1018.

15 Krause, H. D., & Marcus, P. (1977-1978). Privacy. *The American Journal of Comparative Law*, XXVI, 377.

16 *Op.cit.* Warren, S.; & Brandeis; L.

17 *Op.cit.* Whitman, J. Q. p. 1167.

18 Cole, D., & Fabbrini, F. (2016). *Reciprocal privacy: Towards a transatlantic agreement.* In V. C. Federico Fabbrini (Ed.), Constitutionalism Across Borders in the Struggle Against Terrorism (pp. 169-189). Cheltenham UK: Edward Elgar, p. 454.

consumers.[19] While recognizing the importance of the user's autonomy over the information, the report warns of the potential of mass processing of data for marketing practices as a form of social control, due to the possibility of creating normative patterns in the users.[20]

II. Privacy as control: Outdated or ignored?

17 The increasing technological developments and the generalized and ubiquitous flow of personal data has led some to identify a change in the social perception of privacy, in support of a more negligent view that benefits an interconnected world. In support of a new and broader concept of privacy, Facebook's CEO and founder, Mark Zuckerberg, stated in 2010 that "(p)eople have really gotten comfortable not only sharing more information and different kinds, but more openly and with more people. That social norm is just something that has evolved over time."[21]

18 Now, it is undeniable that the concept and perception of privacy have dramatically evolved during the last decades. Nevertheless, the fact that in the past there was a broader understanding of the information that was considered important for the users, does not necessarily mean that people have dismissed the possibility and need for control and the importance of privacy and anonymity.

19 A survey carried out by the European Commission in June 2015 on 28.000 EU citizens, showed that 67% percent of the respondents were concerned about not having control over the information they provide on the internet. The survey showed that although 71% percent of the respondents accept that providing information is part of modern life, the majority of the people still feel uncomfortable about the fact that companies use this information to tailor advertisement. It is interesting to notice that, in comparison to the same survey done in 2010, there is not a substantial change in perception.[22]

20 A similar survey carried out in the United States in early 2015 showed an even higher distrust in the activity of online service providers. This survey showed that for more than 93% of adults, it was important to have control over who can get their information, and 90% considered important to have control over the type of information that can be collected. The survey also evidenced that the majority of respondents have little trust that online service providers keep the collected information private and secure, and 55% believe that people should have the ability to use the internet in a completely anonymous way.[23]

21 This information shows the contradiction between the perception and concerns of the public, with the real life application of data processing. A big part of the problem is based on the fact that, as accurately stated by Professor Lilian Edwards, "(...) users care deeply about their privacy but can't be bothered to read privacy policies."[24]

C. Current state of affairs: An unbalance between regulation and social perception

22 The proposal to modify the DPD introduced on January 25 2012 had as one of its main aims, the idea to strengthen the online privacy of the users. As stated by the EU Justice Commissioner Viviane Reding,"(m)y proposals will help build trust in online services because people will be better informed about their rights and in more control of their information."[25] It is interesting to note that the concept of control has been embraced by the European Union when drafting the original proposal for the GDPR. Nevertheless,

19 Hasty, A. (2014-2015). Treating Consumer Data like Oil. *Federal Communications Law Journal*, 67(2), pp. 307, 308.

20 Silbey, S. S. (1984). Who Speaks for the Consumer? Nader's No Access to Law and Best's When Consumers Complain. *American Bar Foundation Research Journal,* 2, p. 177.

21 Matyszczyk, C. (2010, January 10). *Zuckerberg: I know that people don't want privacy*. Retrieved 7 17, 2016, from <http://www.cnet.com/news/zuckerberg-i-know-that-people-dont-want-privacy/>.

22 European Commission. (2015). *Data Protection Eurobarometer-Factsheet.* For empirical investigations about the value of data for consumers, see: Aquisti, A. (2014), The Economics and Behavioral Economics of Privacy. In Lane, J., Stodden, V., Bender, S., Helen Nissenbaum, H., (Eds), *Privacy, Big Data, and the Public Good: Frameworks for Engagement.* Cambridge University Press. p 8 Retrieved 4 7, 2018, from: <http://wpressutexas.net/cs378h/images/b/b3/LaneEtAlPrivacyBigDataAndThePublicGood.pdf#page=55>.

23 *Op.cit.* Madden, M., & Rainie, L.

24 Edwards, L. (2013). *Privacy, Law, Code and Social Networking Sites*. In I. Brown (Ed.), Research Handbook On Governance Of The Internet (pp. 1-35). London: University of Oxford This phenomenon has been called by some authors as the "privacy paradox", in which internet users have concerns about privacy and know about the privacy terms, but they will not read these terms and will still disclose the information. For more information about the privacy paradox, see: Zuiderveen, F.J. (2014). Improving Privacy Protection in the Area of Behavioural Targeting. *University of Amsterdam Digital Academic Repository,* pp 293-296. See also: Monteleone, S. (2015). Addressing the 'Failure' of Informed Consent in Online Data Protection: Learning the Lessons from Behaviour-Aware Regulation. *Syracuse Journal of International Law and Commerce*, 43(1). p. 75.

25 European Commission. (2015). *Data Protection Eurobarometer-Factsheet*. Commission proposes a comprehensive reform of data protection rules to increase users' control of their data and to cut costs for businesses. Brussels: European Commission.

this possibility in practice does not seem to provide enough protection or control.

23 There are several lawful bases for the processing of data according to Article 6 of the GDPR, including legal obligations and the protection of the data subject's interests. Nevertheless, the consent is the main tool to legitimate data processing,[26] and the primary tool for the data subject to exercise any control.

I. Consent in the EU regulations

24 Article 4 (11) of the GDPR, defines that consent "(...) means any freely given, specific, informed and unambiguous indication of the data subject's wishes by which he or she, by a statement or by a clear affirmative action, signifies agreement to the processing of personal data relating to him or her".

25 The concept of consent adopted in the GDPR relies on perfectly valid grounds, and contains legitimate aims, in the sense that it should be freely given and informed. Nevertheless, so far, the model of implementation of consent that provides control to the data subject and the tools for this end have been inconvenient and not appropriate for the purpose.

26 It is worth mentioning an interesting experiment that was performed by the Norwegian Consumer Council, where volunteers read the terms of use and privacy policies of the apps of an average Norwegian smartphone. The process of reading the terms of use and privacy policies of 33 apps containing around 250.000 words, lasted more than 24 hours,[27] and led the Norwegian Consumer Council to the obvious conclusion that "mobile apps' terms of use and privacy policies fail to uphold privacy obligations and users' consumer rights."[28]

27 A major part of the problem lays in the outdated nature of the current model of consent. As stated by the privacy advocate Simon Davies, "most consent mechanisms were conceived in the pre-dawn of the Internet age. They were developed at a gentler time in history – a time when it was possible to build a simple flow chart of personal data relationships."[29]

28 However, even if it is accepted that consent, as drafted in the GDPR, is a proper tool for control, other provisions further diminish the autonomy of the data subject. Indeed, Article 6 (4) of the GDPR allows the processing of data for purposes that have not been subject to the consent of the data subject, as long as the controller proves compatibility between the initial and the new purposes. The criteria to determine such compatibility (Article 6 (4) (a-e)) are conspicuously broad, with plenty of space for interpretation.

29 With reason, critics of a consent-based approach point out its lack of suitability as a practical solution.[30] Some of these critics aim to prove that the concept of consent is currently an illusion, as the users give it on a non-negotiable, non-informed, and pressurized basis.[31] In a broader way, some authors believe that the sole concept of control is an illusion, since data subjects constantly and willingly disclose their information.[32]

30 Professor Anita L. Allen for example, stresses the practical difficulties of providing control[33] on the grounds that "control over personal data appears to be neither necessary nor sufficient for states of privacy to obtain",[34] since people that may have control over their information, choose to give up this faculty.

31 The position of the author, shared by other authors in the field,[35] is that the consent is a valuable and important tool for the data subject that should not be easily disposed on the grounds of attaining to reality. On the contrary, the concept of consent should maintain its importance in the data protection field, but its direction and implementation should be reconsidered.

26 Enerstvedt, O. (2015). *Consent as a Basis for the Processing of Personal Data under the European Data Protection Directive: case study on Facebook* (Thesis). Oslo: University of Oslo. p. 1.

27 For more information about the experiment, see: The Norwegian Consumer Council. (2016). *250,000 words of app terms and conditions*. Retrieved 7 17, 2016, from <http://www.nbcnews.com/technology/ftc-says-flashlight-app-left-consumers-dark-2D11702823>.

28 The Norwegian Consumer Council. (2016). *Appfail: Threats to Consumers in Mobile Apps*. Oslo: The Norwegian Consumer Council, p. 4.

29 Davies, S. (n.d.). *Why the Idea of Consent for Data Processing is Becoming Meaningless ad Dangerous.* Retrieved 7 17, 2016, from <http://www.privacysurgeon.org/blog/incision/why-the-idea-of-consent-for-data-processing-is-becoming-meaningless-and-dangerous/>.

30 *Op.cit.* Koops, B.J, p. 3.

31 *Op. cit.* Edwards, L. p. 24.

32 *Op.cit.* Allen, A.L. p. 869.

33 *Ibid.*

34 *Ibid.* p. 867.

35 Staben, J. (2012). *"Consent under pressure and the Right to Informational Self-Determination."* Internet Policy Review, 1(4). See also: *Op.cit.* Zuiderveen, F.J. (2014). pp. 201, 236 and 237.

II. Data protection regime: The unbalance between regulation and social perception

32 The previous sections have made evident different problematic issues. One of these issues is that the concept of control of the data is still an important basis for the right to privacy in Western traditions, both from the academic and the social point of view. On the other hand, the previous sections state that consent, as a mechanism of control included in the GDPR and other legislation, has not contributed to create a better suited and rightly entitled data subject.

33 The disparity shown on the previous sections between the ideal of control and the real practice of data mining and processing is largely a result of outside pressure and economic interests.

34 It is interesting to notice how the efforts of the OECD in the elaboration of the *Guidelines on the Protection of Privacy and Transborder Data Flow*, were primarily driven by economic interests. Indeed, the effort of creating the Guidelines mainly answered to the need of establishing a set of principles that would guard against economic protectionism.[36] The influence of the OECD's instrument has been extensive and can be found in the Safe Harbor Agreement of 2000 between the European Commission and the US Department of Commerce, invalidated by the Court of Justice of the European Union,[37] and in the data privacy laws of several countries outside Europe.[38]

35 This is also true in the context of the European Union with the creation of the former DPD. Indeed, as stated by Professor Lee A. Bygrave, the European Commission, although partly motivated by the protection of human rights, was mainly aiming to eliminate barriers to the realization of the internal market.[39] The purpose of the DPD is ambivalent, as expressed in Article 1, which, at the same time seeks to protect fundamental rights and freedoms of natural persons, whilst prohibiting any restriction in the free flow of personal data.[40]

36 The value of information as a fundamental economic asset is a fact that companies have assimilated for several decades. Therefore, the influence of the private actors in the adoption of the proposal of the GDPR is not surprising. Indeed, the GDPR was one of the most lobbied legislations in Europe,[41] with 3999 amendments only by the Civil Liberties, Justice and Home Affairs Committee.[42]

37 The current American approach to the processing of data has also been subject to different pressures that diminish the control of the data subjects. Besides more direct pressure imposed by the public[43] and private[44] sectors, the American legislation, in scenarios not only limited to data protection, has been greatly influenced by economic elites and organized groups representing economic interests, while the average citizen has little to no influence in the elaboration of public policies. This phenomenon has been called an Economic-Elite Domination.[45]

D. A Behavioral Economics Perspective

38 As it has been analyzed in the previous sections, the idea of control of the data subject is not new, but it has been attached to the right to privacy since the moment that Western doctrines identified the emerging threats in an increasingly more technological world.

39 This article supports the revitalization of the concept of informed consent as an appropriate tool of control of the data subjects. Nevertheless, the analysis of

36 *Op.cit.* Bygrave, L., p. 43.

37 Case C-362/14 Maximillian Schrems v Data Protection Commissioner (2015).

38 *Op.cit.* Bygrave, L. p. 50.

39 *Ibid.* p. 55.

40 *Ibid.* p. 57.

41 European Digital Rights (EDRi). (2016, February 24). *Data Protection Lobbyotomy Part 1: Influencing the Dutch government.* Retrieved 7 17, 2016, <from https://edri.org/data-protection-lobbyotomy-part-1-influencing-the-dutch-government/>.

42 Albrecht, J. P. (2015, January 7). *EU General Data Protection Regulation: State of play and 10 main issues.* Retrieved 7 17, 2016, from <http://www.janalbrecht.eu/fileadmin/material/Dokumente/Data_protection_state_of_play_10_points_010715.pdf>.

43 Due to the attacks on 9/11, the NSA, with the help of legislative measures, created extensive networks of collaboration with telecommunication companies. The Protect America Act of 2007, reinvigorated by the FISA Amendments Act of 2008, gave immunization to private companies that voluntarily cooperated with the US intelligence, culminating in the PRISM program, which managed to create partnerships with Microsoft, Google, Yahoo, Facebook, among several others. See: Laberge, C., 2010. *To What Extent Should National Security Interests Override Privacy in a Post 9/11 World?* Victoria University of Wellington Working Paper, pp. 1-134. See also: Tucci, L., 2013. *Putting a Price on Information: The nascent field of infonomics.* [Online] Available at: <http://searchcio.techtarget.com/opinion/Putting-a-price-on-information-The-nascent-field-of-infonomics>.

44 For more information in relation to the data processing practices of private companies in the post 9/11 legal regime, see the documentary: *Terms and Conditions May Apply*. 2013. [Film] Directed by Cullen Hoback. USA: Hyrax Films.

45 Gilens, M., & Page, B. I. (2014). Testing Theories of American Politics: Elites, Interests Groups, and Average Citizens. *Perspectives on Politics*, 12(3), p. 565.

the informed consent of the previous sections shows that models that may work in theory, very often prove to be unsuccessful in practice. The previous statement does nothing different than recognizing the complexity of the human mind and the effect of such complexity in the individual behavior and social environments.

40 The field of behavioral economics relies on the idea of economy and society as complex phenomena.[46] In this sense, behavioral economics seeks to understand the behavior of individuals and its consequences, grounded on the experimental knowledge of the good or bad choices of people. In other words, it means to reorient the interest of economy from formal theoretical assumptions to psychology and real human actions.[47]

41 A behavioral economics-oriented legal approach explores the actual human behavior in connection to law[48] over purely hypothetical or ideal scenarios. In comparison with a regular legal analysis, the inclusion of the economic factor, as stated by Posner, "(...) tries to explain and predict behavior of participants in and persons regulated by the Law".[49] But also, while the standard model of economics is based on strong assumptions,[50] such as ideal decision-making scenarios, behavioral economics tests these models in real life situations, to find evidence of the actual behavior of people.

42 Indeed, one of the most important differences of the behavioral economics approach in contrast to traditional approaches is that it aims to increase the explanatory power of economy by relying on psychological foundations.[51] This means that while it is possible to rely on certain assumptions, the ultimate test of the theory must prove accurate or congruent with reality.[52] The main reason for this is that the sole idea of implementing a behavioral approach, especially in the legal field, comes from the challenges and contradictions that the experiments have shown in relation to economic assumptions that have been paradigmatic.[53] More importantly,from these experiments, new assumptions have arisen, some of which will be explained in the next sections due to their relevance in the field of data privacy.

43 This article argues that in the current state of affairs, the problematic issues presented in the data protection field, due to the lack of consideration of the data subjects' point of view are arguably creating a favorable scenario for the application of a behavioral economics-oriented analysis that takes into account both psychological and sociological factors. This position has been examined by other authors in the field who have highlighted the importance of the economics of privacy and the behavioral economics from a privacy perspective,[54] grounded on the problematic issues for the data subject due to an asymmetric access to the information.[55] Also, other authors have approach the notion of consent from a behavioral economics perspective,[56] albeit arriving to different proposals to overcome biases and impasses.[57]

I. Bounded rationality

44 The concept of bounded rationality recognizes that people have constraints in their rational capacities, which implies that very often, people use "rules of thumb" to make decisions[58] that rely more on automatic impulses than on conscious thinking.

45 A good explanation of this phenomenon is provided by the Nobel-prize winner, Daniel Kahneman. He distinguishes between two systems of the mind: in System 1, the mind operates automatically with no sense of voluntary control; while in system 2, there is effortful and demanding mental activity. What is interesting is that the effortless impressions of System 1 tend to be the source for the conscious

46 Frantz, R. (2013). Friedrich Hayek's Behavioral Economics in Historical Context. In R. Frantz, & R. Leeson (Eds.), *Hayek and Behavioral Economics* (p. 1.34). Hampshire: Palgrave Macmillan. P. 3.

47 Camerer, C. F., & Loewestein, G. (2004). Behavioral Economics: Past, Present, Future. In C. F. Camerer, G. Loewenstein, & M. Rabinn (Eds.), *Advance in Behavioral Economics* (pp. 1-51). Princeton NJ: Princeton University Press. p. 39.

48 Jolls, C., Sunstein, C. R., & Thaler, R. (1998). Behavioral Approach to Law & Economics. *Stanford Law Review*, 50, p. 1476.

49 Posner, R. A. (n.d.). Values and Consequences: An Introduction to Economic Analysis of Law. University of *Chicago Law School, Program in Law and Economics,* Working Paper 53, p. 2.

50 Cartwright, E. (2011). *Behavioral Economics* (3 ed.). London: Routledge. p. 4.

51 *Op.cit.* Camerer, C. F., & Loewestein, G. p. 3.

52 *Ibid.,* p. 4.

53 Aaken, A. v. (2014). Behavioral International Law and Economics. *Harvard International Law Journal*, 55(2), p. 422.

54 *Op.cit.* Acquisti, A. (2014).

55 Acquisti, A., Grossklags, J. (2007). What Can Behavioral Economics Teach Us About Privacy? In Acquisti, A.; Grossklags, J. (Eds), *Digital Privacy: Theory, Technologies and Practices*. Taylor and Francis Group.

56 *Op.cit.* Zuiderveen, F.J. (2014). pp. 286-298. See also: Zuiderveen, F.J. (2013). Consent to behavioural targeting in European Law: What are the policy implications of insights from behavioural economics. *University of Amsterdam Law School Research Paper No. 2013-43.* Also: *Op.cit.* Monteleone, S. (2015). Addressing the 'Failure' of Informed Consent in Online Data Protection: Learning the Lessons from Behaviour-Aware Regulation. *Syracuse Journal of International Law and Commerce*, 43(1).

57 *Op.cit.* Zuiderveen, F.J. (2014). pp. 299-342.

58 *Op.cit.* Cartwright, E. p. 10.

choices of System 2. Moreover, the choices of System 1 may also arise with a prolonged practice that creates an automatic conduct.[59]

> *"(w)hen all goes smoothly, which is most of the time, System 2 adopts the suggestions of System 1 with little or no modification. You generally believe your impressions and act on your desires, and that is fine—usually (...) When System 1 runs into difficulty, it calls on System 2 to support more detailed and specific processing that may solve the problem of the moment. System 2 is mobilized when a question arises for which System 1 does not offer an answer".*[60]

46 This categorization of decision-making systems has an important influence on the way in which consent is provided online. A vital issue with consent is that, in Kahneman's words, "we can be blind to the obvious, and we are also blind to our blindness".[61]

47 Arguably, ticking boxes of acceptance for the provision of online services on the current state of affairs seems more reflective of System 1 than System 2, and when the data subjects provide consent for countless data processing activities, they do not read such terms or understand their implications.

48 Nevertheless, it should be recognized that this situation of automatic response of online users cannot be exclusively blamed on the data subjects. The current model of data processing and the economic interests behind them provide the proper condition for this problem.

49 The generalized use of cookies gives a good example of a model that, due to its omnipresence, leads to automatic decisions (System 1) of the data subject. According to Article 5(3) of the Directive 2002/58/EC, the user must give his consent to the use of cookies. Although for some authors, the possibility of consenting the use of cookies means a positive change that represents an almost informed opt-in mechanism;[62] the problem arises with the fact that 50.1% of all websites on the internet are currently using some type of cookies, while a big part of these sites are the ones that contribute most of the traffic of the Internet, such as Youtube, Amazon, and Wikipedia, among others.[63] The fact that the majority of websites and the most important and frequently visited sites on the internet require the users to constantly provide their consent, makes the act of taking a responsible and informed decision impractical and costly. Thus, accepting the use of cookies becomes an automatic action of the System 1.

50 It is also worth mentioning that while accepting the use of cookies requires a costly and imperfect consent described above, the legislation allows the data controller to do without consent when the cookies are considered strictly necessary.[64] The fact that the controller may use cookies even without the users' consent, arguably creates a perception of futility in the action of accepting the privacy or cookies policies, further affecting the amount of effort that the data subjects will invest in accepting such terms.

51 The application of a behavioral-oriented perspective in this matter may provide valuable contributions for a different approach. According to the bounded reality concept, "one of the tasks of System 2 is to overcome the impulses of System 1. In other words, System 2 is in charge of self-control".[65] What this means is that, in a situation where there is an automatic impulse of System 1, such as providing consent for the use of cookies, System 2 can have the power to overcome said impulse, and by overcoming an automatic decision of providing consent, the user may take a better-suited decision.

52 Therefore, a mechanism that seeks to ensure the right to privacy of the users should, in its foundation, create tools that encourage conscious and mindful decision-making. The purpose of a measure in this sense is not to create unnecessary burdens for the users or to make online browsing tedious, but to properly inform the users of the nature of the data that it is being processed and the important implications that the activity of data processing may have for them. As will be exposed later in this article, the measures to overcome a bounded reality phenomenon may consist in a better provision of information of the processing and its practical implications for the data subject, as well as in the implementation of a value-oriented model that may encourage the users to be more involved in the data processing activity.

II. Loss aversion

53 Another interesting phenomenon that may be initiated relates to "loss aversion". This concept understands that people, when facing losses in a

59 Kahneman, D. (2011). *Thinking Fast and Slow.* New York: Farrar, Straus and Giroux. pp. 22-23.

60 *Ibid.,* p. 26.

61 *Ibid.*

62 Bond, R. (2012). The EU E-Privacy Directive and Consent to Cookies. *The Business Lawyer*, 68, p. 215.

63 W3Techs. (2016, 7 18). *Usage of Cookies for websites*. Retrieved 7 18, 2016, from <https://w3techs.com/technologies/details/ce-cookies/all/all>.

64 See: Article 5(3), Directive 2002/58/EC. See also: Information Commissioner's Office. (2012). *Guidance on the Rules on use of Cookies and Similar Technologies.*

65 *Op.cit.* Kahneman, D. p. 28.

certain transaction, tend to give more weight or importance to said loss than to the gains that the transactions may bring.[66] This concept is better understood when analyzed together with the "status quo bias", which states that unless there is a good reason to change, people tend to stick to what they already have, even if the alternative seems more promising.[67] People therefore have the tendency to stay on the safe side, by giving a higher value to what may be lost than to the reward or retribution of the transaction.

54 In a general way, it is safe to say that one of the main threats to the right to privacy in the activity of data processing is the ignorance of the losses that the unlimited processing of data entails for the data subject.

55 A control-based model must, therefore, tackle this issue in different ways. As with the phenomenon of bounded reality, the data subject should be informed of the consequences of providing consent for the activity of data processing. Since the loss aversion phenomenon relies on the fact that people give higher value to their "belongings" in a transaction, in order to use this tool to shape behavior, the data subject must be aware both of the losses and the gains of a data transactions.

56 However, by itself, the sole recognition of the losses and gains may not be enough when there is not a real consequence with regard to the person's interests, economic or the like. As explained in relation to the bounded reality, the users must be able to take a mindful decision on the provision of data, not inclined to automatic impulses. This situation leads to the proposal of a mechanism that relies on the attention of the data subjects by directly affecting their interests and also by benefitting them. This can be more easily tackled in a value-oriented model that will be proposed later in this paper.

III. Time inconsistency

57 The phenomenon of time inconsistency shows that people tend to grab immediate rewards at the same time that avoid immediate costs. For example, procrastination comes from the avoidance of immediate costs in performing a task, even when performing this task may have future rewards. Overeating comes from embracing immediate rewards over foreseeable problems, such as becoming overweight. In summary, this phenomenon shows that people tend to prefer present or immediate rewards over future costs, and prefer to avoid present or immediate costs, even if they carry future rewards[68].

58 The activity of acquiring a service on the internet through consent to provide personal data constitutes an immediate reward. The service, provided immediately, outweighs the negative consequences for the users of providing such data; consequences that in most cases are not clear or evident. Moreover, even if the user has the will to provide a responsible decision, the action of reading terms and conditions would be too costly in comparison to the reward.[69]

59 In order to expect responsible behavior from the users in the disposition of data, the information of the data processing must be provided in a less costly way that allows the user to easily identify the different aspects of the activity. More importantly, a less costly solution for the user must also consider a more unified way of data management, which will be exposed as a proposal in this contribution.

IV. Bargaining impasse and self-serving bias

60 Another concept that may find an interesting application is the bargaining impasse and self-serving bias. According to this behavior, there is a tendency of people to identify or to consider something as fair when the outcome represents a benefit for them.[70] Moreover, this tendency shows that people tend to believe that their notion of what it is fair is impartial, so when the other party bargains, this action is considered aggressive and unfair.[71]

61 In general terms, users are kept uninformed or misled of the outcome of a transaction that implies the processing of data.[72] In this sense, the tendency to identify fairness or unfairness in a self-serving

66 For more information about the phenomena of "loss aversion", see: Tversky, A., & Kahneman, D. (1991). Loss Aversion in Riskless Choice: A Reference-Dependent Model. Quaterly Journal of Economics, 106(4), p 1038.

67 Thaler, R. H. (2015). *The Making of Behavioral Economics: Misbehaving.* New York: W.W. Norton & Company, Inc. pp. 131 and 154.

68 O'Donoghue, T., & Rabin, M. (2004). Doing it Now or Later. In C. F. Camerer, G. Loewenstein, & M. Rabin (Eds.), *Advance in Behavioral Economics* (pp. 223-251). Princeton NJ: Princeton University Press. p 224.

69 *Op.cit.* Zuiderveen, F.J. (2014). p. 299.

70 Babcock, L., & Loewenstein, G. (2004). Explaining Bargaining Impasse: The Role of Self-Serving Biases. In C. F. Camerer, G. Loewenstein, & M. Rabin (Eds.), *Advance in Behavioral Economics* (pp. 326-343). Princeton NJ: Princeton University Press. p 236.

71 *Op.cit.,* O'Donoghue, T., & Rabin, M. pp 326-327.

72 Kerber W. (2016). Digital Markets, Data, and Privacy: Competition Law, Consumer Law, and Data Protection. *Join Discussion Paper Series in Economics*, p. 11.

bias requires knowledge of the value, as well as of the type, degree of sensibility, and final use of the data; however, this is all information that is not always easily accessible. In other words, for the users to determine the fairness of a bargain, they should be aware of both the reward (the provision of a service) and the costs and implications of this reward.

62 This bias or impasse relies on the stringent viewpoint of people when facing a bargain. It also implies that under a better-suited model, this jealous conduct that people have over their belongings may be used to create a more responsible and conscious data subject. In any case, the possibility of control that this phenomenon brings, not only demands a better provision of information, but also requires a real possibility of bargaining from the data subject, which further supports a value-oriented model and a unified system of data management.

V. Confirmatory bias

63 The confirmatory bias implies that individuals tend to positively rate new information that is consistent with their initial opinion, and negatively rate the information that contradicts said initial opinion.[73] The confirmatory bias denotes the misinterpretation of ambiguous information, as evidence that confirms an initial opinion.[74]

64 More importantly, it has been determined that an agent with a confirmatory bias habitually believes in hypotheses that are wrong, which at the same time, represents an opportunity for an observer to take advantage.[75] Very often, private and public agents use the confirmatory bias to shape or strengthen wrong ideas.

65 Thus, although people show concerns for their privacy and crave better control over their information, the extent to which people know how their privacy is being protected tends to be more limited, and it is often subject to intentionally provoked biases. Indeed, even while there is a general distrust of the public in the activities of data processing, the perception of people in this regard is frequently inaccurate.[76]

66 There are no few examples of corporate power and media coverage diminishing privacy scandals, or supporting wrong ideas by implying that technology companies are strongly protecting the privacy of their users.

67 An example of the influence of technology companies' perpetuation of inaccurate ideas about data processing is the dispute between Apple and the FBI. Apple denied the requirement made by the FBI to create a backdoor and, therefore, unlock an iPhone belonging to an alleged terrorist, arguing the defense of civil liberties and the protection of people's privacy.[77] While this refusal of Apple has been praised as a strong protection of users' privacy,[78] it should not be forgotten that it has also served as an effective advertisement for the iPhone's security and encryption.[79] Moreover, Apple's strong stand for security and privacy has signified great economic benefits for the company by providing successful access to markets like China, where people are becoming increasingly concerned about state surveillance.[80]

68 Other technology companies have crafted their media image in similar ways. Facebook's Kathy H. Chan stated: "our philosophy is that people own their information and control who they share it with".[81] In the same way, Google's Eric Schmidt stated: "(m)y interpretation is that there is concern that we might be misusing this data and we're not telling you [about it], which I assure you is not the case. We're very committed to telling you what we do".[82]

69 In this context, is it is ironic that according to the NSA, these three companies were aware and gave access to people's data in the activities of mass surveillance performed within the PRISM program.[83]

73 *Op.cit.,* Cartwright, E. p 177.

74 Rabin, M., & Schrag, J. L. (1999). First Impressions Matter: A Model of Confirmatory Bias. Quarterly Journal of Economics, p 38.

75 *Ibid.*

76 *Op.cit.,* Staben, J.

77 Kharpal, A. (2016, March 29). *Apple vs FBI: All you need to know.* Retrieved 7 18, 2016, from <http://www.cnbc.com/2016/03/29/apple-vs-fbi-all-you-need-to-know.html>.

78 MacGregor, S. (2016, February 18). *Apple isn't protecting a shooter's iPhone data – they're defending digital privacy.* Retrieved 7 18, 2016, from <https://www.theguardian.com/commentisfree/2016/feb/18/san-bernardino-shooter-iphone-apple-tim-cook-fbi-decrypt-unlock>.

79 Grossman, L. (2016, March 29). *Here's Who Really Lost in the Apple-FBI Showdown.* Retrieved 7 18, 2016, from <http://time.com/4275033/apple-fbi-iphone-case/>.

80 Benner, K. (2016, February 20). *Apple Sees Value in Its Stand to Protect Security.* Retrieved 7 18, 2016, from <http://www.nytimes.com/2016/02/21/technology/apple-sees-value-in-privacy-vow.html?_r=2>.

81 Chan, K. H. (2009, February 16). *On Facebook, People Own and Control Their Information.* Retrieved 7 18, 2016, from <https://www.facebook.com/notes/facebook/on-facebook-people-own-and-control-their-information/54434097130>.

82 Grabham, D. (2013, May 24). *Google: "we have a clear incentive to protect your privacy.* Retrieved 7 18, 2016, <from http://www.techradar.com/news/internet/google-we-have-a-clear-incentive-to-protect-your-privacy--1154069>.

83 Kleinman, A. (2016, March 20). *NSA: Tech Companies Knew*

70 Moreover, other factors that create confirmatory bias in the data processing activity have been recognized. For example, Julian Staben argues that consumers are used to having protective warranties and cancellation policies in commercial purchases, which lead them to assume that the same protection applies to privacy policies.[84]

71 A confirmatory bias may be used in a positive way in terms of empowerment of data subjects, in the sense that under appropriate conditions and with enough information, the data subjects may be more critical in their perception of the commercialization of data and, therefore, be more careful in the disposition of such data.

E. Making a responsible data subject: Applying behavioral economics to create informed consent

72 The models of behavioral economics previously mentioned are crafted after experimental results that have evidenced that certain economic models, based on ideal behavior, do not correspond to reality. Instead, the experiments have discovered that the actions of people can be more counterintuitive. In the application of these models to the data privacy field, potential conclusions and proposals arise.

73 The following sections will analyze a proposal of a model of information provision and data management from the users' perspective, composed of three components. The first component will analyze alternative methods of providing user-friendly information online, mainly using the example of Creative Commons. Since the GPDR currently suggests the provision of information in combination with standardized icons and in a concise way, it is expected that this component will form an already important aspect of the regulation.

74 The second and third components will explore alternatives of data management and data disposal that rely more heavily on the intervention and awareness of the data subjects, thus helping to overcome some of the bias and impasses of behavioral economics. Indeed, the second component considers the difficulties of having an informed data subject in a fragmentary scenario and, therefore envisages the need for creating a system that contains unified information of the data circulating online of the users. Finally, the third component, considers the need for a more involved and aware user in relation to the costs and rewards of the data exchange, thus it will analyze the possibility of implementing a model based on the value of the data to the benefit of the data subject.

75 Although these last two components are only hypothetical and are not expected to be adopted literally by any jurisdiction, this article will argue that they are not contradictory with the GDPR and therefore, do not rely on the infeasible scenario of abolishing existing data protection laws.[85] Consequently, even if the following sections envisage a proposal based on more control of the data subject, it is not the intention of this article to stop relying on the protectionist and arguably paternalistic provisions of the GDPR in relation to consent,[86] but to explore alternative ways of providing and managing information that, while not a novelty, may have relevant implications in the assessment of the behavior of the data subject and in the analysis of future measures to ensure a conscientious decision-making scenario in the data protection field.

I. Providing data processing information

76 Arguably, one of the main issues in the current model of data processing is the assumption that actual informed consent can be expected from the data subjects, in the way in which the information of the data processing is being provided.

77 According to the GDPR, there is a substantial amount of information that must be provided to the data subjects. Mainly, Article 13 contains such a requirement, which includes the identity of the controller, purpose of the processing, and the identity of the recipients, among others. Article 14 includes the information that must be provided if the data has not been obtained directly from the data subject.

78 The purpose of providing this information and obtaining consent is to properly inform the users on the basis of the principles of transparency (Article 5(1)(a) and Article 12(1) of the GDPR) and to put people in control of their personal data.[87] This condition is therefore laudable in light of this work, but the way in which this information has so far been provided is not user-friendly and, mainly, does not comply with its purpose.

About Prism the Whole Time. Retrieved 7 18, 2016, from <http://www.huffingtonpost.com/2014/03/20/nsa-prism-tech-companies_n_4999378.html>.

84 *Op.cit.,* Staben, J.

85 *Op.cit.* Zuiderveen, F.J. (2014). p. 299.

86 *Ibid.* pp. 242-247.

87 *Op.cit.* Zuiderveen, F.J. (2014) p. 201.

79 Providing information for the data processing does not have to be this costly, however. In this line of thought, an example of providing legal information to a broad public is the Human Readable layer of the Creative Commons license. This tool was crafted with the understanding that most creators of content may not have a legal background, and therefore, require more suitable information. Indeed, Creative Commons explains the purpose of the Human Readable in the following way:

> *(...) since most creators, educators, and scientists are not in fact lawyers, we also make the licenses available in a format that normal people can read — the Commons Deed (also known as the "***human readable***" version of the license).*[88]

80 Creative Commons managed to summarize difficult copyright concepts in user-friendly images. Concepts of copyright rules, such as the right to communicate, distribute or reproduce a work, the attribution of moral rights, and other legal concepts, are contained in figures that do not require specialized knowledge.

81 In terms of privacy, some efforts have been made to provide better information to the users. In Germany, Wikimarx[89] highlights the most critical or important provisions in the terms of service, although it requires diligent and concerned users.

82 It is certainly valuable to recognize that, in most cases, the receivers of legal information online, especially in the field of data privacy, are not lawyers. In this sense, it is self-evident that relying on difficult and long privacy policies to prove the informed consent of a user is not an accurate way to create control. However, the way in which the information is provided should be reconsidered, without necessarily modifying the set of information required. As the information required in Article 13 of the GPDR aims to create an entitled and informed data subject, this contribution does not challenge the importance of this information, but the costly and ineffective way in which it is delivered by the service providers.

83 It is important to notice that the GPDR already contemplates the provision of information in a user-friendly way. Recital 60 of the GDPR states that "(...) information may be provided in combination with standardized icons in order to give in an easily visible, intelligible and clearly legible manner, a meaningful overview of the intended processing". In addition, Article 13 states that the controllers must take appropriate measures to provide any information in a "(...) concise, transparent, intelligible and easily accessible form, using clear and plain language (...)". Therefore, the adoption of a user-friendly method of provision of information, such as the one used in Creative Commons, is an interesting approach that does not contradict the GDPR but, on the contrary, may help develop it.

84 On the other hand, it is relevant to consider that, as stated by Wolfgang Kerber, data subjects are intentionally kept uninformed about the data processing by service providers.[90] Therefore, it may be assumed that a simplified way of providing information without existing standardized icons or a stringent regulation - while attractive for service providers due to its simplicity - has the potential to diminish the information received by the data subjects and create confusion. In other words, a simplified way of providing information of the data processing may be used by service providers arbitrarily to create ignorance and confusion among data subjects.

85 Thus, while the use of icons in a similar way to Creative Commons may be attractive for service providers due to its simplicity, some legislation would be expected to specify the recommendation of the GDPR, but mainly to establish the guidelines of these icons and ensure that their use is indeed standardized, understandable, and effective to transmit the information required by the GDPR.

II. Unifying the information

86 The practical problems that arise from the huge amount of information, which the users are supposed to read should not be undermined. Even if the privacy terms and conditions of a service are provided in a user-friendly way, the disparities with the terms and conditions of other services, and the difficulties of understanding their differences will arguably not encourage users to take more responsible decisions. In this sense, the possibility of creating a unified system for data management may be a viable proposal to encourage control.

87 The idea of creating a unified system for the management of data is not a novelty. The FTC Commissioner Julie Brill created an interesting initiative called "Reclaim your Name," by which she encouraged data brokers to create a consumer-friendly online service that would give access to the information that data brokers have of them.[91]

88 Creative Commons. (n.d.). *Licensing Considerations: What our licenses do.* Retrieved 7 18, 2016, from <https://creativecommons.org/share-your-work/licensing-considerations//>.

89 Wikimarx. (n.d.). *Wikimarx.* Retrieved 9 8, 2016, from <http://www.wikimarx.de/>.

90 *Op.cit.* Kerber W. p. 11.

91 Brill, J. (2013, October 28). *Data Industry Must Step Up to Protect Consumer Privacy.* Retrieved 7 18, 2016, from <http://adage.com/article/guest-columnists/data-industry-step-protect-consumer-privacy/244971/>.

88 Moreover, Data Management Platforms (DMP) have emerged during the last years, offered by companies like Oracle or Adobe.[92] These platforms store data which, after a process of analysis, provide useful information for businesses, mainly profiles for targeted online ads.[93] The DMPs, although mainly used for companies in the monetization of data, can be used as models of data management for data subjects.

89 In this line of thought, startups like Datacoup have started the path of creating a value-oriented platform and marketplace for the users to sell their data for a fixed amount per month.[94] The company Citizen Me provides a similar service for consumers with the possibility of earning cash or donating data to charity.[95] Although the payment of a small amount of money in exchange for the data of all the social networks of the users is still far from an actual management and marketplace of data, the approximation to control of the data subject is certainly present, as it provides the possibility not only to manage unified sets of information but also to make this information a valuable good.

90 This contribution argues that a unified system for data management does not contradict the GDPR; on the contrary, a unified system may help develop and create an effective right of data portability (Article 20 of the GDPR). This right, that obligates controllers when requested to provide the personal data of the data subject in a structured, commonly used and machine-readable format can be, in practice, effectively exercised with the use of a unified data management system.

91 A system of this type is already attractive for companies like Adobe and Oracle and may be profitable for others. Therefore, the presence and control of the supervisory authorities would be required, especially during the examination of a data protection impact assessment (Article 35 of the GDPR). Certainly, it is expected that the eventual adoption of a unified data management system would require said assessment, where service providers must conduct an evaluation of the risks and impacts of the data processing, based on the use of what may be considered a new technology,[96] and the fact that it may imply large-scale data processing (Article 35 (b)). Also, and due to the high volume of data that a measure of this nature requires, there may be a risk of illegal and systematic profiling or monitoring of data holders, thus a tightly regulated scenario would be desirable.

92 Eventually, a better-controlled scenario of a unified data management model may include other possibilities of control different than the sole possibility of receiving and selling data. For example, in order to build trust in the user, a unified system should create standard privacy policies oriented to data processors and controllers. Eventually, a unified system may provide the user with tracking tools that identify the current controllers and processors of the data, or mechanisms that allow one to choose the frequency and type of intrusiveness of advertisement.

III. A value-oriented model of data management

93 There is a generalized idea that the most common services provided online are free of charge. In reality, these services are profiting from the data gathered from the data subjects.[97] Indeed, corporations are increasingly treating information as a commodity,[98] as there is a commercial exchange of value, where the internet service provider offers a service in exchange for data and attention,[99] and where these providers gain economic benefits based on the detailed knowledge of the data subject's preferences and behavior.[100]

94 The approximation of data as a valuable good is mostly discussed in the enterprise scenario. So far, most analyses on this matter focus on considering the benefits for big companies in the technological market to treat data as a "natural resource".[101] In this

92 More information about the Data Management Platforms of Oracle and Adobe can be found in the following links: <https://www.oracle.com/marketingcloud/products/data-management-platform/index.html>. <http://www.adobe.com/uk/marketing-cloud/data-management-platform.html>.

93 Marshall, J. (2014, January 15). *WTF is a data management platform*? Retrieved 7 18, 2016, from <http://digiday.com/platforms/what-is-a-dmp-data-management-platform/>.

94 For more information about Datacoup, see: <https://datacoup.com/>.

95 For more information about Citizen Me, see: <http://www.citizenme.com/>.

96 Article 35 of the GDPR states that "(w)here a type of processing in particular using new technologies, and taking into account the nature, scope, context and purposes of the processing, is likely to result in a high risk to the rights and freedoms of natural persons, the controller shall, prior to the processing, carry out an assessment of the impact of the envisaged processing operations on the protection of personal data. A single assessment may address a set of similar processing operations that present similar high risks".

97 *Op.cit.,* Kerber, W. p. 9.

98 Victor, J. M. (2013). The EU General Data Protection Regulation: Toward a Property Regime for Protecting Data Privacy. *The Yale Law Journal*, p .517.

99 *Op.cit.,* Hasty, A. pp. 297, 307, 313.

100 *Op.cit.,* Acquisti, A. p. 6.

101 Deutscher, M. (2013, March 11). *IBM's CEO Says Big Data is Like Oil, Enterprises Need Help Extracting the Value*. Retrieved 7

scenario, issues, such as antitrust have been analyzed when the value of the data and the intensive and disproportionate mining of said data is comparable to charging high prices.[102] Moreover, the emergence of big data has allowed companies like Google, Apple, and Amazon to offer bank-like services, all possible due to the enormous databases and a so far unattainable understanding of the consumer's behavior.[103]

95 This enterprise-oriented approach, which conspicuously recognizes the economic importance of data, not only disregards the possibility of individuals to dispose of their data but, on the contrary, aims to provide tools to monetize the data of the users for the exclusive benefit of companies.[104] This contribution argues that this enterprise-oriented approach not only contradicts the social perception of the data subjects but almost completely excludes the users from real economic benefits.

96 Some authors in the legal field dismiss the debate of personal data as a tradable good on a market, especially under the argument of inalienability.[105] Although this approach is certainly valuable for debate and future regulations, the following analysis will not enter this discussion, but will seek to propose measures that better reflect the current economic approach and the general understanding of data as a valuable good.

97 In any case, legal requirements for a value-oriented model should not be inferior to the requirements for data processing in the current data protection regulations, specifically the GDPR. In other words, the recognition of data as a valuable good from a data subject perspective, should not substantially affect the development of the right to privacy or the extent of the informed consent; on the contrary, it should be focused on strengthening them in a way that creates awareness and seeks to overcome the biases and impasses explained by behavioral economics.

18, 2016, from <http://siliconangle.com/blog/2013/03/11/ibms-ceo-says-big-data-is-like-oil-enterprises-need-help-extracting-the-value>.

102 Cooper, J. C. (2013). Privacy and Antitrust: Underpants Gnomes, The First Amendment, and Subjectivity. *George Mason University Law and Economics Research Paper Series*. Vol. 20, 20(4), p. 1131.

103 Packin, N. G., & Lev-Aretz, Y. (2016). Big Data and Social Netbanks: Are you ready to replace your bank? Houston Law Review, 53(5), p. 1216.

104 Twogood, C. (2014, November 19). 5 *Essential Steps Toward Monetizing Your Data*. Retrieved 7 18, 2016, from <http://www.forbes.com/sites/teradata/2014/11/19/5-essential-steps-toward-monetizing-your-data/#3c8786973b85>.

105 *Op.cit.* Zuiderveen, F.J. (2014) p. 252.

1. What do we get by providing value?

98 There are several advantages in the approach of providing value to the data to the benefit of the data subjects. From a competition point of view, the value-oriented approach provides better tools for controlling the activity of data processing.[106] Indeed, Andrew Hasting notices that "a value approach may be more efficient in providing proves of deceptive practices where the agencies would be able to compare the value of the service in comparison with the 'value' of the data provided in exchange, furtherly analyzing an unfair unbalance".[107] In other words, assigning value to the data arguably creates a more objective basis to identify the abusive market behavior of technology companies.

99 In the same way, the value approach is clearly market-oriented, thus it can have beneficial situations for consumers. Companies will be encouraged to provide better and more competitive services; contrary to the current scenario, where users download products and use services without necessarily considering their quality.

100 But more importantly for the behavioral-oriented approach of this work is the awareness that the value-oriented approach may create in the data subjects. It is undeniable that for most users, there is a lack of understanding of the flow of personal data and the ways in which this flow can be controlled[108]. This phenomenon has been mainly grounded on the lack of awareness of legislation and the acquaintance of the private and public actors, but has, so far, not taken into consideration the user's perspective and actions.

101 The behavioral economics situations previously analyzed benefit from a value-oriented approach. The phenomena of bounded reality and time inconsistency, where data subjects accept privacy limitations in an automated way and expect an immediate reward, and the issue of "loss aversion," where users give more weight to the losses and, therefore, stick to their possessions, are all strongly connected. Indeed, tackling these issues as a whole may be done by relying on the awareness of the value of the data and the possibility of disposing of it, which makes the users mindful of the loss that implies a transaction, and leads them to give more weight to the loss of data than to the reward.[109]

102 In the same way, the loss aversion derives from a more protective use of the self-serving bias. The bias of considering something fair when the outcome

106 *Op.cit.*, Kerber, W. p. 16.

107 *Op.cit.*, Hasty, A. p. 318.

108 *Ibid.* p. 302.

109 *Op.cit,* Tversky, A., & Kahneman, D. p. 1038.

represents a benefit for the person,[110] requires, also, an awareness of the value of the data, and the possibility of disposing of it. Furthermore, the bargaining impasse, where users believe in the fairness of their position in a transaction, certainly requires the possibility of having something to be bargained.

2. Value vs. Property

103 Several theoretical approaches have tried to change the perspective of the data protection model to orient it toward a right to property of the data subject.[111] Indeed, Professor Lessig states that "(t)he laws of property are one such regime. If individuals can be given the rights to control their data, or more precisely, if those who would use data had first to secure the right to use it, then a negotiation could occur over whether, and how much, data should be used".[112]

104 It is interesting to note that critics of this model are based on the dangers of allowing consumers to treat their data as commodities, without being properly informed and having information disparities with the processors.[113] The current state of technological developments, with the acquaintance of legislative rules, already did the job of putting the data subjects in this situation, with or without their knowledge. Although implementing a right to property of the data seems to bring control to the data subject, the whole concept of property lays on nebulousness and theoretical difficulties marked by endless conceptual disputes (is property an interest or a dominion of a thing?).[114]

105 The classification of the type of data that may be subject to property may also present different problems. Indeed, several authors, especially in the medical field, have acknowledged the importance of certain types of data to be part of comprehensive databases for public health and safety.[115] The conceptual issues of propertize data would require a thorough classification, which may provide weak protection for certain types of data or too strong protection for other.

106 Also, the recognition of the data subject as the owner of the data may not be completely effective. As stated by Professor Barbara J. Evans, the recognition of property does not necessarily imply legal property protection, as "(l)aw recognizes that there are many situations where consensual transactions cannot be relied on as a way of ordering an owner's relations with the larger community."[116]

107 Moreover, the model of property of data, for some writers, requires the implementation of a highly regulated market,[117] which requires, at the same time, a better suited but currently inexistent online context. Issues on the like of territoriality make a model based on property impractical and hard to implement, due to the fact that, despite certain and significant convergences, legal disparities on national laws in relation to outsourcing, data mining, or information security[118] make a global implementation of policies very difficult.

F. Conclusion

108 This work has shown that the people's perception in relation to data privacy seems to maintain an ideal of control and self-determination. Nevertheless, there is resistance from maintaining a control-oriented approach since the informed consent, the most important tool for this matter, has so far proved to be ineffective in practice.

109 This situation, greatly influenced by the lobby and economic objectives of both public actors and businesses, arguably rests on the lack of consideration of other scenarios and perspectives. The proposal of this contribution is to consider some of the alternative perspectives, which may provide mechanisms that empower the data subjects.

110 The field of behavioral economics, which takes into account psychological considerations, can be a valuable tool for this purpose. More importantly, a change to a behavioral-oriented perspective has as its main objective, to shape a desired behavior on the users, in the sense of creating a truly responsible data subject that can take informed decisions over the data. This work supports the idea that a more user-friendly way to provide information can be a strong mechanism to empower the data subjects, and that initiatives such as a Creative Commons, offer interesting examples.

110 *Op.cit.,* Babcock, L., & Loewenstein, G. p. 326.

111 *Op.cit.* Kerber W. pp. 14-16.

112 Lessig, L. (1998). The Architecture of Privacy (Draft 2). Taipei: Taiwan Net 98's Conference. pp 17.

113 *Op.cit.,* Victor, J. M. p. 518.

114 *Ibid.*

115 Evans, B. J. (2011). Much Ado About Data Ownership. Harvard Journal of Law and Technology, 25(1), p. 88.

116 *Ibid.*

117 *Op.cit.*, Victor, J. M. p. 519.

118 Gunasekara, G. (2006). The "final" privacy frontier? Regulating trans-border data flows. *International Journal of Law and Information Technology*, 15(3), p. 375.

111 On the other hand, a model that allows unified data management and provides a tangible value to the data should encourage the data subject to consciously choose the purpose of such data. For this to happen, the data subject must be fully aware of the type of data, the purpose of the processing, and the retribution received for the processing.

112 The proposal of this work, although specifically focused on creating tools of control, does not aim to create a property right on the data subjects, considering that this measure, in itself, may not be enough. In other words, the approach of this proposal seeks to be practical, relying on the need of control of the data subjects, based on the privacy values that drive the data protection field, and the fact that the data acquired the characteristics of a valuable commodity.

113 In any case, the previous contribution should be understood as a proposal to change the direction in which the activities of data processing have been so far oriented. The new direction that the current online scenario demands must be oriented to the benefit of data subjects and recognizing the actual conduct and behavior of the users of the internet.

Open Science and Public Sector Information

Reconsidering the exemption for educational and research establishments under the Directive on re-use of public sector information

by **Heiko Richter***

Abstract: The article discusses the possibilities of including public research and educational establishments within the scope of the Directive regulating the re-use of public sector information (2003/98/EC – 'PSI Directive'). It subsequently evaluates the legal consequences of such an inclusion. Focusing on scientific information, the analysis connects the long-standing debates about open access and open education to open government data. Their common driving force is the call for a widespread dissemination of publicly funded information. However, the regulatory standard set out by the PSI Directive is characterized by considerable legal uncertainty. Therefore, it is difficult to derive robust assumptions that can form the basis for predicting the effects of extending the PSI Directive's scope to research information. A potential revision of the PSI Directive should reduce this uncertainty. Moreover, PSI regulation must account for the specific incentives linked to the creation and dissemination of research results. This seems of primary importance for public-private research collaborations because there is a potential risk that a full application of the PSI Directive might unduly affect incentives for such collaborations.

Keywords: Public Sector Information; Scientific Information; Open Access; Open Education; Database Protection; PSI Directive; Open Government Data; Re-Use of Information; Copyright; EU Copyright Reform

© 2018 Heiko Richter

Everybody may disseminate this article by electronic means and make it available for download under the terms and conditions of the Digital Peer Publishing Licence (DPPL). A copy of the license text may be obtained at http://nbn-resolving.de/urn:nbn:de:0009-dppl-v3-en8.

Recommended citation: Heiko Richter, Open Science and Public Sector Information – Reconsidering the exemption for educational and research establishments under the Directive on re-use of public sector information, 9 (2018) JIPITEC 51 para 1.

A. Introduction

1 The PSI Directive[1] regulates the **re-use of public sector information (PSI).** Since the Directive entered into force in 2003, it has contained an exemption for research and educational (R&E) establishments in Article[2] 1(2)(e). As a consequence, rules for re-using a large amount of valuable information (such as research datasets, publications or educational material) are either non-existent or fragmented across the EU. In the process of amending the PSI Directive in 2013, the European Commission discussed whether to bring R&E establishments within its scope. However, this was rejected for three reasons: the high burden for clarifying the legal status of research data to make them re-usable under the rules of the PSI Directive would exceed the benefits; the existence of a dynamic and well-established system for disseminating and exploiting research findings and results; and the distinct character of the open access (OA) debate, which is separate from the PSI debate.[3]

* LL.M. (Columbia), Max Planck Institute for Innovation and Competition, Munich.

1 The term 'PSI Directive' refers to the Directive 2003/98/EC on the re-use of public sector information of 17 November 2003 as amended by Directive 2013/37/EU of June 2013. As there are different Recitals to both Directives, it is distinguished between Recitals (2003/98/EC) and Recitals (2013/37/EU) if necessary. Recitals of both Directives are relevant for the interpretation of the Directive in its current form, see Richter, H. (2018), Informationsweiterverwendungsgesetz (IWG), Einl para. 37 et seq.

2 Articles refer to the PSI Directive if not stated otherwise.

3 See SEC(2011) 1152 final, 33 et seq.

2 In the framework of the review of the PSI Directive, which needs to be carried out by July 2018 in accordance with Article 13, the European Commission is now reconsidering the exemption for R&E establishments. In its stakeholder consultation on the PSI Directive, the Commission has not only addressed re-use of scientific information but also explicitly considered access to it ; in fact, in its recently published draft for a recast of the PSI Directive (COM(2018) 234 final), the Commission explicitly proposes to include research data.[4] This unites two worlds, which were separated 20 years ago[5] – the scientific OA-world and the general PSI-world. As discourses have developed strictly in parallel, so far there is no in-depth analysis of R&E information from a PSI point of view.[6] However, times have changed dramatically: the OA-discussions and models have matured, and digitization and "datafication" have also significantly advanced since 2013. Given these technological and socio-economic changes,[7] a closer look at the functioning of the PSI Directive and its possible application to R&E establishments is desirable.

3 The analysis **focuses** on the legal aspects and mechanisms of the PSI Directive. The central question is what the legal consequences of applying the PSI Directive to R&E establishments would be. This can serve as a basis for economic research, which is necessary for predicting regulatory impact. What can already be said in general is that crucial provisions of the PSI Directive are not entirely clear. Therefore, their possible interpretations have to be discussed before applying the PSI Directive to R&E establishments. The analysis is structured as follows: At the outset, OA-policies and open education approaches are contrasted with the origin and general concept of PSI and information of R&E establishment in particular (*sub* B.). Subsequently, the core of the analysis elaborates on the hypothetical question of what would happen if one removes the exemption for R&E establishments (*sub* C.). The discussion focuses on what sort of R&E information would effectively fall within the scope of the PSI Directive under which circumstances and continues with elaborating the legal consequences for such information. In the next step, the analysis observes possible modifications (as opposed to a strict deletion of the exemption) of PSI rules that address R&E establishments (*sub* D.). The current provisions that address public libraries, museums, and archives can give some guidance. The final section draws a conclusion (*sub* E.).

B. 'Open Access', 'Open Education' and 'PSI'

I. Overview

4 In the last two decades, the advancement of digitization and the global connection have brought up seminal debates and changes regarding the dissemination of information and knowledge. With respect to publicly funded information, the developments are driven by the general political thought that if production of information is financed by taxpayers' money, it should be widely distributed for (almost) free and without any restrictions that apply to using such information. This basic rationale finds support in information economics and pervades three particular sorts of information that regulation has so far treated quite separately: first, the broad concept of 'PSI', which largely refers to Open Government Data (OGD); second, the 'Open Access' movement, which addresses scientific information (publications and data) in particular; third, the term 'Open Education', which relates to publicly funded teaching materials and coursework. All of these strands of debate converge when discussing the R&E exemption in the PSI Directive. Therefore, describing particular developments and frameworks sets a starting point for discussion.

II. Open access to scientific information

5 The OA-debate has a long-standing tradition in science.[8] Basically it centers on the political claim of widely disseminating publicly funded scientific information.[9] Historically, it was a reaction of

4 For the stakeholder consultation, see especially question 12b of the European Commission's "Public Consultation on the Review of the Directive on the Re-Use of Public Sector Information (PSI Directive)", running from 12 September 2017 to 12 December 2017. The Commission has published the proposal for the recast of the PSI Directive on 25.4.2018, COM(2018) 234 final, which includes research data in its Article 10 (see also Recitals 23 and 24 of the proposal). However, as the focus is put on the current law, this article does not explicitly comment or discuss the recent proposal. Instead it shall provide the necessary background knowledge.

5 See the Commission's Green Paper on Public Sector Information in the Information Society of 1998 as a starting point for the PSI Directive, COM(1998) 585 final.

6 As opposed to the title, focusing almost exclusively on cultural institutions Jančič, M./Pusser, J./Sappa, C./Torremans, P. (2012), Policy recommendation as to the issue of the proposed inclusion of cultural and research institutions in the scope of PSI Directive – Working Group 5, 6 Masaryk University Journal of Law and Technology 353.

7 For a recent overview see COM(2017) 228 final.

8 Seminal Suber, P. (2012), Open Access; for a comprehensive history of recent OA movements Scheufen, M. (2015), Copyright Versus Open Access – On the Organisation and International Political Economy of Access to Scientific Knowledge, 65 et seq.

9 See Recital 5, Recommendation 2012/417/EU.

academia to the increasing prices of scientific publications and subscriptions controlled by publishers and distributors.[10] The idea is to make research output from **publicly funded research** establishments and projects available for free, with as few restrictions as possible. Central non-binding declarations and frameworks for OA are the "Berlin Declaration on Open Access to Knowledge in the Science and Humanities", the "Bethesda Statement on Open Access" and the "Budapest Open Access Initiative". Known as the "BBB-Declarations" they promote the free availability of works on the public Internet, permitting any use (e.g. read, download, copy, distribute, print, search or link). While there are slight differences, the basic idea of free access is to shift financing from the subscriber to the institution and/or the author.[11] There is a substantial body of literature that discusses and elaborates on the OA-movement and its legal and economic interfaces and implications.[12]

6 In the policy arena, basically two **reference points** can be identified when discussing OA-initiatives and regulating publicly funded research. The *organizational* reference point distinguishes between research establishments on the one hand and public funding organizations (e.g. governmental agencies or research councils) on the other hand. The *informational* reference point relates to the sort of information addressed. A main distinction can be drawn between publications and research data, both being subsets of "scientific information".[13] In general, research can be understood as "a systematic investigation intended to establish facts, acquire new knowledge and reach new conclusions".[14] However, categorizations and definitions are far from being exact or harmonized.[15]

7 Recent **EU policies** advocate that scientific information resulting from public funding should be openly accessible and re-usable as far as possible.[16] There is considerable activity when it comes to EU funding policies. In its Communication of 2012, the Commission has set out OA-policy objectives for research funded by "Horizon 2020".[17] The main idea is to lead by example and to request all funded projects to deposit an electronic version of their publication after an embargo period and to set up a pilot scheme on access and re-use to generated data.[18] Therefore, each beneficiary must ensure OA to all peer-reviewed scientific publications relating to its results of research projects under Horizon 2020.[19] Since January 2017, OA is also the default setting for research data generated in Horizon 2020. However, projects can opt out at any stage.[20]

8 Regarding the **Member States' policies addressing funders and institutions**, the Commission's non-binding "Recommendation on access to and preservation of scientific information" calls to put measures into place that address OA to scientific publications, research data, preservation and infrastructure.[21] Publicly funded research should be widely disseminated through OA-publication of scientific data and papers.[22] Member States have developed different strategies in addressing these issues.[23] Policies of public research institutions and funders largely differ, although there is a clear trend towards openness.[24] Many universities, research institutions and funders have adopted mandates that require their researchers to deposit their findings and provide OA to them. There is, however, considerable uncertainty about the degree to which binding measures, rather than voluntary recommendations, are legitimate.[25]

10 See Papadopoulos, M./Bratsas, C. (2015), Openness/Open Access for Public Sector information and works – the Creative Commons Licensing Model, European Public Sector Information Platform – Topic Report No. 2015/06, 9 et seq.

11 For the different modes of OA, see Suber, P. (2008), available at: <http://legacy.earlham.edu/~peters/fos/newsletter/08-02-08.htm#gratis-libre>.

12 See Scheufen (*supra* n 8).

13 See Recommendation 2012/417/EU.

14 See Information Commissioner's Office U.K. (2017), Information intended for future publication and research information (sections 22 and 22A), No. 45.

15 See Guibault, L./Wiebe, A. (2013), eds., Safe to be open – Study on the protection of research data and recommendations for access and usage, 17.

16 See SMART 2017/0061, 3.

17 See COM(2012) 401 final.

18 See COM(2012) 401 final, 9.

19 See Article 29.2. of the Model Grant Agreement, which sets out detailed legal requirements on OA to scientific publications; see H2020 Programme, Guidelines to the Rules on Open Access to Scientific Publications and Open Access to Research Data in Horizon 2020, Version 3.2 of 21 March 2017, 5.

20 *Ibid.*, 8.

21 See Recommendation 2012/417/EU, which will be replaced by the Commission's Recommendation of 25.4.2018 on access to and preservation of scientific information, C(2018) 2375 final in due course.x

22 See Recommendation 2012/417/EU, Recital 2.

23 See for example the U.K. Research Excellence Framework, as a system for assessing the quality of research in UK higher education institutions, available at: <http://www.ref.ac.uk/>.

24 For a comprehensive overview on the situation in the Member States, see the Report on the implementation of Commission Recommendation C(2012) 4890 final, "Access to and Preservation of Scientific Information in Europe" of 2015; see also <http://roarmap.eprints.org>, a comprehensive searchable database covering OA-mandates of more than 600 public research institutions.

25 See the current debate before the German Constitutional Court on the OA-mandate of the Universität Konstanz: <https://www.lto.de/recht/hintergruende/h/vgh-mannheim-normenkontrollantrag-9-s-2056-16-professoren-universitaet-konstanz-open-access-wissenschaft-urheberrecht/>.

III. Open education

9 When it comes to education, one can also advocate that everything which is ultimately financed by taxpayers' money should be available for everyone at no cost. The broader term **"open education"** covers different aspects of that claim, such as open educational resources (OER) as well as distance learning and massive open online courses (MOOCs) in particular. A major debate centers on OER, which UNESCO defines as "teaching, learning and research materials in any medium, digital or otherwise, that reside in the public domain or have been released under an open license that permits no-cost access, use, adaptation, and redistribution by others with no or limited restrictions."[26]

10 Open education has been discussed in parallel with the OA-debate.[27] However, overarching policies and regulation on the EU-level[28] and coherent strategies of the Member States and their educational establishments seem much **less developed**. As compared to the opening of research information, one can say that this field seems to be in a premature stage. At the same time, the provision of distance learning and MOOCs can be seen as an emerging, quite dynamic field. Many private educational institutions also provide education as a service on the market. The effect of an inclusion of educational establishments within the scope of the PSI Directive is difficult to forecast. Without careful observation and consideration of markets and practices in the light of re-use, one should abstain from overhasty regulatory changes.

11 Therefore, the **prematurity, heterogeneity, and dynamics** in the education sector have a major implication for this study: the focus is deliberately put on research information rather than on educational information. However, the article contrasts educational information from research information, because both types of information follow different legal treatment when it comes to access and copyright regimes. Finally, should the PSI Directive only address research establishments and their information, delineation from educational establishments and their information is necessary.

26 See UNESCO, Paris OER-Declaration of 2012.

27 See Miao, F./Mishra, S./McGreal, R. (2016), eds., Open Educational Resources: Policy, Costs and Transformation.

28 For a comprehensive overview see Inamorato dos Santos, A./Punie, Y./Castaño-Muñoz, J. (2016), Opening up Education: A Support Framework for Higher Education Institutions, JRC Science for Policy Report.

IV. PSI Directive and exemption for research and education

1. PSI Directive

12 On 31 December 2003, the Directive on the re-use of public sector information (2003/98/EC) entered into force. It was revised by Directive 2013/37/EU, which entered into force on 17 July 2013. According to Article 1(1), the **PSI Directive** "establishes a minimum set of rules governing the re-use and the practical means of facilitating re-use of existing documents held by public sector bodies of the Member States". This definition implies a very broad concept of PSI that accommodates a vast variety of information, such as weather, geographical, tourist, economic, legal, and business information.[29] Legal literature has widely dealt with the PSI Directive from various perspectives.[30]

13 The **Directive's goal** is to stimulate further development of the market for services based on PSI and to enhance cross-border use and application of PSI.[31] Also, the PSI Directive addresses divergence as to re-use rules between the Member States and seeks to strengthen competition in the internal market. While initially designed as a regime with a strong competition rationale,[32] the revision extended the PSI Directive to a regulatory instrument that supports OGD efforts of the Member States.[33] This

29 See also Recital 4 (2003/98/EC).

30 In the course of the PSI Directive's revision, considerable legal research had been conducted by the research network LAPSI (Legal Aspects of Public Sector Information) which published many of its results in the Masaryk University Journal of Law and Technology of 2012 and 2015. A comprehensive analysis of the PSI Directive (2003/98/EC) was conducted by Janssen, K. (2010), The Availability of Spatial and Environmental Data in the European Union; elaborate analyses from a competition perspective were undertaken by Drexl, J. (2015), The Competition Dimension of the European Regulation of Public Sector Information and the Concept of an Undertaking, in: Drexl, J./Bagnoli, V. (eds.), State-Initiated Restraints of Competition, 64 and Lundqvist, B. (2013), Turning Government Data into Gold: The Interface between EU Competition Law and the Public Sector Information Directive – With some Comments on the Compass-Case (September 19, 2012), 44 International Review of Intellectual Property and Competition Law (IIC) 79; see also for background of the PSI Directive in connection with the German Informationsweiterverwendungsgesetz (IWG) Richter (*supra* n 1); for a comprehensive examination of the rules for cultural PSBs see Wirtz, H. (2017), Die Kommerzialisierung kultureller Informationen der öffentlichen Hand – Auswirkungen der Einbeziehung kultureller Einrichtungen in den Anwendungsbereich der PSI-Richtlinie.

31 Recitals 1, 8, 25 (2003/98/EC).

32 Cf. Recital 9 (2003/98/EC), see Drexl (*supra* n 30) and Lundqvist (*supra* n 30).

33 Recitals 3, 5 (2013/37/EU).

is in line with several national and multilateral OGD initiatives, which have gathered momentum in the last 10 years.[34]

14 For achieving its objectives, the Directive requires public sector bodies (PSBs) to make information **re-usable** for commercial and non-commercial purposes under non-discriminatory conditions for comparable categories of re-use. Charges are limited to the marginal costs of reproduction, provision, and dissemination. Also, PSBs may not unnecessarily restrict re-use and have to justify if they grant exclusive rights for re-use. It is important to mention that the PSI Directive sets out a minimum standard. Therefore, Member States are free to pursue more re-use-friendly policies. Furthermore, in the course of the Directive's revision in 2013, libraries, museums and archives have also been included within its scope. However, they are subject to a specific regime regarding re-use, charging, and exclusive arrangements.

2. Research and educational exemption

a.) History of the exemption

15 Amongst other types of information and institutions, **research and educational establishments have been explicitly exempted** from the Directive's scope. Article 1(2)(e) states that the PSI Directive is not applicable to "documents held by educational and research establishments, including organizations established for the transfer of research results, schools and universities, except university libraries". While this exemption had been included in the PSI Directive in 2003,[35] its deletion was considered for the revision in 2013. Even though the high economic and social value of the re-use of R&E establishment's information holdings had been recognized, several reasons – as already outlined in the introduction – were put forward in disfavor.[36] Only a definition of university has been included, which should enable a more accurate delineation of university libraries to which the PSI Directive was extended.

b.) Affected establishments

16 In general, to fall within the scope of the PSI Directive, information must be held by a PSB. Article 2(1) and (2) define PSB by following the definition of public procurement rules.[37] While the legal entity's form is irrelevant, it must be predominantly controlled by the state, be it by financial or managerial means. Many R&E establishments across the EU would meet this definition. There is neither a PSI-specific nor an EU-wide definition of **'research establishment'**. The legislature has deliberately refrained from defining this term, due to the problem of subsidiarity and the different traditions within the Member States.[38] However, a recent approach has been made in the course of the ongoing copyright reform, where the Commission defined a research organization as an organization "the primary goal of which is to conduct scientific research or conduct scientific research and provide educational services".[39] There is no doubt that the PSI Directive would accommodate establishments dedicated to basic and applied research. Such establishments can be independent or held by universities or other organizations. Article 1(2)(e) makes clear that this also includes organizations established for the transfer of research results. However, should they be organized as public undertakings, the definition of PSB is not met (Recital 10 (2003/98/EC)).

17 Also, the PSI Directive does not clearly define **educational establishments**. Article 2(9) defines the term 'university' as a PSB that provides "post-secondary-school higher education leading to academic degrees". But 'educational establishment' not only refers to higher education, but also to schools for primary and secondary education.[40]

18 As the PSI Directive is already applicable to **university libraries** (Article 1(2)(e)), a deletion of the exemption would also bring research related libraries other than university libraries within the scope of the PSI Directive. This would also eliminate major uncertainty associated with the problem to determine whether a university library has legal personality – and is therefore to be considered as PSB – or forms a mere part of the university itself.[41]

34 See the "Memorandum on Transparency and Open Government" (2009) of former U.S.-president Obama; the G8 Open Data Charter of 18 June 2013; for the EU strategy on Open Data see COM IP/11/1524.

35 On the attempt to remove the exemption during the legislative procedure already in 2003 see Pas, J./De Vuyst, B. (2004), Re-Establishing the Balance Between the Public and the Private Sector: Regulating Public Sector Information Commercialization in Europe, 9 Journal of Information Law & Technology.

36 SEC(2011) 1152 final, 33 et seq., explicitly referring to generated scientific data (observational, experimental data, databases), patents, scientific publications, unpublished material (pre-prints, non-refereed publications), output of educational establishments (such as theses, lectures, conference proceedings).

37 See Recital 10 (2003/98/EC).

38 See SEC(2011) 1152 final, 34.

39 See Article 2(1) COM(2016) 593 final – Proposal for a Directive of the European Parliament and of the Council on copyright in the Digital Single Market.

40 More detailed Richter (*supra* n 1) at § 1 para. 514 et seq.

41 See Guibault, L./Salamanca, O. (2017), Public sector

c.) Affected information

19 The PSI Directive applies to all existing 'documents' held by a PSB (see Article 1(1)).[42] According to Article 2(3) the rather old-fashioned term 'document' means "any content whatever its medium (written on paper or stored in electronic form or as a sound, visual or audiovisual recording)". Therefore, 'document' can be understood as reproducible[43] information. For the sake of simplicity, the term 'information' is used in this sense in the following.[44] **The broad definition** of 'document' shows that the PSI Directive applies to a vast range of information. This contains structured and unstructured data, raw data, meta data or compiled data. Also, the form of the media in which data are recorded is irrelevant. Information can therefore be digital files or physical devices of text documents, numerical data, spreadsheets, charts, notebooks, questionnaires, test responses, transcripts, codebooks, images, videos, audios, slides, reports etc. However, the PSI Directive does not apply to one important form of research output: according to Recital 9 (2003/98/EC) the definition of 'document' does not cover computer programs. Therefore, the PSI Directive does not apply to software.

20 One can see that both the amount and variety of **information held by R&E establishments** are infinite. However, the most valuable – and therefore relevant – information for re-use can be grouped into three categories – research, educational and administrative information:

- **Research information** can be broadly divided into research data and publications. Due to the broad meaning of 'document' according to Article 2(3), the scope of both categories is much wider than defined under the conventional OA-regimes: research data can be collected or created and is usually stored in databases. In particular, research data consists of a broad variety of *observational data*, such as data captured and transmitted in real-time by sensors, survey data, sample data, neurological images, or clinical trials. Just as important is *experimental data*, which is the outcome of a test method or an experimental design. This concerns e.g. data from lab equipment, such as gene sequences, chromatograms or magnetic field data. In general, research data can be derived from other data elements or compiled from a number of different sources. Furthermore, all publications can be qualified as 'documents' within the meaning of Article 2(3). This contains scientific publications,[45] no matter if refereed or not, as well as datasets linked to them. In principle, monographs and articles are affected, as well as drafts and unpublished material.[46]

- Information that is usually directly related to **education** comprises multimedia material, lecture manuscripts and slides, recorded lectures, theses and conference proceedings,[47] as well as exams.

- Both educational and research establishments hold huge amounts of **administrative information**. This includes e.g. information regarding planning, budgets, correspondence, human resources and statistics. Some information is closely related to research, such as project information, contracts with funders and plans for future research. Other administrative information is related to teaching, such as timetables, room plans, curricula, examination regulations, enrolment statistics, course descriptions and evaluations.

C. Deletion of the research and educational exemption

I. Overview

21 What would happen if one completely deletes the R&E exemption of the PSI Directive? Taking this **extreme position** as a starting point helps to understand how the legal mechanisms of the PSI Directive work and to what extent and in what way this would effectively influence the creation and use of information held by R&E establishments. At first glance, the PSI Directive becomes applicable to all the information as outlined above. However, a detailed look at the scope of the PSI Directive shows that a significant amount of information would be excluded due the filtering function of other exemptions provided for in Article 1(2). Moreover, Member States and R&E establishments themselves can influence

information and university libraries, in: Wiebe, A./Dietrich, N. (eds.), Open Data Protection, Study on legal barriers to open data sharing – Data Protection and PSI, 228 et seq.

42 See also Recital 11 (2003/98/EC): "A document held by a public sector body is a document where the public sector body has the right to authorise re-use." However, it seems debatable if the legal question of the right to authorize should be part of the definition that constitutes what a 'document' or an 'information' is; for discussion see Richter (*supra* n 1) at § 2 para. 73.

43 The requirement of reproducibility is based on the idea that a PSB can transfer the information to persons without losing the information itself.

44 Also national implementation refers to both, see e.g. § 2(2) IWG (Germany) referring to 'information', while § 4(2) addresses 'documents'.

45 See SEC(2011) 1152 final, 33.

46 *Ibid.*

47 *Ibid.*

whether they would fall under the exemptions to a considerable extent. For the remaining information of R&E establishments that fall under the scope of the PSI Directive, the legal consequences must be carefully considered. While the aim is not to draw a universally applicable conclusion on the economic effects, some crucial problems of the PSI Directive can be spotted with relevance for information held by R&E establishments in particular.

II. Applicability of the PSI Directive

1. Overview

22 Article 1(2) sets out several **exemptions** under which the PSI Directive does not apply. Many of them seem clear at first glance, but at second glance there appears to be considerable legal uncertainty about their interpretation. This causes a general problem for answering the question regarding how relevant the PSI Directive ultimately is for information held by R&E establishments. In the following, the analysis focuses on three exemptions that appear to be most relevant for the information in question. Put in positive terms: the PSI Directive can only apply if the information concerned is accessible in an unrestricted manner (*sub* 2.), if no intellectual property rights are held by a third party (*sub* 3.), and if the supply of this information falls under the scope of the PSB's public task (*sub* 4.). In principle, the PSI Directive can be applicable to personal data, but it does not affect data protection laws of the EU or the Member States in any way (*sub* 5.).

2. Unrestricted accessibility

a.) Legal standard

23 The PSI Directive only concerns the re-use of information and **does not regulate access to information**. Article 1(3) makes clear that the PSI Directive "builds on and is without prejudice to access regimes in the Member States". The Directive does not contain an obligation concerning access to documents.[48] Therefore, it largely[49] remains in the domain of the Member States and PSBs what information they choose to make accessible.[50] The reason lies in the limited competencies of the EU legislature to generally regulate access to information of national PSBs.[51] For that reason, Article 1(2)(c) exempts "documents which are excluded from access by virtue of the access regimes in the Member States". Article 1(2)(c) also exempts "documents access to which is restricted by virtue of the access regimes in the Member States, including cases whereby citizens or companies have to prove a particular interest to obtain access to documents". Furthermore, Recital 9 (2003/98/EC) states that the PSI Directive "should apply to documents that are made accessible for re-use when public sector bodies license, sell, disseminate, exchange or give out information."

b.) Unrestricted right of access to information

24 Taking all these references together, there is no doubt that documents fall under the PSI Directive if **national access regimes provide unrestricted access** to these documents to everyone. The exemptions make sure that even those documents shall not be re-usable that are accessible on request in privileged cases or under additional requirements. Otherwise the PSI Directive would undermine legitimate intentions of national legislators for differentiating access regimes and introducing certain requirements, to prevent the risks associated with uncontrolled circulation of information as a consequence of its mandatory re-use.[52]

25 Ultimately, it depends on national legislation whether and to what extent information of R&E establishments is affected. Member States have different regimes in place that grant individual rights of access to information. However, only such access legislation is relevant that grants **unrestricted and unconditional** access rights. This means that the access right may neither apply only to a particular group (e.g. the press or other researchers) nor require a particular justification or proof of interest (e.g. research or educational interest). There are some sector-specific regimes in place that oblige PSBs to make their information accessible in an unrestricted manner.[53] On a cross-sectoral level, freedom-of-

48 See Recital 7 (2013/37/EU).

49 There are cases in which EU legislation provides access to information, e.g. the Directive 2003/4/EC of the European Parliament and of the Council of 28 January 2003 on public access to environmental information and repealing Council Directive 90/313/EEC and Directive 2007/2/EC of the European Parliament and of the Council of 14 March 2007 establishing an Infrastructure for Spatial Information in the European Community (INSPIRE).

50 See Recital 8 (2013/37/EU).

51 Janssen, K./van Eechoud, M. (2012), Rights of Access to Public Sector Information, 6 Masaryk University Journal of Law and Technology 471 at 476 et seq.

52 Access rules are usually the result of a balancing of interests by the legislature.

53 See Directive 2003/4/EC and INSPIRE Directive (2007/2/EC); for details see van Eechoud, M. (2015), Making Access to Government Data Work, 9 Masaryk University Journal of

information (FOI) regulation[54] – should it be in place at all[55] – typically follows this rationale. While FOI sometimes only addresses the core administration (executive agencies), it can also reach out to more independent PSBs, such as R&E establishments in particular.[56]

26 FOI-legislation, however, might contain explicit block **exemptions for R&E establishments** or explicitly exclude certain types of information. The FOI-legislation of the German *Länder* frequently allows access to information of R&E establishments only insofar as the information does not relate to research, education, arts, performance evaluation and examinations.[57] According to the broad interpretation of a higher administrative court, this also covers the underlying contracts for third-party research assignments.[58] Ultimately, this significantly narrows down the scope to mere administrative information. In the U.K., the Freedom of Information Act (FOIA) contains Sec. 22A, which is explicitly dedicated to "information obtained in the course of, or derived from, a programme of research" and therefore to ongoing research. The exemption is subject to a public interest test. So even if access is granted under this rule, the PSI Directive would not apply to this information. The same applies to Sec. 22 of the U.K. FOIA. This clause exempts information that is intended for future publication and is therefore highly relevant for research establishments. On the basis of Sec. 22, access to a PhD thesis[59] and clinical trial data[60] has been successfully denied. In both cases, interests were considered and balanced. So even if access had ultimately been granted, the PSI Directive would not apply to this information.[61]

27 Also, **other exemptions in access regimes** that do not explicitly address R&E establishments can prevent access to information they hold. First and most importantly, access can be denied for reasons of secrecy or sensitivity. Many of the U.S. open-records laws contain exemptions to protect sensitive and research information.[62] There have been cases where these exemptions have effectively prevented public disclosure of their information.[63] Information related to business secrets and unpublished patents are exempt from rights to access. Second, copyright can already prevent access, especially if works are unpublished.[64] As a third rather general category, FOI-legislation can also exempt such information concerning internal operations or activities of bodies, in case the disclosure of such information would cause disturbances in operations or activities of the body.[65]

28 The application of FOI-regulation is particularly peculiar when it comes to research institutions.[66] This

Law and Technology 61 at 70 et seq.

54 The term refers to 'freedom of information acts', 'right to information laws', 'transparency acts' and 'public records acts', see van Eechoud (*supra* n 53) at 64.

55 Not every Member State has FOI-legislation, see e.g. the complex situation in Germany, where four *Länder* do not have FOI-legislation in place (for details see Richter (*supra* n 1) at § 1 para. 142).

56 See e.g. U.S., where the Federal FOIA (1966) and several state 'open-records laws' govern access to records in the possession of federal agencies and state entities, such as public universities. For the situation in the U.K. see <https://www.gov.uk/government/publications/university-and-business-collaboration-agreements-model-agreement-guidance/university-and-business-collaboration-agreements-model-agreement-guidance> at para. 3.68.

57 See § 2(3) IFG North Rhine-Westphalia: "Für Forschungseinrichtungen, Hochschulen und Prüfungseinrichtungen gilt dieses Gesetz nur, soweit sie nicht im Bereich von Forschung, Lehre, Leistungsbeurteilungen und Prüfungen tätig werden."; § 2(3) No. 2 IFG Baden-Wurttemberg: "Dieses Gesetz gilt nicht gegenüber [...] den Einrichtungen mit der Aufgabe unabhängiger wissenschaftlicher Forschung, Hochschulen [...], Schulen [...] sowie Ausbildungs- und Prüfungsbehörden, soweit Forschung, Kunst, Lehre, Leistungsbeurteilungen und Prüfungen betroffen sind."; see also § 2(2) AIG, § 1 SIFG, § 3(1) No. 9 LSA IZG, § 2(5) ThürIFG, § 1(1a) BremIFG.

58 See OVG Münster of 30 September 2016 (4B 601/16) "unmittelbar wissenschaftsrelevante Angelegenheiten wie Drittmittelverträge über Forschungsvorhaben".

59 See U.K., ICO Decision Note FS 50349323.

60 See Queen Mary University London v. Information Commissioner & Mr Robert Courtney [EA/2012/0229] of 22 May 2013.

61 See also, Sec. 39 of the Irish FOIA, which follows a balancing approach when it comes to research.

62 See for a good overview Cardon, A./Bailey, M./Bennett, B. (2012), The Effect of Public Disclosure Laws on Biomedical Research, 51 Journal of the American Association for Laboratory Animal Science, 306.

63 For specific cases see *ibid.* at 308.

64 German Federal Administrative Court of 25 June 2015 – BVerwG 7 C 1.14 (ECLI:DE:BVerwG:2015:250615U7C1.14.0). However, copyright protection does not automatically protect the freedom of science (critical in this respect VG Braunschweig ZD 2014, 318, which was heavily criticized for good reasons, see Schnabel, ZD 2014, 318 and Schoch, F. (2016), Informationsfreiheitsgesetz (IFG), § 6 para. 24.

65 See e.g. Article 6(1) No. 11 of the Slovenian Zakon o dostopu do informacij javnega značaja (UPB2, Official Gazette of the Republic of Slovenia, No. 51/06); see also standard under U.K. FOIA according to which information is exempt from access if disclosure of the information would prejudice someone's commercial interests and the public interest in withholding the information outweighs the public interest in disclosing it.

66 See the U.S. discussion of the OMB Circular A-110 amendment of 1999, which requires researchers to ensure that "all data produced under a [federally funded] award will be made available to the public through the procedures established under the Freedom of Information Act". For a valuable illustration of the controversy whether FOIA is well placed to allow wider public access or rather harms the traditional process of scientific research, see Fischer, E. (2013), Public Access to Data from Federally Funded Research: Provisions in OMB Circular A-110, Congressional

is because the **fundamental right of the freedom of science** applies to publicly employed and/or funded researchers and seeks to protect their independence and autonomy.[67] This causes systematic tensions with the principle of informational freedom and might unduly affect the way research is being conducted. Most of the FOI-inquiries to research establishments have therefore addressed administrative information.[68] Member States have to strike a balance between the concerned interests. Therefore, the landscape of regulatory regimes that provide access to information of R&E establishments is greatly diverse. One has to keep in mind that the PSI Directive does not hinder the Member States from adjusting their access policies in whatever direction.

c.) Factual accessibility

29 There is some uncertainty about those documents which are in fact made accessible by PSBs, however, without any obligation to do so[69] and without any corresponding individual right to access. The PSI Directive is not clear on this crucial issue.[70] Commentators at the time expressed doubt as to whether the Directive covers information made public without a clear legal basis.[71] Nowadays, however, the predominant reading of the PSI Directive, in principle, includes **all generally accessible information**. The U.K. legislature states that the PSI Directive does not apply "unless the information has already been provided to a requester, or the information is otherwise accessible to the applicant".[72] Also, the Irish legislature refers to legitimate 'accessibility' as one decisive criterion.[73] In a landmark judgment, the Federal Administrative Court in Germany ruled that the PSI Directive also applies in cases where a PSB publishes information, even if there is no corresponding individual right to access it.[74] Therefore, the re-use of information made accessible by the PSB is also subject to the rules of the PSI Directive. While the German court based its view on Recitals 8 and 9 (2013/37/EU), its reasoning appears weak even though the outcome seems appreciable.[75] Should one require a corresponding individual right to access already accessible data, a hypothetical test would need to be put into place. However, there would not always be a definite answer to this test, because courts have discretion when applying exceptions.

30 Ultimately, the question about the exact **legal standard** for factual access remains. First, the information must have been published by the PSB itself.[76] Second, the information must be made accessible to everyone. This does not require that the information is for free. Most prominently, this case affects published information on websites or information provided in statutory registers or for a fee.[77] However, if information is provided only to one party but not made accessible to everyone, this does not constitute accessibility that triggers the Directive's application.

31 A lot of R&E information is exempt from unrestricted individual access rights. Therefore, its factual accessibility is the predominant case for the PSI Directive being applicable in principle. This mainly affects all information provided on websites of **R&E institutions**. They supply a lot of research results through their own repositories.[78] This concerns text publications as well as datasets. Policies differ regarding the timing, the sort, and the status of publication. A lot of information is provided after an embargo period. This allows for private

Research Service Report for Congress, 16 et seq.; see also regarding public universities: <https://www.theguardian.com/science/political-science/2014/dec/04/should-universities-be-exempt-from-freedom-of-information-requests>.

67 Article 13 CFREU: "The arts and scientific research shall be free of constraint. Academic freedom shall be respected." See also e.g. Article 5(3) of the Basic Law for the Federal Republic of Germany: "Arts and sciences, research and teaching shall be free. The freedom of teaching shall not release any person from allegiance to the constitution."

68 For a good overview on actual requests in Germany see: <https://fragdenstaat.de/suche/?q=universit%C3%A4t>.

69 See for research institutions e.g. § 71a HG NRW (for transparency on third-party funding); § 5(7) HmbTG; § 16(3) TransparenzG RP.

70 Recital 8 (2013/37/EU) requires Member States to make all documents re-usable "unless access is restricted or excluded under national rules on access to documents"; therefore, one could argue that even if documents were made publicly accessible, the PSI Directive would not apply because in theory, access is only possible after proving a particular interest.

71 Janssen/van Eechoud (*supra* n 51) at 476 et seq.

72 See Guidelines from National Archives on "Links between access and re-use" of 2008, para. 2.

73 See Explanatory note S.I. No. 525 of 2015.

74 See German Federal Administrative Court of 14 April 2016 – BVerwG 7 C 12.14 (ECLI:DE:BVerwG:2016:140416U7C12.14.0).

75 Recital 9 (2003/98/EC) can also be interpreted in a way that it only states examples according to which certain activities can be seen as making information accessible for *re-use* (as an additional requirement in addition to access as such). For a critique see Richter, H. (2016), Zur Weiterverwendung von Informationen der öffentlichen Hand: BVerwG klärt erstmals grundsätzliche Anwendungsvoraussetzungen des IWG, 35 Neue Zeitschrift für Verwaltungsrecht 35, 1143.

76 No matter if the law provides for mandatory publication or not.

77 See Guidelines from National Archives on "the implementation of the Re-use of Public Sector Information Regulations 2015 – For re-users", 9.

78 This may not be confused with situations in which research information is published in private repositories.

commercialization of research results as well as for a research advantage over other researchers. When it comes to educational material, establishments make available e.g. slides, videos or other material before or after the lecture. There are also repositories that offer open educational resources (OER). In any case, the information must be accessible by everyone. This is not the case if access requires a log-in and/or a password provided to a particular group (e.g. university members). Furthermore, it affects all administrative information that is publicly accessible on the website (e.g. schedules or course descriptions and statistics).

3. No intellectual property rights of third parties

a.) Legal standard

(aa) Ambiguous standard of the PSI Directive

32 Intellectual property is of significant relevance for information held by R&E establishments. However, the PSI Directive is not applicable if third parties – meaning other parties than the PSB itself – hold intellectual property rights (IPRs). Article 1(2)(b) states that the PSI Directive does not apply to "documents for which third parties hold intellectual property rights". The term **"intellectual property rights"** within the meaning of Article 1(2)(b) only refers to copyright and related rights (including *sui generis* forms of protection).[79] Therefore, the PSI Directive does not apply to information covered by industrial property rights, such as patents, designs and trademarks.[80] This interpretation is of highest relevance for research establishments and public technology transfer institutions, because a lot of valuable information they hold does not fall under the scope of the PSI Directive at all. Recital 22 (2003/98/EC) underlines that IPRs of third parties are not affected by the PSI Directive and clarifies the relationship between the PSI Directive and IPRs by stating that the PSI Directive does "not affect the existence or ownership of intellectual property rights of public sector bodies, nor does it limit the exercise of these rights in any way beyond the boundaries set by this Directive". At least in theory, the rationale of the PSI Directive regarding IPRs seems straightforward:

- Should **third parties hold IPRs**, the PSI Directive is not applicable. However, under certain circumstances, Article 4(3) PSI Directive obliges PSBs to name the rightholder;
- Should the **PSB itself hold IPRs**, the PSI Directive can be applicable; if so, the Directive affects the exercise of these rights.[81] The PSB has to make such documents re-usable. Licensing according to Article 8 PSI Directive plays a seminal role then. Also, all the other requirements as set out in Article 3 et seq. apply;
- Should the information **not fall under any IPR protection**, re-use is possible also as set out by Article 3 et seq., but IPR licensing plays no role.[82]

33 However, the crucial question is **how to define whether 'third parties hold' IPRs** or the PSB itself does. The PSI Directive is not clear on this binary criterion. This has been criticized for good reasons.[83] As a closer look at the legal status of R&E information demonstrates, clarification on this point is urgently needed.

(bb) Copyright

34 Whether third parties hold copyright under the meaning of Article 1(2)(b) requires looking at the different copyright systems of the Member States. In general, copyright gives an exclusive right to the natural person who has created the work. However, there are significant differences regarding the status of ownership between common and civil law systems.

- In **common law copyright systems**, there are basically three ways PSBs themselves – and not third parties – can be considered to "hold" IPRs. First, the governments of a number of Commonwealth realms are subject to Crown Copyright.[84] Second, if the creation is a "work for hire", the person who employs someone to create the work (usually a legal entity) is the first owner of copyright and not the actual creator, unless there is an agreement to the contrary.[85] Third, common law copyright systems also allow

79 See Recital 22 (2003/98/EC).

80 See Recital 22 (2003/98/EC); see also SEC(2011) 1152 final, 33.

81 See below, especially Article 8 applies.

82 Therefore, Art. 1(2)(b) does not exclude public domain information. However, there is one exception to that rule related to cultural PSBs (see fiction of IP-protection according to Recital 9 (2013/37/EU), addressing the case where a third party was initial owner of the document; for good reasons critical Jančič/Pusser/Sappa/Torremans (*supra* n 6) at 366).

83 See Drexl (*supra* n 30) at 71.

84 According to Sec. 163 of the U.K. Copyright, Designs and Patents Act of 1988, Crown copyright applies "[w]here a work is made by Her Majesty or by an officer or servant of the Crown in the course of his duties".

85 See e.g. Sec. 11(2) of the U.K. Copyright, Designs and Patents Act of 1988.

for an assignment, which enables an irrevocable and permanent transfer of ownership to the PSB.[86] In all cases, no third parties would hold copyright within the meaning of Article 1(2) (b) PSI Directive. Thus, in principle, many copyright-protected works could fall under the PSI Directive in common law copyright systems.

- In contrast, only the natural person as creator qualifies as author in **civil law authors' rights systems**.[87] Due to their roots in personality rights, authors' rights cannot be transferred by assignment. Therefore, "ownership" always stays with the author. The sole form of a contractual transfer of rights is the grant of a license. However, even in that case the author retains ownership and only authorizes certain acts to be carried out. Obviously, in civil law authors' rights systems a literal reading of 'documents for which third parties hold intellectual property rights' would exclude almost all copyright-protected works from the scope of the PSI Directive. This strict interpretation is not just theory. In fact, commentators have put forward this reading of the PSI Directive[88] and also the German government seems to follow this interpretation, even if the PSB as an employer enjoys an exclusive license in works created by its employees.[89] In that regard, Recital 12 (2013/37/EU) causes confusion and has been interpreted contrary to its original intention.[90]

35 Taking the rationale of the PSI Directive into account, this strict interpretation is not convincing. At least in those cases in which a **license granted to a PSB can be seen as a functional equivalent** to a transfer of ownership by assignment, there are good reasons to treat both cases similarly – provided that the re-use does not affect the interests of the author by any means. This is the case if exclusive licenses are granted and the PSB as licensee is the sole party that is allowed to sub-license.[91] However, the situation becomes blurred if the exclusive license is limited or revocable. Moreover, mandatory legislation can limit the duration of exclusivity.[92] What would be the legal consequence? On the one hand, one could argue that the PSI Directive applies, but the PSB has to take restrictions into account when sub-licensing.[93] As a consequence, one has to decide on a case-by-case basis if the PSI Directive applies.[94] On the other hand, there are also good reasons to hold the PSI Directive not applicable in such situations, due to the potential risk of ultimately affecting the author's interests.[95]

(cc) Database protection sui generis

36 Article 7 et seq. of the Database Directive (96/9/EC) (DB Directive) regulate the *sui generis* right for the protection of databases. The Directive has been implemented in the Member States and it harmonizes the legal treatment of databases to a large extent. The beneficiary of the right is the **"maker of the database"**,[96] whom Recital 41 defines as "the person who takes the initiative and the risk of investing; whereas this excludes subcontractors in particular from the definition of maker". In contrast to authors' rights, legal entities can also be qualified as makers and therefore hold the *sui generis* right. As a consequence, the *sui generis* right is highly relevant for PSI and for research establishments in particular as they have vast holdings of databases.

37 There is, however, a considerable and ongoing debate about **whether and to what extent databases of PSBs enjoy protection**.[97] The legal situation

86 See e.g. Sec. 90(1) of the U.K. Copyright, Designs and Patents Act of 1988.

87 There might be few exceptions, however, as e.g. in the case of inheritance (see § 28 of the German Act on Copyright and Related Rights (UrhG)).

88 See Dreier, T./Spiecker gen. Döhmann, I. (2016), Gegenrechte – Datenschutz/Schutz von Betriebs- und Geschäftsgeheimnissen, Geistiges Eigentum, in: Dreier, T./Fischer, V./van Raay, A./Spiecker gen. Döhmann, I. (eds.), Informationen der öffentlichen Hand – Zugang und Nutzung, 191; Wirtz (*supra* n 30) at 165.

89 See BT-Drs. 18/4614, 20. In Germany, employment agreements frequently grant the employer an exclusive license to any works the employee creates within the scope of obligations.

90 Recital 12 (2013/37/EU) states that the PSI Directive "should be without prejudice to the rights, including economic and moral rights that employees of public sector bodies may enjoy under national rules".

91 Without further reasoning Wiebe, A./Ahnefeld, E. (2015), Zugang zu und Verwertung von Informationen der öffentlichen Hand – Teil II, 2015 Computer und Recht 199, 205; see also Drexl (*supra* n 30) at 71 suggesting that this should be within the scope of the PSI Directive, however, casting doubts if this is the case *de lege lata*.

92 Also, there are jurisdictions if lump sum payment for exclusive license, exclusivity is reduced to 10 years (see § 40a UrhG).

93 See Richter (*supra* n 1) at § 1 para. 412.

94 This seems to be the U.K. standpoint: Sec. 5(1)(b) of The Re-use of Public Sector Information Regulations 2015 excludes documents if "a third party owns *relevant* [emphasize added] intellectual property rights in the document".

95 One also has to consider high transaction costs related to rights clearance. However, this is rather a policy consideration than a legal argument in this context.

96 See Article 7(1) DB Directive.

97 See for discussion Derclaye, E. (2008), Does the Directive on the Re-use of Public Sector Information affect the State's database sui generis right?, in: Gaster, J./Schweighofer, E./Sint, P. (eds.), Knowledge rights – Legal, societal and related technological aspects, Austrian Computer Society, 137, 161; Guibault/Wiebe (*supra* n 15) at 32 et seq. The CJEU as acknowledged *sui generis* protection for state databases, see ECLI:EU:C:2012:449 – *Compass Datenbanken* and also for a database created by academic staff of a publicly funded

significantly differs between Member States. The Netherlands explicitly grant no *sui generis* protection for public databases, unless the right is reserved explicitly by public act.[98] German courts have applied the exemption for official works to official databases by analogy.[99] In contrast, Austrian courts have explicitly recognized *sui generis* protection for official databases.[100] The legal standard on the EU-level is not entirely clear.[101] In the *Compass-Datenbank* case the CJEU implicitly accepted *sui generis* database protection for a public register in Austria.[102] However, this does not clarify whether the DB Directive allows Member States to set aside protection for public databases. Also, there are good reasons to doubt whether the making of a tax-funded database is subject to a "risk" as apparently required by Recital 41 DB Directive.[103] Taking this legal uncertainty into consideration, it is more than welcomed that the European Commission has explicitly addressed this issue in its public consultation on the DB Directive.[104]

38 Whether or not a concrete database qualifies for protection depends – according to Article 7(1) DB Directive – on the substantiality of **"investment in either the obtaining, verification or presentation of the contents"**. Besides the vagueness of the substantiality requirement, EU jurisprudence has repeatedly confirmed that the creation of content does not qualify as 'obtaining', and therefore respective investments are not to be taken into account.[105] However, great uncertainty remains about the demarcation between 'creating' and 'obtaining'. This is particularly relevant when it comes to investment into sensor-generated and measurement data.[106]

39 Should the PSB as the legal entity be qualified as 'maker' of the database, the PSI Directive can apply. Before its amendment in 2013, there was some discussion about its relationship with the DB Directive.[107] However, since the general principle (Article 3) of the PSI Directive has been changed, the relationship seems clear: if a PSB is the maker of the database, the database can in principle enjoy protection under Article 7 et seq. DB Directive. At the same time, the PSI Directive can be applicable to the content of the database. Should re-use of that content require a substantial extraction within the meaning of Article 7(2)(a) DB Directive, Article 3 obliges the PSB to license according to Article 7(3) DB Directive. Article 8 sets out further conditions for licensing as discussed below.

b.) Research establishments

(aa) Overview

40 In general, the IPR exemption will exclude a large amount of information held by R&E establishments.[108] But due to the uncertainty about the exact legal standard of the IPR exemption, one can base its relevance for research establishments only on **assumptions**. Especially scientific publications (like articles, books, manuscripts etc.) are **copyright protected works**.[109] Databases can enjoy copyright and/or *sui generis* protection. Mere data will be discussed separately. Should information be non-protected subject matter, re-use is possible as set out by Article 8.

(bb) Works

41 The crucial question concerning protected works is whether copyright has been assigned/transferred to the PSB and no third party holds IPRs. In general, this depends on who creates the work under which circumstances. However, in academia the (fundamental) right of academic freedom affects the interpretation and application of copyright laws. Predominantly, the status of works created by **academic staff** (not by administrative staff) is concerned. Usually, employed researchers have no obligation to write on a particular topic or to publish in a particular form – these matters are for them to determine.[110] Universities or other research

university, see ECLI:EU:C:2008:552 – *Directmedia Publishing GmbH vs. Albert-Ludwigs-Universität Freiburg*.

98 See Guibault/Wiebe (*supra* n 15) at 65 et seq.

99 German Federal Court of Justice of 28 September 2006 (I ZR 261/03) – *Sächsischer Ausschreibungsdienst*.

100 OGH of 9 April 2002 (4Ob17/02g).

101 German Federal Court of Justice of 28 September 2006 (I ZR 261/03) – *Sächsischer Ausschreibungsdienst*.

102 ECLI:EU:C:2012:449 – *Compass Datenbanken*, para. 47.

103 See Guibault/Wiebe (*supra* n 15) at 66 with reference to the Dutch "landmark case" where the District Court of Amsterdam ruled that a City Council did not qualify as a "producer of a database" and therefore did not own any database right in the information it gathered.

104 See <https://ec.europa.eu/digital-single-market/en/news/commission-launches-public-consultation-database-directive>.

105 See ECLI:EU:C:2004:695 – *The British Horseracing Board*; ECLI:EU:C:2004:697 – *Fixtures Marketing*.

106 For a detailed discussion on the protection of industrial data Wiebe, A. (2016), Protection of industrial data – a new property right for the digital economy?, 65 GRUR Int. 877.

107 See Derclaye (*supra* n 97) at 161, answering the question whether the PSI Directive affects the *sui generis* right with "not sure, not really or absolutely no".

108 SEC(2011) 1152 final, p. 33.

109 On the more fundamental discussion whether there should be copyright protection for scientific works at all see Shavell, S. (2010), Should copyright for academic works be abolished? 2 Journal of Legal Analysis 301; see also Scheufen (*supra* n 8) at 47 et seq., 143.

110 See Barendt, E. (2010), Academic Freedom and The Law: A

institutions do not supervise their work.[111] For that reason in all democratic, free societies, academic freedom strengthens their positions – at least to some extent – when it comes to copyright. In particular, the 'works for hire' doctrines do not strictly apply. There is a long-standing discussion about the "teachers' exception"[112] in the U.S., according to which academic freedom exempts professors and other academics from the work for hire doctrine.[113] As there is, however, still uncertainty about the current legal status,[114] key provisions concerning the definition of ownership vary by institutional policy and factual context.[115] Also the U.K. effectuates academic freedom in its copyright system[116] and defines ownership of faculty-created works through university policies.[117] German law and jurisprudence put great emphasis on the author's academic freedom.[118] The freedom of science under Article 5(3) Basic Law gives the freedom to determine if, when and how to publish their materials.[119] Courts have ruled that the presumption of employees granting a license to their employer does not apply *per se* to professors of a publicly funded university.[120]

42 These modifications affect mostly **those works** that are freely created by academic staff, especially scholarly work, such as articles, papers and books.[121] Also theses and dissertations usually fall into this category.[122] Even if there is a default assignment of rights for theses, their creators are usually eligible to apply for a waiver.[123] Furthermore, should an academic employee transfer ownership of copyright or grant licenses to the research establishment due to OA-policies, he retains some rights or grants licenses to the PSB on a non-exclusive basis.[124] Thus, even under a strict interpretation of Article 1(2)(b), the academic author as a 'third party' (still) 'holds' IPRs. As a consequence, the PSI Directive is inapplicable.

43 For **non-scholarly works and works of non-academic employees**, academic freedom does not apply (as e.g. administrative staff). Whether Article 1(2)(b) applies depends on the general principles and the interpretation of the (unclear) standard as stated above. The same holds true for third parties, in case the research establishment commissions or funds a work. The copyright status depends on the particular policy or agreement.

44 Should works be **funded by third parties**, the copyright status depends on the funding policies as well. Specific provisions regarding ownership, retention of and access to data can be included into the agreement. The funder can retain rights and set OA-mandates as mandatory.[125] While one can see some natural tensions with academic freedom here as well, the concern seems less than in case

Comparative Study, 216.

111 See for example § 1(6) of the Act of Higher Education (1992:1434) in Sweden: Research problems are to be freely chosen, research methods are to be freely developed, research results are to be freely published, see Carlson, L. (2016), Academic Freedom and Rights to University Teaching Materials: A Comparison of Swedish, American and German Approaches (January 10, 2016). Available at SSRN: <https://ssrn.com/abstract=2713421/>, 359.

112 Barendt (*supra* n 110) at 217 with reference to C. McSherry (2001), Who owns academic work?; see for a history in the U.S. Rooksby, J. (2016), Copyright in Higher Education: A Review of Modern Scholarship, 54 Duquesne Law Review 197; also Carlson (*supra* n 111) at 375 et seq.

113 Barendt (*supra* n 110) at 217 et seq. referring to Williams c. Weisser 273 Cal App 2d 726 (cal App 1969); Weinstein v. University of Illinois 811 F 2d 1091 (7th Cir 1987); Hays v. Sony Corporation of America 847 F2d 412 (7th Cir 1988); but with no clarification regarding the survival of the teacher exception, see Gertz, G. (2013), Copyrights in faculty-created works: How licensing can solve the academic work-for-hire-dilemma, 88 Washington Law Review 1465, 1473.

114 See Gertz (*supra* n 113) at 1482 et seq.

115 See Rooksby (*supra* 112) at 216 regarding U.S. with further reference; see also for different policies at U.S. universities Carlson (*supra* n 111) at 379 et seq.

116 See for particular examples for universities in the U.K. Barendt (*supra* n 110) at 217.

117 See Gertz (*supra* n 113) 1465; Rooksby (*supra* 112) at 206: But rather based on university policies than on case law or statute.

118 See Carlson (*supra* n 111) at 383 et seq.

119 See Barendt (*supra* n 110) at 218.

120 German Federal Court of Justice of 27 September 1990 (I ZR 244/88).

121 See University of Reading (2010), Code of Practice in Intellectual Property, Commercial Exploitation and Financial Benefits of 16 June 2010, para. 7.21, which defines this as works "produced solely in the furtherance of an academic career, such as articles in journals, papers for conferences, study notes not used to deliver teaching and books not commissioned by the universities".

122 Dissertation in Germany are clearly no works for hire, see Leuze, D. (2006), Die Urheberrechte der wissenschaftlichen Mitarbeiter, 108 GRUR 552, 553; see also the exemption from ownership assignment at University of Reading (*supra* n 121) at para. 7.17.

123 See MIT copyright policies on theses, available at <https://odge.mit.edu/gpp/degrees/thesis/copyright/>: "The Institute will retain ownership of the copyrights to theses only if the thesis research is performed in whole or in part by the student with financial support in the form of wages, salary, stipend, or grant from funds administered by the Institute, and/or if the thesis research is performed in whole or in part utilizing equipment or facilities provided to the Institute under conditions that impose copyright restrictions. In all other cases, ownership of a copyright shall reside with the student."

124 While there is no 100 % proof, one gets a fairly good sense when comparing OA-policies and -mandates of research institutions on <http://roarmap.eprints.org>.

125 On the crucial question of who retains the rights, see Scheufen (*supra* n 8) at 116 et seq.

of scholarly work.[126] Usually, there is no transfer of ownership made or an exclusive license granted to research establishments.

(cc) Data

45 The situation becomes even less clear, when it comes to **data**. Mere data, if defined as "raw data" or as processed data only to a very limited extent (put into a database), are not protected by IPRs.[127] As there is no right *in rem*, there is no IP-protected ownership right. This is for good reasons – facts are free and usually data document these facts.[128] As a consequence, research data – whether empirical, observed or measured – are in the public domain, assuming they are not protected as works.[129] This applies to separate items of research data as well as to datasets.[130]

46 Whether the *sui generis* right ultimately protects **databases** must be decided on a case-by-case basis. If employees of the research institution create databases, the *sui generis* protection right – leaving aside the problem if PSBs can be seen as makers at all – is usually held by the PSB as legal entity. The *Directmedia* case[131] nicely illustrates that: a research project at the University of Freiburg led to a publication of an anthology, a collection of verse from 1720 to 1933. The CJEU acknowledged the University of Freiburg, a public university, as the maker of the database and therefore as beneficiary of the *sui generis* right. At the same time, the project leader had been acknowledged as copyright holder for a database work.[132] The requirements as set out by Article 7 et seq. DB Directive have to be met. There is, however, considerable uncertainty about their interpretation.[133] A research database will not necessarily involve a substantial investment (meaning time, money and effort). Furthermore, even if investment is held to be substantial, it must be made in either the obtaining, verification or presentation of its contents. According to settled jurisprudence, the 'obtaining' means collecting existing data and not 'creating' new data.[134] This distinction is important when it comes to different ways of creating databases in the research process. It can make a decisive difference whether the independent items of the database are to be qualified as empirical, observed or measured data.[135] Ultimately, it might occur quite randomly if a research database is protected or not. Should it be protected, extraction and re-utilization require the PSB's consent in principle.[136]

47 Research establishments have developed different **data "ownership" policies**. Frequently, they establish guidelines or mandates that claim ownership of primary data.[137] However, the term ownership is misleading, as there is – by definition – no ownership in facts and data documenting those facts. For that reason, such policies can be seen as mere contractual terms, which bind the establishment's employees or members only. While one could discuss these policies under the aspect of academic freedom as well, there is not (yet) much debate about that as compared to publications. However, when it comes to OA-mandates that require scientists to disclose their data to publications, this might have a negative impact on data generation and on the timing of publications.[138]

126 See Leuze (*supra* n 122) at 559.

127 Only if datasets do not contain protected works; see SEC(2011) 1152 final, 33; definitions for research data largely differ, see e.g. Guibault/Wiebe (*supra* n 15) at 17; Hartmann, T. (2013), Urheberrechtliche Schutzfähigkeit von Forschungsdaten, in: Taeger, J. (ed.), Law as a Service (LaaS) – Recht im Internet- und Cloud-Zeitalter, 508.

128 For a differentiated discussion Kim, D. (2017), No One's Ownership as the Status Quo and a Possible Way Forward: A Note on the Public Consultation on Building a European Data Economy, 66 GRUR Int. 697.

129 Certainly, photos, diagrams etc. can be seen as "data" that are protected subject matter.

130 When it comes to metadata of datasets, copyright protection significantly depends on the content. For a practical overview: <https://irights.info/artikel/eigentum-an-metadaten-urheberrechtliche-aspekte-von-bestandsinformationen-und-ihre-freigabe-2/26829>.

131 ECLI:EU:C:2008:552 – *Directmedia Publishing GmbH vs. Albert-Ludwigs-Universität Freiburg*.

132 ECLI:EU:C:2008:552 – *Directmedia Publishing GmbH vs. Albert-Ludwigs-Universität Freiburg*, para. 15; the DB Directive distinguishes between copyright protected databases (Art. 3 et seq. DB Directive) and the *sui generis* right (Art. 7 et seq. DB Directive).

133 The ambiguous results of the database consultation 2017 confirm this, see <https://ec.europa.eu/digital-single-market/en/news/summary-report-public-consultation-legal-protection-databases>.

134 See ECLI:EU:C:2004:695 – *The British Horseracing Board*; ECLI:EU:C:2004:697 – *Fixtures Marketing*.

135 As guidance for research data in particular see de Cock Buning, M./Ringnalda, A./van der Linden, T. (2009), The legal status of raw data: a guide for research practice, Center for Intellectual Property Law (CIER), 25 et seq.

136 Of course, limitations as set out in Article 8 DB Directive may apply.

137 See e.g. Columbia University, <https://research.columbia.edu/content/ownership-data>: "research data and other records of University research belong to University, except [...]"; see also University of Bristol, <http://www.bristol.ac.uk/research/environment/governance/research-data-policy/>: "Where no external contract exists, the University normally has ownership of primary data generated in the course of research undertaken by researchers in its employment".

138 See Mueller-Langer, F./Andreoli Versbach, P. (2014), Open Access to Research Data: Strategic Delay and the Ambiguous Welfare Effects of Mandatory Data Disclosure (June 20, 2014). Max Planck Institute for Innovation & Competition Research Paper No. 14-09, available at <https://ssrn.com/abstract=2458362>, who point to the problem of

c.) Educational establishments

48 Educational establishments hold a vast amount of protected subject matter. Their staff creates **teaching and learning materials** of which the copyright status is just as disputed as the status of research results.[139] However, academic freedom in relation to the work for hire doctrine could be limited, taking the argument into account that – as opposed to research – the academic staff is obliged to do coursework. Universities follow different policies. E.g. the University of Reading presumes an assignment of ownership to the establishment if the material had been produced within the context of the employee's course duties, meaning in connection with a university course/module/program.[140] This also includes handouts, summaries, case studies, seminar papers, exams[141] and syllabi.[142] In general there is a tendency towards a stronger position of the educational establishment when it comes to Internet-based materials for distance learning courses or MOOCs. In that case, materials are typically commissioned by a university itself, which affects the presumption about copyright ownership.[143]

49 As **students** are not employed, the work for hire rationale cannot apply. However, an assignment of copyright can be required under special circumstances.[144]

d.) Ongoing EU copyright reform

50 As can be seen, clarifying the relationship between copyright protection and PSI Directive is crucial. Currently, the interface is further complicated in the course of the ongoing **copyright reform**.[145] Over the last years, there has been a tendency to introduce copyright exceptions for research purposes. A major step into this direction is an EU-wide copyright exception for text and data mining. However, there is some considerable uncertainty about beneficiaries, addressees and purposes of such a provision.[146] In the context of PSI, the relationship between copyright exceptions and the PSI-standard needs to be clarified. Copyright exceptions can effectively allow for less or more re-use than Article 3(1) does. The problem has been recognized already for cultural PSBs and was implemented in Article 3(2), as will be shown below.

4. Activity falling within the public task

a.) Legal standard

51 The public task is a **crucial criterion for delineating the scope and the application of the PSI Directive**. Article 1(2)(a) contains another important exception by constituting that the PSI Directive does not apply to "documents the supply of which is an activity falling outside the scope of the public task of the public sector bodies concerned". Also, it follows from Article 2(4) that the PSB itself is considered to be a re-user if it uses the information for purposes outside of the public task.[147] Due to this double relevance, the application of the PSI Directive heavily relies on where and how the demarcation line between falling 'within' and 'outside' the public task is drawn. This is to be determined by the Member States and by the PSBs in particular.

52 Recital 9 (2003/98/EC) states performing activities falling outside the public task "will typically include supply of documents that are produced and charged for exclusively on a commercial basis and in competition with others in the market." One can see that the conception of the PSI Directive is influenced by the Anglo-Saxon perception that the State can lawfully act outside of its public task. The National

"strategically delay" in order to fully exploit their data in subsequent research.

139 For a comparison of copyright status between Sweden, U.S. and Germany Carlson (*supra* n 111) at 357 et seq.; Rooksby (*supra* n 112) at 203; in the U.K. [1951] 69 RPC 10 on lecture notes of an accountant; the situation is not clear in Sweden, see Carlson (*supra* n 111) at 366.

140 See University of Reading (*supra* n 121) at paras. 5.2.6., 7.4 and 7.5; only providing for an exemption if learning materials are produced by the member of staff for personal use and reference in teaching (produced outside normal course duties).

141 See for Germany Leuze (*supra* n 122) at 557 for Germany; for the U.S. Rooksby (*supra* n 112) at 203.

142 See for the U.S. Rooksby (*supra* n 112) at 203.

143 See Barendt (*supra* n 110) at 217; Leuze (*supra* n 122) at 557 for Fernuniversität Hagen; Rooksby (*supra* n 112) at 205 with further references.

144 See e.g. University of Reading (*supra* n 121) at para. 5.7.

145 See <https://ec.europa.eu/digital-single-market/en/modernisation-eu-copyright-rules>; for an overview Hilty, R./Moscon, V. (2017), Position Statement of the Max Planck Institute for Innovation and Competition on the Proposed Modernisation of European Copyright Rules, Part A – General Remarks, version 1.1.

146 For a discussion see Hilty, R./Richter, H. (2017), Position Statement of the Max Planck Institute for Innovation and Competition on the Proposed Modernisation of European Copyright Rules, Part B – Exceptions and Limitations (Art. 3 – Text and Data Mining), Max Planck Institute for Innovation and Competition Research Paper No. 17-02, 2017, for a recent comprehensive overview see Geiger, C./Frosio, G./Bulayenko, O. (2018), The Exception for Text and Data Mining (TDM) in the Proposed Directive on Copyright in the Digital Single Market – Legal Aspects, <http://www.europarl.europa.eu/RegData/etudes/IDAN/2018/604941/IPOL_IDA(2018)604941_EN.pdf>.

147 See also Recital 9 (2003/98/EC).

Archive's "Guidance on public task statements" clarifies that a public task relates to the core role and functions of a PSB.[148] As one criterion to determine whether PSBs produce information as part of the public task, the guidance mentions if the "creation and maintenance is funded through taxation rather than revenues or private investment".[149] It contrasts core responsibilities with "those of a more optional (and often commercial) nature".[150] One can see that according to this view, **PSBs have the discretion to determine where to draw the line**. In theory, even commercial activities could be explicitly designated as a public task – provided the transparency and review requirements according to Article 2(4) have been met.[151]

53 This functional conception of the PSI Directive, however, more or less breaks down in jurisdictions that regard state activities falling outside of the public task as illegitimate by definition.[152] According to this view, all legitimate activities of PSBs are performed in fulfillment of a public task. Then Article 1(2)(a) would be entirely irrelevant as well as Article 10(2). The only way to escape that dead end is to apply a functional reasoning according to which one has to recall that the main goal of the PSI Directive in 2003 was to **avoid cross-subsidies and its anti-competitive effects** on markets for value-added information services.[153] Therefore, it makes sense to exempt such information from the scope of the PSI Directive, where market forces have determined both its production and its distribution. According to this rationale, it seems reasonable to see a re-use of information if it has been *produced* on the basis of public funding and is then commercialized by the PSB in (potential) competition with private providers (see Article 10(2)). Recital 9 (2003/98/EC) supports that view as it refers to information produced and charged for exclusively on a commercial basis and in competition with others in the market.[154]

148 See National Archives, "Guidance on public task statements" of July 2015, 3.

149 See National Archives, "Guidance on public task statements" of July 2015, 3.

150 See National Archives, "Guidance on public task statements" of July 2015, 4.

151 See the problem in Office of Public Sector Information, Report on its investigation of a complaint, PinPoint Information Limited and The Coal Authority, December 2014, para. 23 et seq.

152 For Germany see Richter (*supra* n 1) at § 1 para. 260 et seq.

153 See Recital 9 (2003/98/EC).

154 See also Sec. 6(4) No. 1 ACT of 25 February 2016 on the re-use of public sector information, Poland, which refers to "which is not *produced* [emphasize added] by obliged entities as part of their public tasks defined by law", borrowing from Recital 9 (2003/98/EC) rather than from Article 1(2)(a), which refers to *supply* of information.

b.) Research and educational establishments

54 The practical relevance of Article 1(2)(a) for **R&E establishments** thus depends on the interpretation of the ambiguous legal standard. Following a literal interpretation, R&E establishments could escape the application of the PSI Directive if they define the public task narrowly. When following a more functional and competition related interpretation, information that has been produced on a commercial basis would fall under the exemption of Article 1(2)(a). This is often the case if the research establishment provides contract research to private parties under market conditions.[155] Such information produced by the PSB would not fall under the scope of the PSI Directive then. However, the situation might be different in cases of research collaboration with private partners. Due to the relevance of exclusive arrangements in the course of such collaboration, this topic is discussed below (*sub* III.4.).

5. Personal data

55 Article 1(4) clarifies that the PSI Directive "leaves intact and in no way affects the level of protection of individuals with regard to the processing of personal data" under EU and national data protection legislations. As a consequence, the PSI Directive is applicable to information that qualifies as personal data in general. However, **data protection law prevails** and normally sets significant limits to the re-use of personal data. Data protection – especially the GDPR and national data protection rules – is highly relevant for certain kinds of research data, such as contained in surveys or trials. The separation of regulatory layers is appreciable. However, the interface between PSI and personal data needs clarification, especially when it comes to balancing approaches. For good reasons, research establishments might be overly hesitant to make information re-usable if the interface is not clearly defined. Quite surprisingly, the interface between data protection regime and PSI is not (yet) really much discussed. While the European Data Protection Supervisor (EDPS) has provided some valuable guidance in 2013, an update is urgently needed.[156]

155 See National Archives, "Guidance on the implementation of the Re-use of Public Sector Information Regulations 2015 – For the cultural sector" of July 2015, 12.

156 See the Opinion of the European Data Protection Supervisor on the 'Open-Data Package' of the European Commission including a Proposal for a Directive amending Directive 2003/98/EC on re-use of public sector information (PSI), a Communication on Open Data and Commission Decision 2011/833/EU on the reuse of Commission documents of 18 April 2012; furthermore, van Eechoud (*supra* n 53) at 74 who elaborates on data protection complications; Richter (*supra* n 1) at § 1 para. 585 et seq.; see for a recent study

III. Consequences

1. Overview

56 What happens if R&E establishments and their information fall under the scope of the PSI Directive? If the general principle of Article 3 applies, R&E establishments are obliged to make the information re-usable. However, the PSI Directive allows for determining re-use **conditions** under certain circumstances. While Article 8 delineates the general leeway for such limitations, the PSI Directive also regulates conditions regarding formats (Article 5) and charging (Article 6) of re-use. Article 10 and 11 address competition concerns by setting out principles for non-discrimination and against exclusivity. Beyond that, the PSI Directive also contains procedural and transparency requirements.[157] Although these aspects might affect the PSBs practice to a certain extent and cause costs, they are not further discussed in the following.

2. General principle: Re-usability (Article 3)

57 Since its amendment in 2013, the **general principle of Article 3(1)** provides an obligation for PSBs to ensure that documents "shall be re-usable for commercial or non-commercial purposes" in accordance with the conditions as set out in the PSI Directive.[158] Should the information not be excluded from the PSI Directive, permitting re-use is mandatory. Member States have developed different ways to effectuate this obligation.[159]

58 Article 2(4) defines **'re-use'** as "the use by persons or legal entities of documents held by public sector bodies, for commercial or non-commercial purposes other than the initial purpose within the public task for which the documents were produced". The challenges for delineating the 'public task' have already been discussed above. But the definition contains some additional uncertainty, as it requires to determine what "for purposes ... *other* than the initial purpose" means. One could follow a substitutability test, asking if the concrete use of the information satisfies the same needs as the original purpose of the information's production has addressed. Should that be the case (e.g. if a private party wants to create a register that is identical with the public register), no re-use would be given. As a consequence, one would not qualify cases of (mere) imitation as re-use.[160] However, even in such cases, there are good reasons to argue for re-use, as – according to the rationale of the PSI Directive – a private party that is not designated to fulfill the public task must *per se* be understood as a re-user. Otherwise, re-usability would depend on a complicated assessment which Article 2(4) did not intended to stipulate.[161]

59 These different ways to interpret Article 2(4) might become relevant for **research institutions** if private re-users want to offer similar services, e.g. build up a parallel data repository. On the conceptual level, it seems favorable to also consider this as 're-use' because licensing conditions may account for preventing unfair practices. Also, a transfer of research results to other research establishments can be considered as re-use. However, Article 2(4) sets out that this is not the case if an exchange of information between PSBs is purely in pursuit of their public task. Again, the interpretation depends on how the concept of 'public task' is understood. There are good reasons to acknowledge re-use at least in those cases in which research establishments commercialize research data supplied by other PSBs.[162]

3. Conditions for re-use

a.) Restrictions

60 As a **general rule**, Article 3(1) PSI Directive obliges the Member States to allow for re-usability of information without any restrictions. The wording of the provision makes clear that this applies to commercial and non-commercial purposes. As a consequence, PSBs cannot allow only non-commercial re-use while prohibiting commercial re-use. The only thing they can do is to differentiate conditions for such re-use categories (see Article 10). Should a PSB hold IPRs, it has to license for re-use purposes if requested. The PSI-regime can be read as a duty to license in this respect.

61 However, there are **exceptions** to the rule of unrestricted re-usability, which can fall into two categories. In the first category, the PSB itself *must* restrict re-use. This is the case if data protection law does not prevent access for everyone, but restricts

Wiebe/Dietrich (*supra* n 41).

157 See Articles 4, 7, 9.

158 See Recital 8 (2013/37/EU) as opposed to Recital 9 (2003/98/EC).

159 Either by administrative procedure or by permission *de lege*.

160 See the problem in Office of Public Sector Information (OPSI), Report on its investigation of a complaint, PinPoint Information Limited and The Coal Authority, December 2014, para. 27 et seq.

161 See Richter (*supra* n 1) at § 2 para. 113, 114, 119.

162 Unless Article 1(2)(a) applies, see rationale according to Recital 9 (2003/98/EC).

re-use.[163] The PSB is also obliged to limit re-use due to contractual (and not IPR-related) obligations towards third parties. This can be very relevant for the re-use of mere datasets that usually lack IP-protection.[164] Furthermore, if one interprets the IPR-exemption narrowly, the sub-licensing PSB has to obey restrictions stemming from its own license with the licensor. In the second category, the PSB *can* restrict re-use and set conditions, either by contractual terms or by license according to Article 8.

b.) Licensing (Article 8)

62 Article 8 is the central provision that allows the PSB for setting **re-use conditions** and defines possibilities and limits. Article 8 – unlike its title might suggest – applies to information no matter if protected by IPRs or not. However, a distinction must be made due to the practical relevance of licensing IPRs for re-use under Article 8. In general, differential licensing is possible.[165] However, discrimination needs to be justified according to Article 10. It is common to differentiate conditions between commercial and non-commercial users.

63 If the information is **protected by IPRs** (of the PSB), consent is required for re-use. In general, Article 3(1) obliges the PSB to give consent and – as a consequence – to license re-use. As has been shown, unless the PSB itself is restricted when it comes to sub-licensing, it must allow for non-commercial as well as for commercial re-use.[166] As a consequence, the PSB cannot reserve commercialization of the information for itself. When it comes to other licensing restrictions, Article 8(1) gives some discretion by stipulating that imposed conditions "shall not unnecessarily restrict possibilities for re-use". Recital 26 (2013/37/EU)) specifies that licenses should "in any event place few restrictions on re-use as possible". Therefore, restrictions shall be the exception.

64 If the information is **not protected by IPRs**, Article 8(1) allows for setting terms and conditions and also applies to individual contracts that concern the respective information. This is of particular relevance for setting terms and conditions for the re-use of datasets that do not qualify for IP-protection. One could argue that PSBs have even less discretion as compared to situations in which they hold IPRs, because there are good reasons for why the legal regime does not protect the particular information as IP.

65 The PSI Directive encourages the use of **standard licenses** and the Commission's PSI-notice gives more guidance in detail.[167] OA-policies usually make use of standard licenses, but which licenses are used in particular differs widely.[168] As re-usability is a general claim of OA, a significant number of OA-licenses would match the requirements of Article 8(1). Non-commercial-clauses are, however, not consistent with Article 3(1) and would normally be invalid. Furthermore, especially non-derivative and share-alike clauses can be seen as an obstacle for re-use. They have been identified as a source for potential incompatibilities between scientific projects.[169] In contrast, attribution clauses are a well-established practice regarding OA[170] and not of any concern (see Recital 26 (2013/37/EU)).

66 At the end of the day, Article 8 requires PSBs to **justify re-use conditions** and brings their justifications under scrutiny of the courts or other impartial bodies[171]. Whether such a particular restriction is justified needs to be decided on a case-by-case basis. Should the PSI Directive accommodate R&E institutions, a clarification in the recitals could give guidance by referring to the specific features and practice of licensing from R&E institutions in general and OA-practices in particular.

c.) Formats (Article 5)

67 Article 5 PSI Directive obliges PSBs to provide information in all pre-existing **formats** and – proportionality provided – even to create or adapt

163 In detail Richter (*supra* n 1) at § 1 para. 585 et seq.

164 For that reason, universities implement different "ownership" policies on a contractual basis.

165 See Recital 19 (2003/98/EC).

166 Therefore, e.g. CC-NC licenses would be void, see in detail van Eechoud, M. (2011), Friends or Foes? Creative Commons, Freedom of Information Law and the European Union Framework for Reuse of Public Sector Information, in: Guibault, L./Angelopoulos, C. (eds.), Open Content Licensing – From Theory to Practice, 199; Wiebe/Ahnefeld (*supra* n 91) at 207; Richter (*supra* n 1) at § 4 para. 82 et seq.

167 See Article 8(2) and Commission "Guidelines on recommended standard licences, datasets and charging for the reuse of documents" (2014/C 240/01); when it comes to differential licensing, it can be problematic to use standard licenses, see Papadopoulos/Bratsas (*supra* n 10) at 26 et seq.

168 See e.g. in particular for datasets: Open Data commons: <https://opendatacommons.org/licenses/pddl/1.0/>.

169 Guibault, L. (2013), Licensing Research Data under Open Access Conditions, in: Beldiman, D. (ed.), Information and Knowledge: 21st Century Challenges in Intellectual Property and Knowledge Governance, 63, 73, stating that the "Berlin declaration remains vague regarding the freedom to make changes and improvements and to distribute derivative works".

170 See e.g. the Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities.

171 See regarding the German jurisprudence on particular re-use conditions Richter (*supra* n 1) at § 4 para. 86 et seq.

information accordingly. This clause becomes relevant in cases where the PSB hold a desired format, but only another format has been made accessible.[172] According to the re-use friendly interpretation of the German Federal Administrative Court, Article 5 obliges the PSB to provide this non-accessible but existing format, even if there is no individual right to access to it.[173] Limits are set for reasons of data protection.[174] Also, this case must not be confused with situations in which one desires e.g. access to the underlying raw data of an accessible complied dataset. In this case, the contained information is not equivalent and Article 5 PSI Directive does not apply.

d.) Charging (Article 6)

68 Charging provisions were the most contested issue in the frame of the PSI Directive's amendment in 2013.[175] Article 6(1) sets **marginal cost pricing** as the default model for charging re-use. This does not affect charging for access, which lies beyond the scope of the PSI Directive. However, exceptions in Article 6(2) allow PSBs for charging above marginal costs if revenues based on charges cover a substantial part of the costs relating to the public task of the PSB or to the collection, production, reproduction and dissemination of the concerned information. These exceptions seek to not hinder the normal operations of PSBs[176] and they provide for discretion to determine their financing mix. In practice, it remains to be seen whether the exception is in fact the rule.[177]

69 For this reason, one cannot say with certainty if charging provisions of the PSI Directive would affect pricing policies of **R&E establishments**. At least OA-policies can be held to be consistent with Article 6, as they call for a free dissemination of information by definition. Repositories use standard licenses and their financing is based on funds of the institution and/or by the authors who make their research available there.

e.) Non-discrimination (Article 10)

70 Article 10(1) lays down a **non-discrimination principle** and requires that re-use conditions must be non-discriminatory for comparable categories. It allows PSBs to differentiate conditions, however, it does not allow for a mere approval of non-commercial re-use while prohibiting commercial re-use. As a specific non-discrimination rule, Article 10(2) concerns the case in which a PSB generates information within its public task but then uses this information as an input for commercial activities that fall outside its public tasks. In that case the PSB is obliged to apply the same conditions to the supply of information to third parties. This provision is rooted in competition reasoning. It sets a level playing field by preventing anti-competitive effects of cross-subsidization on the markets for value-added products or services. The application of this provision, however, requires the PSB to clearly distinguish between public task and its own commercial re-use.[178] The respective problems have already been discussed above.

71 When observing the relevance of this standard for **R&E establishments**, one has to keep in mind the context of the PSI Directive's creation in 2003: the Internet was developing, Google's search engine was in a premature phase, smartphones and Facebook did not exist yet, and systematic digitization efforts of cultural institutions were only beginning. Classical cases mainly concerned weather-, hydrological-, geo- and legal information[179] as well as public registers, since a lot of PSBs were about to implement online accessibility of that information. Nowadays the situation is different. Massive digitization and interconnection have enabled PSBs themselves to implement entirely new "business models". Distance learning is a good example for that. Should it be offered on a commercial basis and fall outside the public task, educational institutions might have to provide the basic material for similar conditions to third parties – provided its initial creation falls within its public task and the application of the PSI Directive is not exempt for other reasons according to Article 1(2).

4. Prohibition of exclusive arrangements (Article 11)

72 Due to its **competition rationale**, the non-discrimination principle is closely related to Article 11, which requires the PSB to justify exclusive arrangements on re-use. Such arrangements would

172 See e.g. decision notice of the ICO U.K. of 4 April 2017 – Cambridgeshire County Council – FS50619465.

173 See German Federal Administrative Court of 14 April 2016 – BVerwG 7 C 12.14 (ECLI:DE:BVerwG:2016:140416U7C12.14.0).

174 E.g. if information is available on a single request base, but not as bulk export.

175 See Beyer-Katzenberger, M. (2014), Rechtsfragen des „Open Government Data", Die Öffentliche Verwaltung 144, 150.

176 See Recital 22 (2013/37/EU).

177 For the application see Commission "Guidelines on recommended standard licences, datasets and charging for the reuse of documents" (2014/C 240/01); see also EFTA-court of 16 December 2013, Case E-7/13 – Creditinfo Lánstraust hf.

178 See Drexl (*supra* n 30) at 75.

179 See for use cases e.g. de Vries, M. et al. (2011), Pricing Of Public Sector Information Study (POPSIS).

prevent the PSB from fulfilling its obligation of non-discriminatory treatment. Therefore, Article 11 states that exclusive agreements between PSBs and private partners concerning the re-use of information should be avoided as far as possible unless the exclusivity is not necessary for the provision of a service in the public interest (Article 11(2)).[180]

73 **Research collaboration** between public research establishments and private partners (e.g. the industry) is widespread and prone to exclusive arrangements. Collaborations have to be distinguished from contract research, where research is provided as a service on a commercial basis under market conditions primarily in order to generate a financial return.[181] Such contract research would most likely be exempt from the PSI Directive according to Article 1(2)(a). Third-party funding for collaborations significantly contributes to the budget of many public research establishments. The performance of such collaborations is widely regarded as falling within the scope of a university's public task.[182] Usually the underlying agreements regulate ownership and exploitation of results (and therefore also information) stemming from the collaboration. While this predominantly concerns inventions and therefore patents (which are not covered by the PSI Directive), provisions of the agreements can also concern copyright and related rights and research data in general.

74 Collaboration agreements have to be carefully analyzed when determining whether information falls under the **scope of the PSI Directive**. In general, the allocation of rights follows the parties' respective contributions to the project.[183] Should the agreement allocate IPRs to the industry collaborator, a 'third party' holds IPRs according to Article 1(2)(b) and the PSI Directive is not applicable. Should the PSB retain all rights, the PSI Directive can be applicable in general. However, the collaboration contract might contain provisions that exclusively reserve re-use of this information for the industry partner (i.e. commercial exploitation of value-added products on the basis of datasets generated in the course of the collaboration). While one might intuitively hold Article 11(1) applicable in that case, it must not be forgotten that the PSI Directive only applies if the information that has been generated in collaboration and is held by the PSB[184] is accessible without restrictions. This requires either that there is a right to unrestricted access, or that the PSB made the information accessible. As has been shown above, FOI-regulations provide for many exemptions that can prevent access in such cases. At the same time, collaboration contracts might often be used to derogate from public disclosure.[185] It can be seen that the PSI Directive does not tackle the quite common situation where non-accessible information is generated in public-private collaboration and its re-use is reserved to the private collaborator only.

75 However, the collaborator might allow for publication of the concerned information by the research establishment.[186] Should the research establishment make the information accessible,[187] Article 11(1) would be applicable if there was an exclusive license for its commercial re-use as set out in the collaboration agreement.[188] The exclusivity then has to meet the justification standard of Article 11(2) for not being rendered void. As a reaction, the collaboration partners can avoid the application of the PSI Directive and Article 11(2) by assigning initial ownership to the collaborator who then has to license back non-exclusive rights to the research institution. In that case, the PSI Directive will be not applicable according to Article 1(2)(b). One can see that – without the introduction of safeguards – an application of the PSI Directive **might significantly shift the incentive curve for the terms of collaboration** and might also diminish the general accessibility of research results.

IV. Analysis

76 The effects of removing the exemption for R&E establishments are not entirely clear. As can be seen in general, the exemptions are vague and they can be interpreted and applied narrowly or broadly. Practice differs between the Member States. Removing the exemption for R&E establishments cannot be done without carefully **assessing and clarifying**, what interpretation should set the minimum standard for re-use. Concerning the IPR exemption of Article 1(2) and copyright in particular, it must urgently be clarified how exclusive licenses are to be treated.

180 See also Recital 20 (2003/98/EC).

181 See U.K. University and business collaboration agreements: model agreement guidance of 6 October 2016, 4.6.

182 See U.K. University and business collaboration agreements: model agreement guidance of 6 October 2016, under 6., qualifying that as "charitable research"; see also §§ 3(3), 71 Gesetz über die Hochschulen des Landes Nordrhein-Westfalen.

183 See U.K. University and business collaboration agreements: model agreement guidance of 6 October 2016, 3.47.

184 Also meaning: not held by the private collaborator and not referring to the value added, information-based product.

185 See also U.K. University and business collaboration agreements: model agreement guidance of 6 October 2016, under 3.68, for different ways to achieve that.

186 See U.K. University and business collaboration agreements: model agreement guidance of 6 October 2016, 3.62.

187 E.g. by publication in a university repository.

188 This was the problem with digitization partnerships – the whole purpose of digitization is to enable accessibility of the information.

Clarification would have general impact as it would affect all authors' rights systems, but also common law copyright systems allow for granting licenses. Furthermore, the debatable[189] *sui generis* protection of databases mostly raises factual uncertainty about whether a particular database enjoys protection. This is a problem that the DB Directive itself needs to address. In contrast, the different perceptions of the concept of 'public task' can only be harmonized to a certain extent. The actual underlying question is how much space for exclusive commercialization shall be left to the R&E establishments.

77 Even if one follows the narrowest reading of the exemptions and therefore the re-use-friendliest view, a large amount of information held by R&E establishments will effectively not fall under the **PSI Directive's scope**. This predominantly affects scientific publications (works) and teaching materials to a considerable extent. Also, information that has been produced as contract research for private parties under market conditions falls outside of the PSI Directive's scope. In contrast, many datasets and databases fall under the scope of the PSI Directive, at least if they are accessible without restrictions. University repositories play a significant role here. However, data protection has an important function and further narrows down the application of the PSI Directive.

78 Besides the general **benefits** of having as few exemptions as possible, a complete deletion of the R&E exemption would reduce the costs of delineating the PSI Directive's scope. There is no need to define R&E establishments anymore, which is especially relevant for cross-purpose organizations.[190] Furthermore, there seems to be no reason why re-use of administrative information of R&E establishments should be treated differently from the information held by other PSBs.

79 However, potential **costs of the deletion** might arise from dysfunctionalities concerning operations directly related to R&E due to the seminal role for the knowledge society and economic growth. One has to keep in mind that the creation of information and dissemination of knowledge and information is the main goal of R&E establishments and not a mere by-product. A deletion might interfere with well-established systems of knowledge creation and dissemination. Furthermore, IPRs have been identified as crucial and their allocation determines whether the PSI Directive is applicable or not. Institutions therefore have to constantly screen and monitor the IPR status of the information they hold.[191] This can be costly. One has to acknowledge, however, that this is not a new challenge for R&E establishments – OA-repositories have already found ways to address ownership disclosure and to formulate re-use conditions. Furthermore, in the framework of research collaborations, IP ownership is a central point, even though practice is much more advanced when it comes to inventions and patents as compared to mere datasets.

80 There is further **uncertainty** about the effect. R&E establishments themselves have discretion to submit information to the application of the PSI Directive for three reasons: first, Member States can still design access regimes and R&E establishments can still decide what information they choose to make accessible without restrictions. Legislative and institutional OA-policies can be authoritative in this respect. Second, it lies in the discretion of the establishments to determine copyright policies. Third, they can also determine their public task autonomously – provided this is done in a transparent way and subject to review. Depending on one's standpoint, this can be seen as favorable or problematic. On the one hand, it can be argued that it leaves enough autonomy and flexibility to the R&E establishments to respond to particular organizational needs, including those relating to the specific features of the information they hold. On the other hand, this flexibility can effectively water down the minimum standard for the re-use of R&E information in the internal market and a lot of valuable information would not be affected. It also increases the risk for activities or rules designed to circumvent the application of the PSI Directive and to uphold barriers to competition.

81 In **conclusion**, while empirical research is urgently needed for finding a prudent regulatory approach, it does not seem too far-fetched to delete the R&E exemption from the PSI Directive. As can be shown, some fears are not justified, however, other potential problems are highlighted. An inclusion of R&E establishments can also be understood as a chance to eliminate ambiguities of the PSI Directive, which can be beneficial to re-use in all other fields. However, cases have been identified where broadening the scope of the PSI Directive might result in less openness (contrary to the ambition of OA-policies) and change collaboration incentives. One can think of accounting for that by providing clarification in the recitals that would accompany a deletion of the R&E exemption. Should this not

189 See already Kur, A. (2006), Erste Evaluierung der Richtlinie 96/9/EG über den rechtlichen Schutz von Datenbanken – Stellungnahme des Max-Planck-Instituts für Geistiges Eigentum, Wettbewerbs- und Steuerrecht, 55 GRUR Int. 725.

190 There has been some confusion in cases where research information is held by non-research organizations (e.g. weather services); see also National Archives, "Guidance on the implementation of the Re-use of Public Sector Information Regulations 2015 – For the cultural sector" of July 2015, 12.

191 See Jančič/Pusser/Sappa/Torremans (*supra* n 6) at 359.

be sufficient, modification of the substantial rules of the PSI Directive tailored to R&E establishments might be a solution.

D. Modification of the research and educational exemption

I. PSI principles for cultural PSBs as model?

1. Legal standard

82 In 2013, the amendment of the PSI Directive included libraries (including university libraries), museums and archives (cultural PSBs) in its scope.[192] The inclusion was based on a careful assessment[193] and was politically highly sensitive.[194] The deletion of the original exemption for these institutions[195] was justified with the advancement of digitization and rights clearance.[196] The PSI Directive contains **modified provisions** for the re-use of these institutions' information that account for the special features of cultural PSB and their information. One can easily see a general parallel, as information of R&E establishments have their specific features and their re-use policies are not less politically sensitive.

83 The **general principle** for re-use of information of cultural PSBs is set out in Article 3(2). It affects only such information in which cultural PSBs hold IPRs.[197] Given that that is the case, Article 3(2) – as opposed to Article 3(1) – contains a mere expectation to allow re-use but no enforceable obligation. Therefore, initially it lies in the hands of the cultural PSB whether to submit itself to the re-use regime. However, once the PSB allows re-use of the information, it is obliged to make it available for others to re-use. Therefore, the crucial question is whether the PSB has in fact allowed **re-use**. The different ways to interpret re-use as defined in Article 2(4) can lead to different outcomes. Furthermore, Article 3(2) also covers the re-use of the cultural PSB itself. As stated above, the rationale of Article 3(2) either follows the binary logic 'within' vs. 'outside' the public task can be applied[198] or one might follow a more functional, competition-related interpretation. Under the strictest view, the cultural PSB would submit itself to the PSI regime simply by commercializing information.

84 If the cultural PSB has allowed re-use, it must allow re-use for everyone in accordance with the PSI Directive. This includes applying non-discriminatory terms (see Article 10 and Article 8). Conditions may vary for different types of re-use, but not among different types of re-users.[199] One question without a definite answer is whether cultural PSBs have the discretion to allow **only non-commercial** while prohibiting commercial re-use. This view seems to be predominant in some Member States[200] and it can be supported by the thought that allowing only for non-commercial re-use is better than no re-use at all. On the contrary, Article 2(4) clearly defines re-use as commercial or non-commercial use.[201] Should one follow the very re-use friendly view that commercial use by the PSB qualifies as re-use and the PSB then also has the obligation to license for commercial-use, it is likely that some cultural PSBs need to adapt their "business models".

85 The PSI Directive contains specific **provisions privileging cultural PSBs** in their re-use policies.[202] Article 6 allows for charging above marginal costs, including a reasonable return on investment. The provision seeks to not hinder their normal running, as cultural PSBs – rather than other PSBs – systematically rely on revenue-based income streams.[203] Furthermore, Article 11 sets out special rules for exclusive arrangements regarding the digitization of cultural resources. This accounts for the wide-spread public-private-partnerships, which

192 See for the course of the amendment Guibault/Salamanca (*supra* n 41) at 220 et seq.

193 See Jančič/Pusser/Sappa/Torremans (*supra* n 6) at 361.

194 See for a detailed history Wirtz (*supra* n 30).

195 See already discussion in COM(2002) 207 final, 9.

196 See SEC(2011) 1152 final, 34.

197 Also, there must be IP-protection, otherwise Article 3(1) applies. However, there is the exception of Recital 9 (2013/37/EU).

198 See example in National Archives, "Guidance on the implementation of the Re-use of Public Sector Information Regulations 2015 – For the cultural sector" of July 2015, 10: Transfer of information from a museum to its commercial trading arm is to be considered as re-use.

199 See National Archives, "Guidance on the implementation of the Re-use of Public Sector Information Regulations 2015 – For the cultural sector" of July 2015, 22.

200 National Archives, "Guidance on the implementation of the Re-use of Public Sector Information Regulations 2015 – For the cultural sector" of July 2015, 10, argue that re-use request can be declined or it can be allowed for restricted uses such as non-commercial research re-use, but be declined for commercial re-use.

201 See Richter (*supra* n 1) at § 2 para. 118 et seq.: Re-use clearly addresses commercial *and* non-commercial. See for a similar reasoning also the wording of the Cyprus Act 205(I)/2015, Sec. 4(2): "Public sector bodies shall ensure that, where the re-use of documents, information and data for which libraries, including university libraries, museums and archives, hold intellectual property rights, is allowed, these documents, information and data shall be re-usable for commercial or non-commercial purposes".

202 Overview in Richter (*supra* n 1) at § 1 para. 553 et seq.

203 See Recital 23 (2013/37/EU); Jančič/Pusser/Sappa/Torremans (*supra* n 6) at 367 et seq.

can facilitate the use of cultural collections.[204]

2. Suitability for research and educational establishments

86 As can be seen, the general principle and the modifying provisions account for both political consensus and special features of cultural PSBs. In both cases, the organization's main task is related to the production, storage or **dissemination of knowledge**.[205] Information is not a mere by-product of the activities, but the establishment's main purpose centers on them. This explains why cultural heritage organizations are also among the signatories of the Berlin Declaration on OA. Also, IP-protection is relevant for R&E establishments. This corresponds with the general principle for cultural PSBs, which explicitly requires IP protection and would therefore not affect a considerable amount of research data.

87 However, there are also **differences** between cultural PSBs on the one hand and R&E establishments on the other. While digitization might be relevant to some extent, university libraries are already included within the scope of the PSI Directive. In comparison, charging for copyright protected information does not seem to be such a predominant problem for research institutions. Rather, income streams originating from patents are highly relevant, but they fall outside the scope of the PSI Directive. However, public educational establishments – depending on their financing structure – might largely depend on revenue based income streams stemming from the commercialization of information.

88 In **conclusion**, all of these possible aspects, however, need to be carefully assessed when looking for prudent regulatory approaches. What has been shown is that even the interpretation and application of the general principle for cultural PSBs is not entirely clear, especially when it comes to the possibilities and limits to reserve commercialization. Clarification is urgently needed, should the rules for cultural PSBs be used as a model for re-use rules governing R&E establishments.

II. Alternative modifications and limits

89 Whether or not alternative or additional modifications of the exemption are desirable depends on the specific needs and effects. Regarding the **scope** of the PSI Directive, one might even see the application of the PSI Directive to research establishments as less critical compared to educational establishments, because the fundamental right of scientific freedom effectively prevents a considerable amount of research information from the application of the PSI Directive. This effect might be mitigated in the case of educational establishments, though it is sometimes difficult to draw the line (e.g. universities). If this corresponds with a market driven development of education (e.g. universities develop commercial strategies, also for distance learning), one has good reasons to argue that this field should be left entirely to the competition in the market – at least from a competition point of view.[206] One could also take the function and use of the information as decisive criterion (as opposed to the nature of the establishment). However, definitions can be difficult as has been shown for the term 'research data'. Moreover, distinguishing between different sorts of information creates some costs of delineation and legal uncertainty. The classification of funding agreements (whether closer to administration or research) has illustrated that.

90 One could also think about modifying the legal **consequences** of the Directive's application. As exclusivity seems to enable research collaboration with third parties, the need for modifying the standard for exclusive agreements has to be considered in order to prevent cooperation incentives changing in an unfavorable way.[207] In contrast, it does not seem advisable to extend the principle for cultural PSBs according to Article 3(2) to R&E to information that is not protected by IPRs. While one might think that this could foster re-use of non-protected datasets, this would bring back the situation prior to 2013, which has rightly been criticized as creating an 'illusionary property right' for PSBs.[208]

91 Regarding the scope, one could also think about extending it to **public funders** and providing for specific rules that would oblige them to implement OA-mandates in their grants.[209] However, these specific provisions address accessibility and should be an instrument of sector specific regulation.

92 This reminds one of the fact that the current PSI Directive does not regulate access, but requires it

204 See Recital 30 (2013/37/EU).

205 See Jančič/Pusser/Sappa/Torremans (*supra* n 6) at 356.

206 See Drexl (*supra* n 30) at 83.

207 See e.g. the provisions on digitization (Article 11(2a)).

208 See De Filippi, P./Maurel, L. (2015), The paradoxes of open data and how to get rid of it? Analysing the interplay between open data and sui-generis rights on databases, 23 International Journal of Law and Information Technology 1.

209 See e.g. the H2020 Programme (2017), Guidelines to the Rules on Open Access to Scientific Publications and Open Access to Research Data in Horizon 2020, Version 3.2 of 21 March 2017.

for being applicable. **Access regimes** can streamline accessibility and the main challenge is to find a fair and legitimate standard that balances out interests accordingly. It seems likely that extending the PSI Directive to R&E establishments stimulates the general debate on open science. Access regimes can also include re-use rules. Therefore, including research establishments in the PSI Directive would not preclude more re-use friendly, sector specific regimes. One has to make sure that if an access regime provides for a more re-use friendly standard than the PSI Directive, the re-use friendlier regime would prevail. This collision problem has been significantly discussed and has yet to be entirely solved with regards to the relationship between the PSI Directive and the INSPIRE Directive.[210]

E. Conclusion

93 The preceding analysis brought together the discussions about open research data, open education, and PSI. Their common driving force is the call for a widespread dissemination of publicly funded information. While the OA-debate and common regulatory approaches on research information are well developed, regulatory approaches and markets for educational information seem heterogeneous, premature, and quite dynamic at that stage. Therefore, the analysis focused on research information rather than on educational information.

94 In principle, the OA-debate and the PSI Directive follow similar rationales. Thus, it does not come as a surprise that several connections occur. However, due to some general legal uncertainty about the PSI Directive's standard, it is difficult to derive robust assumptions that can form a basis for predicting the effect of including R&E establishments. Without any doubt, the Directive's exemptions will filter out a lot of information held by R&E establishments, especially information protected by IPRs from third parties. How much information is affected in total depends on the interpretation of the exemptions under Article 1(2); namely, accessibility, IP-protection, and the public task. Moreover, it must be understood that R&E establishments have considerable discretion to "opt-in" the application of the PSI-Directive by making information accessible, designing IP-arrangements, and re-defining their public task. This can be regarded as a positive element providing flexibility to reconcile the different needs and traditions of Member States and PSBs. But it is also critical because the application of re-use rules can be circumvented. In any case, clarification of the standard for all of the three exemptions is urgently needed.

95 When discussing the legal consequences of an application of the PSI Directive, the effect on exclusive arrangements seems of particular importance and requires cautious consideration. The PSI Directive might affect several public-private research collaborations. This issue must be addressed to prevent an unintended, significant change of collaboration incentives and terms. There are situations in which one might even end up with less re-use than before. In general, the PSI Directive could address specific features of R&E as it has also been done with cultural PSBs. Should this approach be followed, it definitely needs clarification whether R&E establishments could still reserve commercial re-use of the information for themselves while allowing non-commercial re-use to others.

96 Finally, one has to be reminded of what makes the PSI debate about R&E establishments unique and challenging. The common rationale of OA-initiatives and PSI lies in the claim that what is financed with taxpayers' money should "belong" to everyone. However, there is a seminal difference: unlike in any other PSB that is covered by the PSI-Directive, the employee himself (meaning the researcher and not the institution) decides to a considerable extent what and how information is supplied. Therefore, the researcher's personal incentives and the informal norms of research communities rather than conventional market mechanisms drive the creation and dissemination of information and knowledge. The PSI-Directive should not change these basic rules of the game. One can be optimistic that this will not happen if the crucial aspects mentioned are taken into account, discussed, and tested before PSI regulation might be revised.

210 See Richter (*supra* n 1) at § 1 para. 559 et seq.

"This Video is Unavailable"

Analyzing Copyright Takedown of User-Generated Content on YouTube

by **Kristofer Erickson and Martin Kretschmer***

Abstract: What factors lead a copyright owner to request removal of potentially infringing user-generated content? So-called "notice-and-takedown" measures are provided in the United States under Section 512 of the U.S. Copyright Act (as amended by the Digital Millennium Copyright Act 1998) and enabled in the European Union under the Directive on Electronic Commerce (2000/31/EC). While the combination of limiting liability ("safe harbor") and notice-and-takedown procedures was originally conceived as a means of balancing innovation with the interests of rightholders, there has been limited empirical study regarding their effects. This research investigates, for the first time, the factors that motivate takedown of user-generated content by copyright owners. We study takedowns within an original dataset of 1,839 YouTube music video parodies observed between January 2012 and December 2016. We find an overall rate of takedowns within the sample of 32.9% across the 4-year period. We use a Cox proportional hazards model to investigate propositions from rightholder groups about the factors that motivate takedowns: these include concerns about commercial substitution; artistic/moral concerns; cultural differences between firms; and YouTube uploader practices. The main finding is that policy concerns frequently raised by rightholders are not associated with statistically significant patterns of action. For example, the potential for reputational harm from parodic use does not appear to predict takedown behavior. Nor does commercial popularity of the original music track trigger a systematic response from rightholders. Instead, music genre and production values emerge as significant factors. We suggest that evolving policy on intermediary liability - for example with respect to imposing filtering systems (automatically ensuring "stay-down" of potentially infringing content) - should be carefully evaluated against evidence of actual behavior, which this study shows may differ materially from stated policy positions.

Keywords: Copyright; parody; notice-and-takedown; fair use; YouTube; music industry

© 2018 Kristofer Erickson and Martin Kretschmer

Everybody may disseminate this article by electronic means and make it available for download under the terms and conditions of the Digital Peer Publishing Licence (DPPL). A copy of the license text may be obtained at http://nbn-resolving.de/urn:nbn:de:0009-dppl-v3-en8.

Recommended citation: Kristofer Erickson and Martin Kretschmer, "This Video is Unavailable": Analyzing Copyright Takedown of User-Generated Content on YouTube, 9 (2018) JIPITEC 75 para 1.

A. Introduction

1 When Internet users combine, remix, mash-up or parody existing cultural materials, they may infringe the copyright of the owners in the original work. Two sets of legal norms interact in determining the availability of such user-generated content on Internet platforms.[1] The first set are based on international agreements, which define the exclusive rights under copyright law and restrict possible exceptions that may permit derivative re-use.[2] The second set consist of rules about the liability of intermediaries on whose services such materials may be communicated. While the latter rules vary by jurisdiction (and can be copyright-specific, or applicable to issues such as terrorism, hate speech, or sexual abuse), in practice the great majority of global requests for removing infringing content are based on the formal notice-and-takedown regime established by the United States Digital Millennium Copyright Act (DMCA 1998).[3]

2 The owner of a copyright work may tolerate a derivative use or may act to remove infringing content, bearing resource costs associated with issuing a notice. Due to the complexity of the media ecosystem in which user-generated content is produced, rightholders are faced with a difficult decision about whether and when to act. Particularly in the case of owners of large catalogues of material (such as major record labels), the cost of policing and requesting removal of infringing content may exceed the benefits of doing so. Rightholders must decide which content they will expend resources protecting, and which types of potential infringement they should most aggressively pursue. For example, should mash-ups or parodies be approached in the same way as incidents of outright piracy? If not, where do copyright owners draw the line and what factors in particular trigger a removal request?

3 In 2012, some rightholders were opposed to a UK Government proposal to introduce a new copyright exception for the purposes of parody, caricature and pastiche. They argued that such an exception would potentially cause substitution, deprive them of licensing revenue and damage the artistic integrity of works. Here, we analyze the pattern of takedowns over a 4-year period, to test whether rightholders act(ed) in ways consistent with policy statements. Do economic or moral rights concerns guide rightholder takedown behavior? And what changes, if any, arise from the introduction of a new copyright exception?[4]

4 Our analysis of rightholder behavior complements and offers a new perspective on recent empirical work assessing the appropriateness of notice-and-takedown procedures as a means of balancing the interests of rightholders, innovative services and citizens.[5] We find that our efforts to discern

* Kristofer Erickson is Associate Professor of Media and Communication, University of Leeds.

Martin Kretschmer is Professor of Intellectual Property Law and Director of CREATe (RCUK Copyright Centre), University of Glasgow.

1 Definitional note on "Internet platforms": The safe harbor for internet intermediaries is defined in the United States under Section 512 of the U.S. Copyright Act (as amended by the Digital Millennium Copyright Act – DMCA 1998) for "Online Service Providers" and in the European Union under the E-Commerce Directive for "Information Society Services". Both legislations were conceived in a pre-social media world where the Internet Service Provider (ISP) was the technological orientation point. In recent regulatory efforts, the European Commission has used the term "online platforms" (Commission Communication: Stepping up the EU's efforts to tackle illegal content online, MEMO-17-3522, Brussels, 28 September 2017). Jurisprudence has found it easier to develop the wider concept of internet intermediaries in the context of Article 11 IPRED (IPR Enforcement Directive 2004/48/EC) and Article 8(3) InfoSoc Directive (2001/29/EC). The Court of Justice of the European Union (CJEU) defines an intermediary indistinctly for online and offline contexts: "for an economic operator to fall within the classification of 'intermediary' [...], it must be established that it provides a service capable of being used by one or more persons in order to infringe one or more intellectual property rights, but it is not necessary that it maintains a specific relationship with that or those persons" – see *Tommy Hilfiger* (C-494/15, at 23) and *UPC Telekabel* (C-414/12, at 32 and 35).

2 According to Art. 9(2), Berne Convention for the Protection of Literary and Artistic Works 1886, Art. 9(2) exceptions to the exclusive rights in national laws are required to be specific, non-prejudicial to the author, and not in conflict with normal exploitation (the so-called "three-step-test"). The latest version of the Berne Convention is the Paris Act 1971, as amended in 1979. All EU countries are members, and the US acceded to Berne in 1989. In 1994, the Berne Convention (with the exception of Art. 6bis on "moral rights") was incorporated into the WTO Agreement on Trade-related Aspects of Intellectual Property Rights (TRIPS). The World Trade Organization (WTO) currently has 164 member countries (as of 29 July 2016), making Berne copyright norms binding on most of the world.

3 According to Google's transparency report, Google has received in total more than 3bn copyright takedown requests (available at: <https://transparencyreport.google.com/copyright/overview>, last accessed 20 October 2017). Personal communication from a senior counsel of Google indicated that 99% where submitted as a request using the DMCA formalities. This was regardless of whether the country in which the request was filed prescribed these formalities or had any safe harbor laws. See the Canadian case of *Google Inc. v. Equustek Solutions Inc.* (2017 SCC 34) for forensic details of Google's takedown procedures.

4 Digital literacy is frequently characterised as a requirement for successful engagement in 21st century political life. See W.L. Bennett, Changing citizenship in the digital age, in *Civic life online: Learning how digital media can engage youth* (MIT University Press 2008), pp. 1-24.

5 Urban, Karaganis and Schofield found that automated takedown systems leave little room for human review, with nearly 30% of a randomized sample of 1,826 takedown requests during a six months period in 2013 assessed as being of questionable validity. J. Urban, J. Karaganis, B. Schofield.

rightholder behavior are complicated by the existence of automated and opaque systems for detection and removal of content. This makes it difficult to study and evaluate takedown behavior.

5 Our empirical approach consists of a longitudinal cohort analysis of 1,839 user-generated music video parodies hosted on video sharing platform YouTube. The data were initially collected in January 2012 as part of a consultation carried out by the UK Intellectual Property Office.[6] While the original research was designed to assess the economic effects of introducing a copyright exception for parody in the UK (the political context of the Hargreaves Review[7] is explained in section C below), the further assessment of parody in the context of "takedown" practices offers an opportunity to take into account wider cultural, social and political features of those videos. Parody is controversial, because while it is recognized as engaging fundamental norms of freedom of expression, the creation of a successful parody necessarily draws upon and may copy aspects of an original work. This makes our sample unrepresentative of user-generated content as a whole, but usefully relevant to the study of takedown behavior.[8]

6 The initial sample of 1,839 parody music videos was obtained by searching a list of the top-charting music tracks in the UK for the 12 months preceding January 2012. Along with information such as the number of views, parodic intent and production values present in the user-generated parodies, the research team also recorded the uniform resource locator (URL) of each parody video. An extended group of researchers later revisited URLs of user-generated videos at two intervals: January 2013 and December 2016.[9] At each interval, the original URLs were checked to ascertain whether the video was still accessible on the YouTube website and if not, the reason for its removal (where this was possible to determine).

7 The analysis of information about both parody videos and original works enables examination of the relationship between risk of takedown and features of user-generated parody videos, such as its expressive content, genre, production values, and country of origin. This offers for the first time a window into takedown behavior in the context of stated policy concerns of rightholders.

8 The longitudinal aspect of the study enables us to further explore the rationales underlying rightholder opposition to policy change. An exception for "caricature, parody or pastiche" was introduced into UK Law with effect from 1 October 2014, in the middle of the longitudinal data collection.[10] If rightholders were not rigorously and systematically protecting their copyright from parodic treatment prior to the exception, this weakens public policy arguments opposed to such an exception. If they did not significantly change behavior after introduction of an exception, it raises questions about the salience of national copyright law for regulating online expression.

9 The paper is structured as follows. In Section B, we provide an overview of the technical and legal context, explaining the emergence of YouTube (and its content identification technology) and the status of user-generated services under "safe harbor" regimes (which have developed into a dominant mode of Internet regulation, limiting liability of intermediaries under certain conditions).

10 Next, we offer an analysis of the introduction of an exception for parody into UK law, following the Hargreaves Review of 2011 that recommended a suite of copyright reforms aimed at encouraging

Notice and takedown in everyday practice. Project report 2016. (Available at: <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2755628>, last accessed 29/09/2017). See also P. J. Heald, How Notice-and-Takedown Regimes Create Markets for Music on Youtube: An Empirical Study, 83 UMKC L. Rev. 313, 328 (2014) and D. Seng, "Who Watches the Watchmen?" An Empirical Analysis of Errors in DMCA Takedown Notices (2015), available at SSRN: <https://ssrn.com/abstract=2563202>, last accessed 20/10/2017).

6 The parody research study consisted of three distinct *Independent Reports for the UK Intellectual Property Office* (2013) commissioned in the context of the implementation of the Hargreaves Review of Intellectual Property and Growth (2011): (1) K. Erickson, Copyright and the Economic Effects of Parody: an empirical study on music parody videos on YouTube; (2) D. Mendis, M. Kretschmer, The Treatment of Parodies under Copyright Law in Seven Jurisdictions: a comparative review of the underlying principles; and (3) a synthetic summary applying the identified legal factors to the empirical findings, thus offering a range of policy options. The studies were used in the UK Government's preparatory documents for legislation implementing the recommendation (Hargreaves Review Impact Assessment, BIS1057, 2012, Copyright exception for parody, p. 10).

7 I. Hargreaves, *Digital Opportunity – A Review of Intellectual Property and Growth* (2011) (Available at: <https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/32563/ipreview-finalreport.pdf>, last accessed 20/10/2017)

8 *Campbell v. Acuff-Rose* 510 U.S. 569 (1994), at 588: "When parody takes aim at a particular original work, the parody must be able to "conjure up" at least enough of that original to make the object of its critical wit recognizable". Cf. Taking Forward the Gowers Review of Intellectual Property: Proposed changes to copyright exceptions, Newport: Intellectual Property Office (2008); R. Deazley, Copyright and Parody: Taking Backward the Gowers Review?, The Modern Law Review, pp. 785-823 (2010).

9 We are grateful for research assistance from Hossein Hassani and Andrea Varini at Bournemouth University in collecting the first wave of takedown data in 2013. The second wave of takedown data was added in December 2016 by Sabine Jacques and Morten Hviid at the University of East Anglia.

10 The Copyright and Rights in Performances (Quotation and Parody) Regulations 2014 No. 2356.

innovation and growth. In the UK government's evidence-gathering consultation on the proposal to create a new "fair dealing" exception for caricature, parody or pastiche, certain music rightholders were opposed to the plan, arguing that it undermined their economic and creative interests. By analyzing these policy arguments from rightholders, we identify various propositions about expected behavior.

11 Section C identifies and categorizes factors that may influence takedown of parody videos. The sample selection, variables and analysis methods are explained. We broadly classify four groups of factors that could influence a takedown: (1) commercial factors (including factors intrinsic to the original commercial work and its parodies); (2) moral/artistic factors; (3) cultural factors; and (4) behavioral factors related to the activities of the parodist. A Cox proportional hazards analysis model is estimated to investigate the impact of parody characteristics on the likelihood of removal over time, identifying those factors that are statistically significant.

12 In the concluding discussion we explore specifically whether and how music rightholders used notice-and-takedown procedures to protect their interests, and whether takedown behavior on YouTube is consistent with public opposition to a fair dealing parody exception in the UK.

13 This research is the first attempt at a longitudinal study of takedown for a cohort of user-generated works. The findings make an important advance in the empirical understanding of takedown behavior. Without understanding how current notice-and-takedown procedures are being used, it is impossible to project how future policy reforms might alter the online communication landscape. The findings allow us to evaluate legislative pressure to prescribe automated notice systems and pre-emptive removal (filtering on the basis of content recognition technologies, plus "stay-down" obligation once an initial takedown request has been made).[11]

B. YouTube as a Research Site: Technical and Legal Context

14 Founded in 2005 by former employees of the online payment system PayPal, YouTube is the world's most visited online streaming video platform. YouTube was initially acquired by Google in 2006 for USD$1.65 billion and since that time has integrated contextual advertising and search technology from its corporate owner. As of July 2017, the company claimed 1 billion users, making up a third of total global Internet traffic.[12] Despite the huge visitorship attracted by videos on the website, YouTube has not published public information about its profitability. In 2009 the New York Times estimated that YouTube's revenues might fall anywhere in a range from $200 million to $500 million USD per year, with the company reported to have reached profitability in 2011.[13] In a 2016 interview with Fortune, CEO Susan Wojcicki stated that "the company is still in investment mode" and may not currently be profitable due to technological investment and expansion into foreign markets.[14]

15 Initially, YouTube content consisted almost entirely of user contributions, and was considered emblematic of the "web 2.0" business model, leveraging user-generated content and social interaction to attract a user base.[15] Copyright infringement was initially a significant problem for the platform. The availability of content owned by third parties made YouTube the target of copyright infringement lawsuits, notably by cable provider Viacom in 2007.[16] In Europe, YouTube was sued by RTI in 2008[17] for hosting clips and episodes of the Italian Big Brother TV program and in France by TF1 in 2012[18] for hosting clips of programs belonging to the French broadcaster. In almost all cases (with the exception of the Italian RTI case) YouTube has enjoyed immunity from liability for infringement by its users because of its status as an information service provider (see next section for an explanation of so-called "safe harbor" provisions

11 British Phonographic Industry, "Urgent Reform" Needed to Notice and Takedown as Removal of 200 Millionth Illegal Search Result from Google Approaches, 24 March 2016. Available at: <https://www.recordoftheday.com/news-and-press/urgent-reform-needed-to-notice-and-takedown-as-removal-of-200-millionth-illegal-search-result-from-google-approaches>, accessed 1 July 2017. Stakeholder letter, Creative Sector shows united front to tackle the value gap: "UUC platforms have become major distributors of creative works - all while refusing to negotiate fair copyright licenses, if at all, with the right holders", 4 October 2017. Available at: <http://impalamusic.org/content/creative-sector-shows-united-front-tackle-value-gap>, accessed 20 October 2017.

12 YouTube in Numbers <https://www.youtube.com/intl/en-GB/yt/about/press/>, accessed 2 July 2017.

13 Tim Arango, As Rights Clash on YouTube, Some Music Vanishes (New York Times, 22 March 2009), available at: <http://www.nytimes.com/2009/03/23/business/media/23warner.html>, accessed 1 July 2017.

14 Leena Rao, "YouTube CEO Says There's 'No Timetable' For Profitability", available at: <http://fortune.com/2016/10/18/youtube-profits-ceo-susan-wojcicki/>, accessed 29 June 2017.

15 J. Burgess and J. Green, *YouTube* (2009 Polity Press).

16 *Viacom International, Inc. v. YouTube, Inc.*, No. 07 Civ. 2103.

17 *Reti Televisive Italiane contro YouTube*, Trib. Roma, 24 novembre 2009, n.54218/08 (It.).

18 *TF1, TF1 Video, TF1 droits audiovisuels, LCI and e-TF1 v. YouTube* (RG: 10/11205), Tribunal de Grande Instance, Paris, 29 May 2012.

in the US and EU). In most cases, courts have found that due to the volume of material processed by platforms such as YouTube, service administrators cannot be held liable for unauthorized use without obtaining specific knowledge of infringement. Claimants have been pointed to the notice-and-takedown mechanism as a remedy for the removal of infringing content on sites like YouTube.

16 Over time, conflict with rightholders has led YouTube to develop more sophisticated measures for preventing the uploading of copyright material in the first place and empowering rightholders to locate and remove material hosted by the website via its fingerprint matching technology called ContentID. This system works by comparing existing and newly-uploaded contents to an "index file" of video or audio material provided by a rightholder. If a user-uploaded video is matched to an audiovisual work in the reference file, the appropriate rightholder is notified. Rightholders who participate in the ContentID system may then choose to i) have the video removed, ii) leave the video accessible while muting the infringing audio, iii) leave the video up and monetize it to collect a share of the advertising revenue, or iv) track it and do nothing.[19] Rightholders may issue their own takedown notices independently to the website even if they do not participate in ContentID.

17 While YouTube has strengthened its ability to respond to rightholder complaints, considerable amounts of commercial content has appeared on the platform through partnerships with traditional and emerging media businesses. One of the most significant of these partnerships is the VEVO music channel, which hosts content licensed from Sony Music Entertainment, Universal Music Group, Abu Dhabi Media and EMI. The participating music labels benefit from a revenue share model that divides the proceeds earned from contextual advertising, pre-roll video advertising, merchandise, and iTunes music downloads. VEVO, along with similar channels controlled by Warner music, Sony BMG and Universal Music Group, have proven extremely popular; data compiled by ratings research company ComScore shows that commercial music videos remain the most popular type of content on the platform, accounting for more than 180 million unique monthly viewers in the USA, and making up half of the largest channels in the top ten by viewership.[20]

18 The popularity of commercial music video content, combined with the large volume of user-generated content on YouTube, makes it a compelling site to study the effects of derivative use such as parody. YouTube's business model, which depends equally on traditional and user-generated content, locates it in a precarious position; on one hand needing to placate rightholders concerned about the integrity and commercial viability of their licensed content, and on the other hand requiring participation from users who demand the ability to use and remix copyright material in new ways. This dilemma remains a source of conflict between the various user communities and content creators on the service, with copyright law providing a general framework in which conflicts are resolved.

C. Status of online intermediaries

19 We now review briefly the legal status of online intermediaries under copyright law and examine the notice-and-takedown mechanism that rightholders can employ to remove unwanted infringing content from services such as YouTube.

20 A so-called "safe harbor" for "Online Service Providers" that offers immunity from claims to copyright infringements under certain conditions was first introduced in the United States under Section 512 of the U.S. Copyright Act (as amended by the Digital Millennium Copyright Act – DMCA 1998).[21]

21 Section 512 specifies a formal procedure under which service providers need to respond expeditiously to requests from copyright owners to remove material. Rightholders who wish to have content removed must provide information "reasonably sufficient to permit the service provider to locate the material" (such as a URL) and warrant that the notifying party is authorized to act on behalf of the owner of an exclusive right that is allegedly infringed. The practice is known as "notice-and-takedown". Importantly, "counter notice" procedures are also specified under which alleged infringers are notified that material has been removed and can request reinstatement.

22 Similarly, under the EU Directive on Electronic Commerce (2000/31/EC) hosts of content uploaded by users will be liable only upon obtaining knowledge of the content and its illegality. But unlike the DMCA, the E-Commerce Directive does not regulate the procedure of receiving the necessary knowledge but leaves this up to the Member States. Husovec (2017) summarizes the position concisely: "The case-law of the CJEU only requires that the perspective

19 YouTube, How Does ContentID Work? <https://support.google.com/youtube/answer/2797370?hl=en-GB>, accessed 1 July 2017.

20 ComScore, Top 10 YouTube Partner Channels By Unique Viewers February 2016 <https://www.comscore.com/Insights/Rankings/comScore-Releases-February-2016-US-Desktop-Online-Video-Rankings>, accessed 2 July 2017.

21 See 17 U.S.C. §§ 512, 1201-1205 (2000).

of a 'diligent economic operator' is decisive. The constructive knowledge can be obtained in any situation, including, as a result of an investigation undertaken on the provider's own initiative, as well as a situation in which the operator is notified of the existence of such activity or information, but perhaps not sufficient to constitute actual knowledge."[22]

23 In the majority of cases dealing with copyright infringement, YouTube has been deemed by courts in the USA and Europe to fall within the definition of a Service Provider benefitting from exclusion from liability for copyright infringement. Both the DMCA and the E-Commerce Directive place the burden of responsibility on rightholders to identify infringing material and notify the service provider of its presence. In order to comply with these provisions across different jurisdictions, YouTube has invested significantly in developing an online system to receive and respond to notice-and-takedown requests from rightholders. At the same time, the platform also discourages users from uploading infringing material, and polices remove repeat infringers from their revenue-sharing partnership status and accounts.

24 By placing the burden of policing copyright infringement on the shoulders of individual copyright owners rather than on network service providers, jurisdictions such as the USA and the EU, which have adopted these safe harbor provisions aim to enable early-stage innovation on the Internet, limiting the costs of copyright enforcement. However, the present balance of responsibility has fallen under criticism. Rightholders have protested that this legislation burdens them with disproportionate costs, and that intermediaries – possessed of access to digital technologies and user data – should be obliged to do more to proactively find and eliminate infringing content. On the other hand, online free speech advocates have protested that the notice-and-takedown mechanism is open to abuse by parties who wish to suppress unpopular and dissenting speech, by using the copyright infringement claim as an excuse to force intermediaries to remove content.[23] While notice-and-takedown is an effective measure to stop direct piracy of content, neither rightholders nor Internet intermediaries have developed due process for making judgments about "fair" derivative or transformative uses. Understaffed and risk-averse, online platform operators may simply choose to comply with a takedown notice, rather than risk safe harbor protection by standing up for a user who may indeed benefit from a copyright exception.

25 Takedowns of apparently fair dealing derivative works have proven particularly controversial in recent years. Under section 512(c) of the DMCA, a takedown notice must contain a statement by the copyright holder of a good faith belief that there is no legal basis for the infringing use identified by the complaint. Subsequently, US courts have found that complainants may have an obligation to consider fair use before issuing such takedown notices, or face liability for misrepresentation of infringement. Currently, users who are unhappy about the removal of their videos from YouTube may file a counter-notification consisting of a warranty that they are legally entitled to make use of the work, however, small-scale users may be deterred from doing so because of confusion or fear of further legal action by rightholders.[24]

26 In the case of *Lenz v. Universal Music Corp.*,[25] a California District Court upheld a complaint that the music label had failed to take into account the fair use of material when it issued a takedown notice to YouTube over a video that the complainant had uploaded to the service. The video, a short clip of the complainant's toddler dancing, triggered the takedown request because the song playing in the background was Prince's *Let's Go Crazy*, owned by Universal Music. The case highlighted an important feature of the existing notice-and-takedown mechanism: dependency on automated "fingerprinting" technology used by rightholders to locate infringing material (in this case by Prince) can result in false positives that would otherwise be covered by fair use. A second issue highlighted by this case is that the whim of one artist can generate thousands of takedown notices while derivative uses of other artists' work remains untouched[26]. There is no consistently applied set of rules governing the removal of derivative online use of copyright work.

27 The proposed EU Directive on Copyright in the Digital Single Market (COM(2016) 593 final, 14.9.2016) introduces a new provision (Article 13) that will create (under some readings) an obligation for information service providers to prevent the availability of infringing works in the first place. This new "notice-and-stay-down" obligation has

22 M. Husovec, *Injunctions Against Intermediaries in the European Union,* Cambridge University Press (2017), p. 53 (analyzing *L'Oréal and Others*, Case C-324/09).

23 J. Miller, Fair Use through the Lenz of §512 of the DMCA: A Preemptive Defense to a Premature Remedy?, Iowa Law Review (2010), 95, 1697-1729.

24 Fred von Lohman, "YouTube's Content ID (C)ensorship Problem Illustrated", Electronic Frontier Foundation (2 March 2010), available online: <https://www.eff.org/deeplinks/2010/03/youtubes-content-id-c-ensorship-problem>, accessed 20/10/2017.

25 *Lenz v. Universal Music Corp.,* 572 F. Supp. 2d 1150 (N.D. Cal. 2008).

26 The particularly aggressive stance of Universal Music has been dubbed the "Prince Policy" due to that artist's notoriously strict stance on online use of his work (Miller, *supra* footnote 23).

been welcomed by rightholders (in particular from the music industry) as improving their bargaining position vis-à-vis services such as YouTube. The wording of Article 13 has been criticized as a "censorship filter", [27] and is close to mandating general monitoring (that would be in conflict with Article 15 of the E-Commerce Directive and CJEU case law).[28] This remains a fast moving field of policy.

D. Parody in Copyright Policy: The UK Example

28 The original dataset of 1,839 music video parodies examined here was collected in the context of the UK Hargreaves Review in 2012. The policy consultation process provided an opportunity to gather responses from rightholders to the proposed copyright exception, enabling us to generate propositions about expected takedown behavior. This section explains the context of Hargreaves' proposals and the responses by music industry rightholders.

29 There have been numerous arguments made in support of statutory copyright exceptions for parody, such as in the UK.[29] The 2006 Gowers Review of Intellectual Property recommended that the government adopt such an exception on the grounds that it would promote the creation of valuable new works and reduce transaction costs by removing the need for licensing in certain cases.[30] In his 2011 review, Professor Ian Hargreaves similarly recommended the creation of a new fair dealing exception for parody, on the grounds that allowing unlicensed parody would generate growth for UK media industries, and would "encourage [...] literacy in multimedia expression in ways that are increasingly essential to the skills base of the economy".[31] Parody is understood to be a fundamental part of political and cultural life in the UK, with the Government citing its "long and vibrant tradition" in UK comedy.[32]

30 While both the Hargreaves and Gowers reviews stressed the generative effects of transformative use for the creative industries, some industry groups took a strongly opposing stance toward the proposed legislation. In response to the Hargreaves Review, the Music Publishers Association (MPA) wrote:

> *The proposed exception for parody would undermine the integrity and moral rights of publishers and cut across their normal licensing activities, whether for the purpose of synchronization or straight forward adaptation of the lyrics or musical style. Carving out an exception which meant that "parodists" would not have to pay for comic use of musical material undermines the business model of a music publisher. (MPA 2012)*

31 Distilling arguments contained in the 471 industry responses published in the Government Consultation on the Hargreaves Review, and in particular those that opposed the introduction of a copyright exception for parody, we identify three common concerns on the part of rights owners.[33]

32 First, certain respondents argued that *permitted unlicensed parody would deprive rightholders of a legitimate stream of licensing revenue.* Wider economic interests such as those cited in the Hargreaves review, needed to be balanced against the threat to licensing revenue earned by rights owners for permission to make use a work, including uses that might fall under the proposed fair dealing exception for parody. The Design and Artists Collecting Society (DACS), which represents the interests of visual creators (including photographers and graphic illustrators), stated in its response to the Hargreaves consultation that, "[r]ightsholders will lose an established stream of revenue from the licensing of their work for parodies which go beyond the established limitations of substantial taking and criticism and review" (DACS, 2012: 44).

27 J. Reda, When filters fail: These cases show we can't trust algorithms to clean up the internet (available at: <https://juliareda.eu/2017/09/when-filters-fail/>, accessed 20/10/2017).

28 *L'Oréal/eBay* (C-324/09, 12 July 2011), *Sabam/Netlog* (C-360/10, 16 February 2012). For a contrary view, see A. Lucas-Schloetter, "Transfer of Value Provisions of the Draft Copyright Directive" (March 2017, p. 19). According to Lucas-Schloetter, the prohibition on general monitoring does not apply "when the infringing content to be searched for is identified" (available at: <http://www.authorsocieties.eu/uploads/Lucas-Schloetter%20Analysis%20Copyright%20Directive%20-%20EN.pdf>, accessed 20/10/2017).

29 In the United States, parodies typically are considered as "fair use" under section 107 of the Copyright Act 1976. Under the fair use doctrine, factors to consider include the purpose and character (e.g. commercial/non-profit educational use), substantiality of the portion used, and the effect of the use upon the potential market. The case of *Acuff Rose Inc. v Campbell* (510 U.S. 569, 1994) established that the "transformative" nature of the parodied work is decisive: Does it add "something new, with a further purpose or different character, altering the first with new expression, meaning or message"?

30 Gowers (2006) 68.

31 Hargreaves (2011) 50.

32 Intellectual Property Office, Consultation on Copyright (2012). Available online: <http://www.ipo.gov.uk/pro-policy/consult/consult-closed/consult-closed-2011/consult-2011-copyright.htm>, accessed 20/10/2017.

33 For a discussion of discourse analysis method applied to consultation responses, see K. Erickson, User illusion: ideological construction of "user generated content" in the EC consultation on copyright, Internet Policy Review 3(4), pp. 1-19 (2014) (available at: <https://policyreview.info/articles/analysis/user-illusion-ideological-construction-user-generated-content-ec-consultation>, accessed 27/10/17).

33 The second argument made by rightholders in opposition to the proposed parody exception was that widespread unlicensed parodies might *compete unfairly with original works in the marketplace, either by substituting for the original, or by causing unwanted reputational damage*. These two related arguments were explored by Rogers[34] in a study commissioned by Consumer Focus, the UK statutory body that represents consumers across regulated markets, and they have been cited by commentators on both sides of the debate; although Rogers and co-authors note that there is an absence of empirical evidence with which to evaluate these claims. The first part of the argument, that unlicensed parody might substitute for an original work, seems unlikely given the nature of parody: the successful parodist must conjure up knowledge of an original work in an audience member's mind in order for the parody to be effective, assuming prior knowledge of the original work. There is the additional possibility that the circulation of a popular parody might stimulate consumption of an original work, when new fans of the parody are reminded of the original. The second part of the argument articulated by Rogers et al – that parody might cause reputational harm to an original – is difficult to test empirically, although there are normative questions to be raised about how far copyright protection should impede the free flow of market information regarding the quality of goods, such as that enabled by neighboring copyright exceptions for purposes of criticism and review.

34 A third argument made in opposition to the proposed parody exception in the wake of Hargreaves is that *derogatory treatment of an original by parodists could infringe on the original authors' moral rights*. Outlined in sections 77-85 of the UK Copyright Design and Patents Act (CDPA 1988), moral rights consist of the rights of an author to be identified as the creator of a work (paternity), to prevent misidentification as the author of a work, and to object to derogatory treatment of a work that he or she has authored (integrity). It is principally the latter that opponents argued could be endangered by the introduction of a copyright exception for parody. In fact, the wording of the proposed parody exception was explicitly written so that it shall not infringe on an author's moral right. Nevertheless, it is foreseeable that some authors could object to certain parodic treatments of their work and may wish to prevent transmission of such work by asserting their moral rights.

35 The arguments articulated above are largely theoretical – prior to the Hargreaves consultation exercise, no rigorous empirical studies of the economic effects of parody existed. Much of the prior discussion of parody is either anecdotal, focusing on key cases and disputes involving single works, or represents the views of industry bodies or collecting societies (such as the Music Publishers Association and DACS, cited above). If we assume that the aggregate views expressed by collective bodies are representative of their members' economic interests as a whole, we should expect to find corresponding empirical evidence that supports those concerns expressed in the published responses to the Hargreaves review. For example, if infringement of moral rights is a major concern, we should expect to see some rightholders systematically withholding certain works from parody or objecting to certain derogatory types of parodic treatment. Similarly, if protecting work from substitution by parodic imitators is of concern, we should expect to see those parodies that attract significant viewership taken down with greater regularity. In the following section, we describe the research method used to observe rightholder behavior, using data on takedowns gathered from music videos and their related parodies on YouTube.

E. Research Design and Method

36 We initiated data collection in 2012 for the purposes of assessing the economic implications of introducing an exception for Parody into UK copyright law. The researchers sought to ascertain the quantity of user-generated parody content on YouTube and review their effect on commercial works parodied.[35] We used the top-100 list of monthly songs tracked by the British Charts Company to obtain a list of the 343 most popular songs released in the UK in the previous 12 months. These songs were matched with a corresponding licensed music video hosted on YouTube (such as via VEVO or other record labels' official channels). As a second step, searches for parody videos referencing those commercial works were performed by searching for "song name + parody" in YouTube's internal search engine. The researchers located 8,299 user-generated music video parodies referencing the original 343 commercial music videos. A randomly-selected sample of 1,839 parodies from within that larger population was subjected to closer analysis by human coders to determine the nature of the parody, the severity of critique, the production values used, and the extent

34 M. Rogers, J. Tomalin and R. Corrigan, The economic impact of consumer copyright exceptions: A literature review, Consumer Focus (2009), (available at: <http://oro.open.ac.uk/25604/5/The-economic-impact-of-consumer-copyright-exceptions-Rogers-Tomalin-Corrigan.pdf>, accessed 20/10/17).

35 See K. Erickson, M. Kretschmer and D. Mendis, Copyright and the Economic Effects of Parody: An Empirical Study of Music Videos on the YouTube Platform and an Assessment of the Regulatory Options, Intellectual Property Office 2013 (available at: <https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/309903/ipresearch-parody-report3-150313.pdf>, accessed 27/06/17).

of the original work copied. The research team also recorded the location (URL) of each of the initial 1,839 parody videos to enable future analysis.

37 Following the original collection of data in 2012, it was decided to observe the videos again to obtain a new perspective on "takedown" policies. For this purpose, one year after the original study, in January 2013, the team re-visited the list of parody URLs to check whether those videos were still live on the platform or whether they had been removed.[36] In 2016, colleagues at the University of East Anglia (UEA) collected an additional wave of takedown data based on the original sample, establishing which parody videos were still live four years later. The two research teams pooled these data together (which the UEA team then analyzed for a study on cultural diversity).[37]

38 The two waves of follow-up study allowed inclusion of the additional variable of the removal of user-generated parody videos, first at one year and then at four years after they were first observed. In both waves, researchers differentiated where possible whether the takedown was initiated by a rightholder, or whether the video was removed by the uploader for unknown reasons (see Table 1).

39 Table 1: Music video parody sample decay rate due to takedown 2012-2016

	January 2012	January 2013	December 2016
Total Accessible	1839	1471 (79.9%)	1088 (59.2%)
Cumulative Taken down for copyright (%)	--	265 (15.5%)	606 (32.9%)
Cumulative Taken down for unknown reason (%)	--	103 (5.6%)	145 (7.9%)

40 When the dataset was revisited in January 2013, some 265 (15.5%) of the original 1,839 videos had been removed by a likely copyright complaint. This was ascertained by checking the notice that appeared in front of inaccessible videos. For example, blocked videos could indicate that they were "unavailable due to a copyright complaint" or "no longer available in your territory" (also due to copyright). In 2016 when researchers Jacques et al re-examined the original dataset, they found that an additional 341 videos had been removed for copyright reasons in the intervening period after 2013. The overall yearly rate of decay after 2013 is therefore lower, but still significant, bringing the total number of accessible videos down to 1,088 from the initial 1,839. Some 145 of the missing videos were removed for reasons other than a copyright notice, likely by the uploader themselves. These videos are considered separately from the instances of copyright takedown and are treated as censored in our analysis (see below).

41 Building from policy arguments made by music rightholders identified in the preceding section, we propose and categorize a range of different factors that may influence whether a copyright takedown is initiated (see figure 2). The full list of variables and their parameters is provided in Annex 1. Specifically, we identify four groups of factors that could influence a takedown: (1) commercial factors (including factors intrinsic to the original commercial work and its parodies); (2) moral/artistic factors; (3) cultural factors; and (4) behavioral factors related to the activities of the parodist.

42 Table 2: Factors that may influence takedown of parody videos

Commercial factors	Cultural factors
Sales rank of original	Genre: rock
Parody views	Genre: Electro
Parody production values	Genre: Hip hop
Monetization on parody	Territory: UK
	Territory: USA
	Major/independent record label
Moral/artistic factors	**Behavioral factors**
Parody type: target	Copied sound recording
Parody type: weapon	Copied video recording
Severity of criticism	Lack of intent (mislabeled parody)
	Parodist appears on camera
	Gender of parodist

43 *Commercial factors.* One argument advanced by rightholders in opposition to a copyright exception for parody focused on the potential for reduced commercial revenue from loss of the ability to license to parodists. While the music industry and collecting societies do not often publish information about the frequency of licensing or the agreed terms, we can assess this claim within our study by considering commercial features intrinsic to parodies. In particular, we examine whether a parody video is accompanied by monetization in the form of pre-roll or mid-roll advertisements, and whether it was created with low or high production values,[38] indicating commercial quality. We also include the popularity of the parody video in our analysis, using the number of views originally measured in January

36 It should be noted that in all waves, researchers checked for removed videos using a UK-based IP address.

37 See S. Jacques, K. Garstka, M. Hviid and J. Street, The Impact on Cultural Diversity of Automated Anti-Piracy Systems as Copyright Enforcement Mechanisms: An Empirical Study of YouTube's Content ID Digital Fingerprinting Technology, SSRN 2017 (available at: <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2902714>, accessed 27/06/17).

38 Production values were recorded by asking human research assistants to rate them on a Likert-style scale from 1 (lowest) to 5 (highest).

2012. Based on rightholder statements, we might expect higher rates of takedown for parody videos with higher production values and popularity, reflecting concern over substitution and potential loss of licensing revenue anticipated by rightholders. YouTube carries content of varying quality, ranging from purely amateur, non-commercial video to semi-professional and commercial video produced by entrepreneurs and firms. Since this latter group potentially derives revenue from their activities on and off YouTube, it is reasonable to expect that rightholders would target these potential licensees more readily than non-commercial users, where the likelihood of paying for use of an original work is low.

44 In addition to factors intrinsic to parodies themselves, we also consider the commercial appeal of the underlying musical work. In the current study, the feature of the original music track and video that we examine is sales popularity of the original work (based on its position and duration in the top-100 UK charts). The variable "sales rank" captures the relative position and duration of the original song on the UK top-100 music charts in 2011.

45 *Moral / artistic factors.* An additional set of arguments raised by rightholders in opposition to a parody exception related to artistic qualities of the parodies themselves. Related to the commercial factors above, one source of opposition from rightholders was the apprehension that negative parodies could impact the market for an original work by harming the reputation of the artist or the work itself. Reputational harm in the market is difficult to measure; the impact of a negative review may take years to propagate and produce an effect. Our data provide an opportunity to detect whether rightholders are concerned by reputational damage, independent of whether such damage actually materializes. Approximately 33% of the parodies in the original dataset were "target" parodies, meaning that they explicitly took as an object of ridicule the original work or its creator. By contrast, "weapon" parodies use an original piece of content to draw attention to some third-party individual or issue. We include dummy variables for both types to analyze the importance of parodic intent. If rightholders are concerned about the potential for reputational harm produced by online parody, we should expect to see that they issue more takedowns for negative "target" parodies. Another factor relates to the moral right that the original artist may have to object to a derogatory treatment of their work. It is difficult to assess whether a parody produced under a fair dealing exception such as that available in the UK could infringe the moral rights of artists. However, it is possible that moral rights concerns drive rightholder behavior.[39] To explore this, we analyze a subsample of parodies containing the most explicitly negative messages (severity of critique) to test whether this has a statistically significant effect on the likelihood of a takedown.

46 *Cultural factors.* This group of factors relate to differences in the legal culture between territories, as well as differences in the creative practices of specific musical genres or businesses (music labels), which may influence the observed pattern of takedown. Since the UK did not have a statutory exception for parody until October 2014, the availability of the fair use defense to parodists in the USA may be expected to produce a difference in the level of tolerance for parodies reflecting the different legal culture of the two countries. To assess this influence, we record and include the national territory of the music publisher in our analysis, using a dummy variable for original songs originating in the UK. It has been widely observed in scholarship on media production that different mediums, and even sub-genres are characterized by differing production practices, in particular relating to tolerance of sampling or borrowing from pre-existing works.[40] To assess whether genre has an influence on likelihood of takedown, we include dummy variables in our analysis for Rock, Electronic and Hip Hop music, with "Pop" as the reference category. Finally, the business practices of specific music labels may be a factor in whether parody videos are taken down. Specific businesses may have internal policies that are more or less tolerant of online uses. A young, up-and-coming independent label might actively encourage YouTubers to parody their artists' works, while a more established corporate player might be more restrictive, for example. To capture potential effects from individual music labels, we include a dummy variable for songs owned by major, as opposed to independent music labels.[41]

47 *Behavioral factors.* This group of factors relates to the decision making and behavior of the parodist/uploader when creating and sharing their video. One important set of factors relates to the underlying

39 And indeed, this is possible given the anecdotal reports of displeasure by specific artists concerning online uses of their works. See Miller and the "Prince effect" (*supra* footnote 26).

40 See A. Sinnreich. *Mashed up: Music, technology, and the rise of configurable culture*. University of Massachusetts Press (2010), 107-123.

41 Major labels are defined as belonging to one of the "big three" - Universal Music Group, Sony Music Entertainment, and Warner Music Group, or their sub-labels (including Atlantic, Capitol, Parlophone and EMI, among others). The authors are grateful to Matthew Sag at Loyola University Chicago for his suggestion to include possible label effects in the analysis. The initial collection of music label information was carried out by Sabine Jacques, with additional coding by the authors.

material directly copied or added by the parodist when making their video. In our initial sample, many parodists copied portions of the original composition or sound recording in their uploaded video. A smaller group of parodists copied the original music video itself, although most parodists in our sample chose to create new video content as well as lyrics to accompany their derivative work. This is unsurprising, as YouTube is primarily a platform for video expression and uploaders may lack the musical ability to perform and record a new sound recording riffing on an original song without directly copying it. Another behavioral factor relates to the parodic intent – or lack of discernible intent – of the uploader. As previously discussed, the majority of videos corresponded to two known types of parody: "target" (which takes aim at the original work) and "weapon" (which uses a work to draw attention to a different social issue or phenomenon). However, a further 13% of parodies in the initial sample had no discernible focus of critique, even though the parodist had tagged their uploaded video with the keyword "parody", making it detectable to our initial sampling method. We record this lack of parodic intent and include it as a dummy variable in the analysis, with the reference category being all other parodies where a focus of critique was evident. Finally, we record and include variables which capture the style of address and gender of the parody performer (female solo compared to male solo and mixed groups).

F. Analysis and Discussion

48 The data on YouTube takedowns, comprised of 1,839 cases, presents two challenges for analysis. One challenge relates to censoring of the data: while the observation period took place over 72 months, not all takedowns that may eventually occur are captured in our study. A second challenge is one of survivorship bias introduced by the removal of the most egregious infringing parodies immediately upon upload. In order to address these challenges, we perform a Cox proportional hazards analysis to examine the effect of covariates on time-to-removal.[42] This allows us to identify which variables are associated with an increase or decrease in the risk that a given user-generated video will be taken down. The dependent variable (event) in the analysis was the detection of a takedown (expressed as a binary variable: 1=yes, 0=no). Covariates include features of the parody video itself as well as features of the original commercial work (full descriptive statistics are provided in Annex 1). The time variable is the maximum number of months a video "survived" from upload to detection of a takedown.

49 The results of the Cox regression analysis are presented in Annex 2. Results are reported as hazard ratios, indicating an increased risk of takedown when the ratio is greater than 1, and a reduced risk when the ratio is less than 1. Columns 1-4 present the results for each of the groups of covariates, and column 5 presents the model with all main variables included. The "target" variable is not included in specification 5 due to multicollinearity with the other variable of interest, "severity of critique" (all severely critical parodies were target parodies).

Discussion 1: Commercial factors

50 In the preceding section, we identified one set of factors related to claims by rightholders that parody harms the commercial market for their works. To assess these claims, we analyze variables related to production values, popularity and commercial sales of original works to assess whether these factors influence the probability of rightholder action. A first observation from the analysis is that the commercial success (sales rank) of the underlying commercial release does not appear to have a significant impact on rightholder takedown activity. A second commercial concern for rightholders is the possibility for substitution by parodic works, which might compel them to remove parodies most popular with viewers. We observe a significantly negative effect for number of views on the risk of a takedown. This means that more popular videos (as measured in 2012) had a lower risk of being removed by rightholders.

51 A second, related concern for rightholders is the potential for lost licensing revenue from parody videos that have commercial potential. The proxy variable used to capture commercial potential in this analysis is the level of production values in the parody (initially measured by human coders using a Likert-style rating from 1-5). Parody videos with higher production values may reflect creators with access to more resources and more funding compared to amateur producers. Overall, higher production values reduced the risk of a takedown compared to videos with average or low production values. There are several potential explanations for this result: commercially-minded YouTubers may benefit from pre-existing licensing agreements (for example through membership in multi-channel networks); highly skilled parodists may benefit from knowledge which helps avoid automated takedown (for example by performing their own musical rendition to accompany the parody); or,

42 For further discussion of suitability of the Cox proportional hazards model to analysis of cohort data in an organizational setting, see A. Scherer, N.V. Wünderlich and F. Von Wangenheim, The Value of Self-Service: Long-Term Effects of Technology-Based Self-Service Usage on Customer Retention. MIS Quarterly 2015, 39(1).

rightholders may be engaged in a form of brand management, trimming videos that they feel do not meet standards of quality aligned with their objectives as entertainment brands.

Discussion 2: Artistic characteristics

52 In the preceding section, we characterized two claims originating from rightholders that the artistic qualities of a parody might be harmful to artists. The first proposition relates to the potential for reputation harm arising from negative parodies, which target the artist or the original work. Among the covariates in Annex 2, we include two dummy variables for "weapon" and "target" parodies, to test the impact of negatively targeting the original work on the risk of a takedown. The reference category is all other mislabeled parodies where no clear intent could be ascertained. We observe that the effect on risk of takedown for both weapon and target parodies is negative. It appears that having a clear parodic intent, even if critical of the original work, benefits the survival of parodies.

53 A second claim was the potential for parodies to infringe the moral rights of creators (one rationale for curtailing exceptions to copyright). The influence of moral rights concerns on the takedown rate is complicated by the range of potential objections that an author might have to a transformative use of their work. We assume that "derogatory treatment" in the eyes of a creator is likely to include use that de-values the original for a new audience.[43] One possibility is that a parody could be placing a work in an objectionable context.[44] The variable "severity of critique" was included in column 4. It has a significantly negative effect on takedown risk, strengthening the interpretation that a clear target of attack is more beneficial than having no target at all.

Discussion 3: Cultural factors

54 We analyze another set of factors related to the cultural context of music production and legal culture of the territories of the original artist. A significant factor for likelihood of a takedown is the genre of the underlying musical work. We find that for parodies of rock music, the risk of takedown is significantly reduced compared to pop, hip hop, and electronic music. This finding remains stable and significant across different specifications. The result is surprising, counterintuitive to existing scholarship which suggests greater tolerance for sampling and re-use in art forms such as hip hop music. Our result may reflect an overall permissive tendency in less popular, traditional music. Rock music right holders may not be interested in enforcing copyright on YouTube due to a focus on traditional commercial channels of distribution. Other than musical genre, the other main cultural factor influencing takedown was territory of the original artist. For original works by artists based in the USA, the risk of takedown was significantly lower than for the UK and Europe. This may reflect the influence of fair use, or it may reflect greater tolerance on the part of American music rightholders to online user-generated expressive practices.

Discussion 4: Behavioral factors

55 Finally, we analyze factors originating from the behavior of parodists when creating and uploading their derivative works. One significant factor in this group is a lack of parodic intent on the part of the uploader. The result is positively significant (at the *p<.01* level). This result may reflect elimination of parodies where the uploader has tried to disguise their use as a parody. These could be straight copies of the original music video or could consist of "karaoke" covers. Lack of parodic intent is also correlated with lower production values, so the increase in takedown rate may also reflect brand management "pruning" by copyright holders

43 Section 80(2)(b) of the UK Copyright, Designs and Patents Act 1988 provides that the treatment of a work is derogatory "if it amounts to distortion or mutilation of the work or is otherwise prejudicial to the honour or reputation of the author or director". Reference to the wording of Art. 6 of the Berne Convention suggests that the author can only object to distortion, mutilation or modification of her work *if* it is prejudicial to her honor or reputation. Still one UK court has given a wide interpretation, considering the removal of a forest background from a photograph as a derogatory distortion (*Delves-Broughton v. House of Harlot Ltd* [2012] EWPCC 29).

44 An example is the notorious *Deckmyn* case before the CJEU (Case C-201/13) where the rightholders of *Suske en Wiske* hoped to stop a right wing political party from circulating a pamphlet that spoofs a famous cartoon cover, but this case was decided without reference to moral rights (which are not harmonized EU rights). At (27): "It follows that the application, in a particular case, of the exception for parody, within the meaning of Article 5(3)(k) of Directive 2001/29, must strike a fair balance between, on the one hand, the interests and rights of persons referred to in Articles 2 and 3 of that directive, and, on the other, the freedom of expression of the user of a protected work who is relying on the exception for parody, within the meaning of Article 5(3)(k)." If a discriminatory message is conveyed "which has the effect of associating the protected work with such a message" (at 29) (a case which it is for the national court to assess), (30) "attention should be drawn to the principle of non-discrimination based on race, color and ethnic origin, as was specifically defined in Council Directive 2000/43/EC of 29 June 2000 implementing the principle of equal treatment between persons irrespective of racial or ethnic origin (OJ 2000 L 180, p. 22), and confirmed, inter alia, by Article 21(1) of the Charter of Fundamental Rights of the European Union."

unhappy with poor-quality uses.

56 This group of factors includes measures of the amount of borrowing from original works (copied video or sound recording). Unsurprisingly, we find a significant positive relationship between the presence of an original sound recording and risk of a takedown. When parodists borrowed the underlying recorded track from a commercial work, their video was more likely to be rapidly taken down. Borrowing the original video in a parody also increased risk of takedown, but less significantly. This may reflect the immediate detectability of copied videos, with the most egregious copies taken down immediately, leaving only more robust derivative works that withstood subsequent takedowns. The impact of artistic borrowing on takedown rate may generally be explained by the use of ContentID by rightholders to automatically locate and policy infringing material (sound and video content).

57 **Table 3: Summary of factors influencing takedown of parody videos (waves 1 & 2)**

Commercial factors	Cultural factors
Sales rank of original	Genre: rock .560***
Parody Views .864***	Genre: Electro
Parody production values .898***	Genre: Hip hop
Monetization on parody	Territory: UK
	Territory: USA .724***
	Major record label?
Moral/artistic factors	**Behavioral factors**
Parody type: target	Copied sound recording 1.237***
Parody type: weapon	Copied video recording
Severity of criticism	Lack of intent (mislabeled parody
	Parodist appears on camera
	Gender of parodist

*Note: Significant variables reported as hazard ratios. Significance levels: *** p<0.01, ** p<0.05, * p<0.1*

G. Conclusion

58 This paper has examined the rate of copyright takedown of parody music videos to assess different claims made by rightholders about the features of parody that they find threatening to the artistic integrity and commercial exploitation of their work. Based on public policy statements, we expected to observe a higher probability of takedown for variables related to commercial substitution, as rightholders exercised their copyright to protect the market value of their works. Considering artistic factors, we expected to observe rightholders exercising their copyright to protect the artistic integrity of their expressions and reputation of artists.

59 Other than removal of direct forms of copying, rightholders do not appear to be acting in a way that corresponds to public statements about the artistic or commercial harm posed by parody works. One counterintuitive finding is that rightholders are not targeting the most popular or highest production value parodies, but appear to be doing the opposite. This contradicts the expected result, which is that rightholders should be concerned about substitution, and that they should seek to suppress commercial-quality derivative uses in favor of licensing use of their material. It is likely that the ability of rightholders to track and monetize derivative uses of their copyright material via ContentID partially explains the observed result. High-quality and popular parodies might remain live on the platform because rightholders have determined that the revenue gains from monetizing those unauthorized parodies weigh against any potentially negative effects such as substitution.

60 The use of ContentID monetization does not explain the disproportionate rate of takedown of parodies with lower production values, which is significant across specifications. Poor quality may be linked to a lack of copyright awareness on the part of uploaders. Parodists with less skill may be more likely to directly copy a sound recording, making their output more easily detectable by rightholders. The significance of direct copying on the risk of a takedown reinforces this possibility. Rightholders and their representatives may also be involved in brand management in their online takedown policy – leaving up those videos that are popular or reflect well on the artist's brand, while seeking to remove those that tarnish the artist due to their amateurish production values. Further research is needed to ascertain why high production values appear to be an important factor in why certain derivative uses might escape a takedown request, other factors being equal. Qualitative features of parodic treatment (such as the extent of transformation, and if what was taken from the original was necessary) are commonly considered in legal determinations of infringement.[46] The empirical findings suggest that this is also important in commercial practice.

61 A second finding of our study is that rightholders do not appear to be concerned with the expressive content of parodies, even when they explicitly target or criticize the original artist or work. This contradicts the expectation, based on published opposition by rightholder groups, that widespread parody threatens the integrity of works and therefore the moral rights of creators. In our sample, the "severity" of a parody significantly reduced the risk of a copyright takedown. The outcome suggests

45 *Supra* footnote 8: D. Mendis and M. Kretschmer (2013), p. 19.

46 *Supra* footnote 10: Heald (2014); Seng (2015); Urban, Karaganis, Schofield (2016).

that rightholders are more concerned with direct copying and with commercial licensing than with artists' moral rights.

62 The results obtained in this study suggest potential directions for future research. We have presented data on takedowns and rightholder behavior for a limited sample of internet content. While our results invite comparisons with other studies of notice-and-takedown,<?> in fact user-generated parody videos are a unique form of expression subject to dynamics that may be different in other domains where rightholders seek to protect their work from direct infringement. Comparative research might examine other communities where consumers appropriate commercial work to generate new expressions, for example fan fiction or machinima creator communities.

63 The UK eventually introduced a new fair dealing exception for the purposes of parody, caricature and pastiche with effect from 1 October 2014. In its technical review of draft legislation, the Intellectual Property Office outlined its rationale, stating, "adopting this exception will give people in the UK's creative industries greater freedom to use others' works for parody purposes. Drafting this as a fair dealing exception [...] is intended to allow creators to make minor uses of other people's copyright material for the purposes of parody, caricature or pastiche, without first asking for permission."<?> Because our original data were collected in 2012 and had already undergone takedown effects before the introduction of the new legislation, we are unable to examine effects of the UK exception on takedown rate. The effect of policy change on right holder behavior is a potential direction for future research.

64 This study provides the first empirical analysis of YouTube takedown behavior combining information about content as well as stated policy of rightholders. The central finding is that rightholders appear to make complex choices that are assisted by automatic detection mechanisms, with little concern for the artistic integrity of the creative works they represent. The significant difference between musical genres suggests that rightholders, even in the same medium, behave quite differently from their peers. Further empirical research of tradeoffs between enforcement, innovation, and freedom of expression in online platforms is urgently needed. Our study maps a new methodological path how to do this.

Annex 1: Descriptive statistics for main variables

	Min	Max	Mean	Std. Deviation
Taken down (censored event)	0	1	.3469	.47612
Indicator of sales rank	0	.99	.6043	.20556
Number of views at time of January 2012	1	26,856,003	130,543.547	1,064,833.044
Monetized dummy	0	1	.5856	.49274
Production values (1-5 scale)	1	5	3.0240	.99726
Highest production dummy	0	1	.3121	.46347
High production dummy	0	1	.2478	.43187
Average production dummy	0	1	.4041	.49086
Low production dummy	0	1	.2838	.45095
Parody type: target dummy	0	1	.3484	.47659
Parody type: weapon dummy	0	1	.3065	.46116
Parody type: mislabeled dummy	0	1	.3451	.47554
Highest severity of critique dummy	0	1	.0152	.12248
Music genre: pop dummy	0	1	.4448	.49708
Music genre: hip hop dummy	0	1	.3121	.46349
Music genre: rock dummy	0	1	.1648	.37107
Music genre: electro dummy	0	1	.0783	.26872
Territory: USA dummy	0	1	.7504	.43289
Territory: UK dummy	0	1	.1838	.38742
Major label dummy	0	1	.8124	.39050
Copied sound recording dummy	0	1	.77	.420
Copied video recording dummy	0	1	.01	.107
Parodist appear in video dummy	0	1	.7847	.41111
Female dummy	0	1	.1203	.32541
Time (Months)	13.00	72.00	59.904	17.507

Annex 2: Cox proportional hazards analysis of the effects of video features on takedown rate

	(1)	(2)	(3)	(4)	(5)
Model: covariates	Commercial	Cultural	Moral/ Artistic	Behavior	All factors
Sales rank	1.070 (.736–1.554)				.989 (.673–1.452)
Monetized dummy	1.209* (1.209–1.420)				1.135 (.962–1.340)
Log views	.854*** (.824–.885)				.864*** (.832–.898)
Production values	.901* (0.826–.984)				.898** (.819–.985)
Rock dummy		.600*** (.465–.774)			.560*** (.434–.724)
Electro dummy		1.143 (.859–1.227)			1.058 (.787–1.422)
Hip hop dummy		1.027 (.859–1.227)			1.013 (.861–1.236)
USA dummy		.640** (.471-.868)			.724*** (.604-.867)
UK dummy		1.010 (.720–1.416)			
Major label		.954 (.782–1.163)			.891 (.729–1.090)
Weapon dummy			.613*** (.412–.809)		.908 (.732–1.127)
Target dummy			.776** (.672–.994)		
Severity of critique				.368* (.137–.988)	.504 (.186–1.364)
Mislabeled dummy				1.425*** (1.213–1.674)	1.096 (.908–1.322)
Parodist appears				.976 (.804–1.185)	1.031 (.844–1.259)
Parodist female				1.162 (.917–1.472)	1.028 (.809–1.307)
Copied sound rec				1.472*** (1.202–1.802)	1.237** (1.004–1.523)
Copied video rec				1.761* (.938–3.307)	1.212 (.638–2.302)
Observa-tions	1,839	1,839	1,839	1,839	1,839
Model AIC	8872.71	8959.53	8971.49	8809.66	8691.23

*Notes: Values are exp(β) with 95% confidence intervals (lower-upper) in parentheses. Significance levels: *** p<0.01, ** p<0.05, * p<0.1*

Data-Related Aspects of the Digital Content Directive

by **Axel Metzger, Zohar Efroni, Lena Mischau and Jakob Metzger***

Research Group 4 ("Data as a means of payment") at the Weizenbaum Institute for the Networked Society – The German Internet Institute.

Keywords: Data as counter-performance; Data portability; Conformity; Embedded digital content and services; Digital Content Directive; General Data Protection Regulation (GDPR); Personal Data; Contract law; Consumer protection law

© 2018 Axel Metzger, Zohar Efroni, Lena Mischau and Jakob Metzger

Everybody may disseminate this article by electronic means and make it available for download under the terms and conditions of the Digital Peer Publishing Licence (DPPL). A copy of the license text may be obtained at http://nbn-resolving.de/urn:nbn:de:0009-dppl-v3-en8.

Recommended citation: Axel Metzger, Zohar Efroni, Lena Mischau and Jakob Metzger, Data-Related Aspects of the Digital Content Directive, 9 (2018) JIPITEC 90 para 1.

A. Executive Summary

1 The legislative initiative of harmonising certain aspects of contracts for the supply of digital content and services across the EU via a specific directive (DCD) is certainly a welcome and necessary one. While examining scenarios in which consumers provide data (as opposed to money) in exchange for such content or services, **it is important that the concept and ideally the specific wording of "data as counter-performance" is preserved** in the language of the directive, and that the directive covers both personal and any other data in this context. The directive should further apply to data irrespective of the question whether the consumer provides them actively or passively.

2 It is of crucial importance to establish a harmonised level of consumer protection for embedded digital content and services by covering the digital element of smart goods. The existing **differentiations between stand-alone and embedded digital content / services at the scope level should be removed.** Specific rules for embedded digital content /services should be drafted and applied only when absolutely necessary. In addition, the consumer protection implications arising from multi-party scenarios in the context of supplying smart goods must be more intensively investigated and expressly addressed in the final text of the directive.

3 On the issue of **portability of personal data**, this matter **should be governed exclusively by the GDPR**. Regarding **user-generated content (UGC)** that is not personal information, the **portability of such content should not be undermined by too broadly defined exceptions**. The right to retrieve such content should only be excluded if it cannot be made available without disproportionate and unreasonable effort. Traders should have a clear duty

to apply state-of-the-art technology to guarantee that consumers' UGC can be extracted separately, and that consumers' right to retrieve UGC should apply both against the trader and against any third party that stores and/or processes the content.

4 A **harmonised level of consumer protection** under the directive **in the context of conformity should principally apply in an equal manner to consumers who provide data as counter-performance and paying consumers alike**. Objective conformity requirements play an important role within the harmonised consumer protection scheme, and the type of counter-performance (data or price) should not result in lower requirements in the case of data as counter-performance contracts. However, the application of data protection law to some situations that are commercial in nature (such as the right to termination) marks the limits of the non-discrimination principle in favour of consumers who extend their personal data in exchange for commercial offers. The directive should not intentionally inhibit the ability of domestic contract laws to provide remedies to traders in the appropriate case and to the extent that such remedies are in line with the harmonised data protection law.

B. Introduction

I. The Weizenbaum Institute for the Networked Society

5 The Weizenbaum Institute[1] investigates the current changes in all aspects of society occurring in response to digitalisation. Its goals are to develop a comprehensive understanding of these changes based on rigorous academic analysis and to offer informed strategies to address them at a political and economic level.

6 The Weizenbaum Institute is funded by the Federal Ministry of Education and Research. The consortium is coordinated by the Berlin Social Science Center ("WZB") and includes the four Berlin universities – Freie Universität Berlin, Humboldt-Universität zu Berlin, Technische Universität Berlin, Universität der Künste Berlin – as well as the Universität Potsdam and the Fraunhofer Institute for Open Communication Systems ("FOKUS").

7 The Berlin-Brandenburg Consortium focuses on the interaction of the social sciences, economics and law with design research and computer science. Interdisciplinary basic research and the exploration of concrete solutions in practice-based labs are combined with knowledge transfer into politics, business, and society. The conceptual design of the Institute aims to achieve scientific excellence with a nationwide and international impact, as well as networking with cooperation partners from civil society, business, politics, and the media.

II. Purpose and Methodology

8 Our mission is to highlight a number of important issues within the larger debate around the Digital Content Directive ("DCD")[2] and its legislative process. We focus for the most part on situations where consumers, in exchange for digital content / services, provide data and not money. Within our selected topics, we bring forward several recommendations concerning the preferred approaches with the aim of contributing to the continuing discussions they have evoked. As the legislative process is reaching its most critical stages, we present solutions that will hopefully be taken into consideration while the EU trilogue participants hammer out the final text of the DCD.

9 The structure of this position paper is as follows: first, we present the approaches of the European Commission ("COM"), the Council of the European Union ("Council"), and the European Parliament ("EP") as reflected in their respective proposals in the form of a comparative table juxtaposing the relevant texts one next to the other. Then, for each topic, we add comments concluded by concrete recommendations.

10 Among the topics that are sought to be regulated under the directive, we focus on the principal question of (personal) data as counter-performance in the context of business-to-consumer contracts as well as on related issues of embedded digital content, portability rules, and conformity requirements.

* Prof. Dr. iur. Axel Metzger, LL.M. (Harvard), Founding Director, Weizenbaum Institute for the Networked Society, Berlin, Professor of Law, Humboldt-Universität zu Berlin; Dr. iur. Zohar Efroni, LL.M. (Cardozo), Research Project Lead, Weizenbaum Institute for the Networked Society and Humboldt-Universität zu Berlin; Lena Mischau, Research Associate, Weizenbaum Institute for the Networked Society and Humboldt-Universität zu Berlin; Jakob Metzger, Research Associate, Weizenbaum Institute for the Networked Society and Humboldt-Universität zu Berlin.

1 This position paper represents exclusively the opinion of its authors, who are members of the Research Group "Data as a means of payment" at the Weizenbaum Institute.

2 COM, 'Proposal for a Directive of the European Parliament and of the Council on certain aspects concerning contracts for the supply of digital content', COM(2015) 634 final, 2015/0287 (COD), 09.12.2015 (hereinafter referred to as "DCD-COM").

C. Data as Counter-Performance

I. Relevant Provisions

European Commission (09.12.2015)[I]	Council of the European Union (01.06.2017)[II]	European Parliament (27.11.2017)[III]
Recital (13)	Footnote 15[IV]	Recital (13) (Amendment 19)[V]
In the digital economy, information about individuals is often and increasingly seen by market participants as having a value comparable to money. Digital content is often supplied not in exchange for a price but against counter-performance other than money i.e. by giving access to personal data or other data. Those specific business models apply in different forms in a considerable part of the market. Introducing a differentiation depending on the nature of the counter-performance would discriminate between different business models; it would provide an unjustified incentive for businesses to move towards offering digital content against data. A level playing field should be ensured. In addition, defects of the performance features of the digital content supplied against counter-performance other than money may have an impact on the economic interests of consumers. Therefore the applicability of the rules of this Directive should not depend on whether a price is paid for the specific digital content in question.	An explanation along the following lines will be added in the recitals: "In the digital economy, digital content is often supplied without the payment of a price and suppliers use the consumer's personal data they have access to in the context of the supply of the digital content or digital service. Those specific business models apply in different forms in a considerable part of the market. A level playing field should be ensured. This Directive should apply to contracts where the supplier supplies or undertakes to supply digital content or a digital service to the consumer. Member States should remain free to determine whether the requirements for the existence of a contract under national law are fulfilled. The Directive should not apply where the consumer does not pay or does not undertake to pay a price and does not provide personal data to the supplier. [...]	In the digital economy, information about individuals is often and increasingly seen by market participants as having a value comparable to money. Digital content ***and digital services are*** often supplied not in exchange for a price but against ***data,*** i.e. by giving access to personal data or other data. Those specific business models apply in different forms in a considerable part of the market. Introducing a differentiation depending on the nature of the counter-performance would discriminate between different business models, ***which provides*** an unjustified incentive for businesses to move towards offering digital content ***or digital services*** against data. In addition, defects of the performance features of the digital content ***or digital service*** supplied against ***data as counter-performance*** may have an impact on the economic interests of consumers. ***In order to ensure a level playing-field,*** the applicability of the rules of this Directive should not depend on whether a price is paid for the specific digital content ***or digital service*** in question.
Recital (14)	Footnote 15	Recital 14
As regards digital content supplied not in exchange for a price but against counter-performance other than money, this Directive should apply only to contracts where the supplier requests and the consumer actively provides data, such as name and e-mail address or photos, directly or indirectly to the supplier for example through individual registration or on the basis of a contract which allows access to consumers' photos. [...] This Directive should [...] not apply to situations where the supplier collects information, including personal data, such as the IP address, or other automatically generated information such as information collected and transmitted by a cookie, without the consumer actively supplying it, even if the consumer accepts the cookie. [...]	"[...] This Directive should not apply to situations where the supplier only collects metadata, the IP address or other automatically generated information such as information collected and transmitted by cookies, except where this is considered as a contract by national law. [...] However, Member States should remain free to extend the application of the rules of this Directive to such situations or to otherwise regulate such situations which are excluded from the scope of this Directive."	As regards digital content ***and digital services*** supplied not in exchange for a price but ***when personal data is provided***, this Directive should apply to contracts where the ***trader*** requests and the consumer provides *personal* data, *as well as where the trader collects* ***personal data. It would include, for example, the*** name and e-mail address or photos, ***provided*** directly or indirectly to the ***trader***, for example through individual registration or on the basis of a contract which allows access to consumers' photos, ***or personal data collected by the trader, such as the IP address.*** [...]
Article 3 – Scope	Article 3 – Scop	Article 3
(1) This Directive shall apply to any contract where the trader supplies digital content to the consumer or undertakes to do so and, in exchange, a price is to be paid or the consumer actively provides counter-performance other than money in the form of personal data or any other data.	(1) This Directive shall apply to any contract where the supplier supplies or undertakes to supply digital content or a digital service to the consumer (...). It shall not apply (...) to the supply of digital content or a digital service for which the consumer does not pay or undertake to pay a price and does not provide or undertake to provide personal data to the supplier. [...]	(1) This Directive shall apply to any contract where the ***trader*** supplies or undertakes to ***supply digital content or a digital service to*** the consumer ***whether through the payment of a price or under the condition that*** personal data ***is provided by the consumer or collected by the trader or a third party in the interest of the trader***.

[I] DCD-COM (n 2).

[II] Council, 'Proposal for a Directive of the European Parliament and of the Council on certain aspects concerning contracts for the supply of digital content (First reading) – General approach', 9901/17 ADD 1, 2015/0287 (COD), 01.06.2017 (hereinafter referred to as "DCD-Council"). Footnote(s) in the DCD-Council text omitted.

[III] EP, 'Report on the proposal for a directive of the European Parliament and of the Council on certain aspects concerning contracts for the supply of digital content (COM(2015)0634 – C8-0394/2015 – 2015/0287(COD))', A8-0375/2017, 27.11.2017; (hereinafter referred to as "DCD-EP").

[IV] Footnote 15 is part of Article 3(1) DCD-Council.

[V] At the same time, EP states in Recital 13 – Amendment 20: "In the digital economy, information about individuals is often and increasingly seen by market participants as having a value. ***Specific business models have developed in which traders supply*** digital content ***or a digital service and the consumer is required to provide or give*** access to personal data. Those specific business models ***already*** apply in different forms in a considerable part of the market. ***This Directive does not intend to decide whether such contracts should be allowed or not. In addition, it leaves to national law the question of validity of contracts for the supply of digital content or a digital service where personal data are provided or accessed. This Directive should, in no way, give the impression that it legitimises or encourages a practice based on monetisation of personal data, as personal data cannot be compared to a price, and therefore cannot be considered as a commodity. However,*** introducing a differentiation ***in the rules applying to monetary and non-monetary transactions*** would provide an unjustified incentive for businesses to ***favour the supply of*** digital content ***or digital services on condition that personal data is provided***. In addition, defects of the performance features of the digital content ***or digital service*** supplied ***when no price is paid might*** have an impact on the economic interests of consumers. ***With a view to ensuring a levelplaying-field and a high level of consumer protection,*** the applicability of the rules of this Directive should not depend on whether a price is paid for the specific digital content ***or digital service*** in question" (Emphasis in original).

II. Comments

11 Some of the key questions the Digital Content Directive (DCD) prompts already begin with its intended scope. The following discussion focuses on three of those questions: namely, the inquiry whether data should be considered counter-performance in the first place (1); whether treating data as counter-performance should apply to personal data only, or rather, also to any other data (2); and whether the scope of the DCD should cover actively provided data only or also data provided passively (3). Currently, the positions of the European Commission, the Council of the European Union, and the European Parliament[3] on these essential questions differ quite significantly.

3 Committees responsible: Committee on the Internal Market and Consumer Protection (IMCO) and Committee on Legal Affairs (JURI).

1. Data as counter-performance

12 Article 3(1) and Recital 13, 14 DCD-COM clearly state that counter-performance can be provided not only in the form of money, but also in the form of personal data or any other data. Notably, the General Approach document of the Council does not mention the notion of "counter-performance" by name. It seems that the Council prefers to avoid using this terminology by stating instead that the DCD "shall not apply [...] to the supply of digital content [...] for which the consumer does not pay [...] a price and does not provide [...] personal data".[4] The EP shows a similar tendency by recommending to remove the phrase "counter-performance" from Article 3(1). Its amendment to Article 3(1) stipulates that the DCD "shall apply to any contract where the trader supplies [...] digital content [...] under the condition that personal data is provided or collected [...]".[5]

13 The debate whether data should be considered "counter-performance" or not reflects the tension between two regulative approaches to the intersection between markets, data protection and consumer protection; namely, recognizing data as counter-performance in the context of the DCD and thereby guaranteeing a high factual level of consumer protection might signal to market participants the acceptance of commercialisation of personal data. Alternatively, ignoring that type of counter-performance may signal a rejection of such commercialisation, but this would come at the price of lowering the factual level of consumer protection.

14 There are no clear answers to the general question regarding how far the legal system should "protect consumers from themselves" without risking becoming overly paternalistic.[6] At the same time, there seems to be a consensus around the recognition that "data [provided] against digital content" is today a prevalent business model that cannot be ignored.[7] Accordingly, the COM and EP agree that in the digital economy, information about individuals is being increasingly seen by market participants as having a value comparable to money.[8] Even the European Data Protection Supervisor (EDPS) in principle welcomes the intention of regulatory

4 Article 3(1)(2) DCD-Council.

5 Article 3(1) DCD-EP. However, the EP does mention "data as counter-performance" in its Amendment 19 (regarding Recital 13).

6 *Cf.* Peter Bräutigam, 'Das Nutzungsverhältnis bei sozialen Netzwerken, Zivilrechtlicher Austausch von IT-Leistung gegen personenbezogene Daten' (2012) MultiMedia und Recht 635, 637.

7 Both the Council and EP agree with the COM that those specific business models apply in different forms in a considerable part of the market. *See* Recital 13 DCD-COM/-Council/-EP – Amendment 19, 20.

8 *See* Recital 13 DCD-COM, Recital 13 DCD-EP – Amendment 19.

approaches to protecting consumers in the digital environment, including those who provide personal data in exchange for ostensibly "free services."[9]

15 In fact, excluding data as counter-performance (hereinafter "DACP") situations from the DCD would lead to discrimination between DACP-consumers and price-paying consumers. It is highly questionable whether discrimination solely on this basis across the board is justifiable. Obviously, DACP-consumers do not obtain the digital content "for free" and therefore there is no reason to assume that they deserve a lower level of protection. Their data has a substantial economic value to traders, and their economic interests are surely at stake when the trader deviates from its contractual obligations irrespective of the nature or their counter-performance.[10]

16 Recital 13 DCD-COM makes a double assumption according to which (1) differentiation would boost DACP business models, and (2) incentivising DACP business models in this way would be unjustified and should be avoided. These assumptions call for further scrutiny. Strictly speaking, excluding DACP-transactions from the scope of the DCD would mean a lack of harmonisation in this area, and by extension, result in any type and level of consumer protection a given Member State decides to grant. The DCD would thereby forgo an important opportunity to cover this aspect of digital markets. Increased fragmentation among domestic laws in their respective approaches of DACP-transactions would clearly undercut the harmonisation agenda of the DCD.

17 Apart from this, the emphasis of Recital 13 on discrimination between business models appears somewhat misplaced: The main instrument with which the DCD seeks to achieve the ultimate goal of fostering the growth of the Digital Single Market is not by equalizing incentives to pursue various business models, but more likely by harmonizing the level of consumer protection (or at least, some aspects thereof) and thus significantly increasing legal certainty across the European Union. The discrimination that needs to be avoided is therefore not so much between different business models as it is between different consumer groups. As also suggested by the COM and EP in Recital 13, an important consideration in this context is the impact on the economic interests of consumers: it is the discrimination between classes of consumers that generally should be avoided.

18 In addition, the human-rights aspect of personal data and the capacity of personal data to serve as counter-performance are not mutually exclusive.[11] In other legal disciplines it is well established that personality-related rights (such as authors' rights or publicity rights) can simultaneously have a monetary dimension, which their holders are free to realise.[12] Such duality can equally apply to the interface between data as reflecting a personal right (e.g., under the GDPR[13]) and data as a commodity (e.g., under the DCD).[14] The direct reference from the DCD to the GDPR as having the regulative priority in all data protection-related matters effectively leaves the latter unaffected.[15] No erosion in the status and operation of data protection law is to be feared if the DCD merely targets the commercial facets of a market reality that data protection law cannot wipe away.

19 An impact in the opposite direction, namely, a foray of data protection law into the domain of contract law, should also be considered at this juncture.[16] The right to withdraw consent to the processing of personal data (e.g., as laid down in Article 7(3) GDPR) does not necessarily negate the possibility of a contract over personal data. The conclusion

9 *See* EDPS, 'Opinion 4/2017 on the Proposal for a Directive on certain aspects concerning contracts for the supply of digital content' (EDPS, 14 March 2017) <https://edps.europa.eu/sites/edp/files/publication/17-03-14_opinion_digital_content_en.pdf> accessed 23 March 2018, p. 7.

10 *See* Recital 13 DCD-COM/-EP – Amendment 19, 20.

11 See, e.g., the fundamental right to the protection of personal data as enshrined in Article 8 Charter of fundamental rights of the European Union ("EU Charta"), Article 16 Treaty on the Functioning of the European Union ("TFEU"); different view, EDPS (n 9) 7: "*However, personal data cannot be compared to a price, or money. Personal information is related to a fundamental right and cannot be considered as a commodity.* [...] *There might well be a market for personal data, just like there is, tragically, a market for live human organs, but that does not mean that we can or should give that market the blessing of legislation*".

12 *Cf.* Peter Bräutigam (n 6) 639; Carmen Langhanke/Martin Schmidt-Kessel, 'Consumer Data as Consideration' (2015) Journal of European Consumer and Market Law 218, 219; Artur-Axel Wandtke, 'Ökonomischer Wert von persönlichen Daten, Diskussion des „Warencharakters" von Daten aus persönlichkeits- und urheberrechtlicher Sicht' (2017) MultiMedia und Recht 6, 9.

13 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 ("GDPR").

14 *Cf.* Carmen Langhanke/Martin Schmidt-Kessel (n 12) 219 f. (offering a similar observation: "*consumer protection takes place at two layers, the layer of data protection and the layer of contract law*").

15 *Cf.* Martin Schmidt-Kessel et. al., 'Die Richtlinienvorschläge der Kommission zu Digitalen Inhalten und Online-Handel – Teil 2' (2016) Zeitschrift für das Privatrecht der Europäischen Union, Fokus, 54, 59; *different view, e.g.:* Niko Härting, 'Digital Goods und Datenschutz – Daten sparen oder monetarisieren? Die Reichweite des vom DinhRL-E erfassten Geschäftsmodells' (2016) Computer und Recht 735, 738, 740.

16 *See, e.g.,* Andreas Sattler, 'Personenbezogene Daten als Leistungsgegenstand' (2017) JuristenZeitung 1036, 1038, 1041 (offering a critical perspective on this point).

of a contract remains subject to national law.[17] In addition, the withdrawal of consent does not affect the lawfulness of processing based on consent before its withdrawal (Article 7(3) GDPR).

20 The conclusion must be that the European legislator cannot turn a blind eye to DACP-transactions within the general project of promoting the Digital Single Market despite the potential tension with data protection law. The proposed Article 3(8) DCD-Council already points in a similar direction. That said, certain clarifications, for instance as suggested by the Council (in footnote 15) or by the EP (in Recital 13 – Amendment 20), may be useful in explaining the interplay between these two bodies of law. Avoiding the term "counter-performance" or any comparable terminology from the realm of contract law contributes nothing to achieving the goals of the DCD or serving any other regulative purpose.[18]

2. Personal or any other data

21 Whereas the COM relates to "personal data or any other data" as potentially replacing payment of price, both the Council and EP advocate for limiting the language to "personal data" only. The General Statement (Council) and the Report (EP) do not explain in detail the rationale for excluding "other data" from the scope of the DCD. A possible explanation is the wish to avoid the additional complexity resulting from the necessity to differentiate between the two types of data in the text of the directive. Another possible reason could be the underlying idea that specific regulation addressing non-personal data is less necessary provided that consumers, for the most part and in light of the broad concept of "personal data" under the GDPR, are not likely to provide non-personal data to traders.

22 The latter assumption is weakened if considered against available methods and technologies to anonymise data once it reaches the trader and further down the value chain. But even assuming that data preserves its original identity as personal or non-personal after entering the commercial cycle, differentiation at the scope level would immediately introduce the (sometimes nontrivial) task of determining which category the data provided by the consumer belongs to. Once this has been done, the assumption about the negligent importance of non-personal data in DACP-scenarios would have to be revisited under future business models that might increase the importance of such scenarios.

23 We therefore submit that including both personal and non-personal data would better serve the interests of efficiency, legal certainty and consumer protection. To the extent that the two data categories receive different treatment under the DCD in order to prevent friction with data protection law, a direct reference to the definition of personal data in the GDPR appears advisable. Once included, non-personal data should be controlled by the DCD norms that protect consumers against continued use of their data after termination.[19] The GDPR will continue to apply directly on such matters with regards to personal information.[20]

24 In the context of the DCD referring to GDPR norms where data protection law is implicated due to the nature of data as personal data, a general word of caution is warranted: reference to specific provisions in the GDPR should not necessarily mention provision numbers, but rather the intended data protection principles in order to prevent cross-reference errors in case the legislative texts are to be amended or replaced in the future.[21] Furthermore, a specific reference to the GDPR in one occasion should not open the door to *argumentum e contrario* where the GDPR should apply but is not mentioned in the text of the DCD. It is therefore advisable to explain (possibly in the Recitals) the relationship between the DCD and the GDPR and specifically exclude *e contrario* interpretations.

3. Actively and passively provided data

25 According to the COM, the DCD applies only to data that is actively provided by the consumer, whereas data collected by the trader that is not actively provided, such as the IP address or even data collected after the acceptance of a cookie, do not fall under its scope.[22] The Council, by comparison, would introduce a minimum harmonization standard, allowing member states to also extend the

17 *See* Axel Metzger, 'Data as Counter-Performance, What Rights and Duties do Parties Have?' (2017) Journal of Intellectual Property, Information Technology and Electronic Commerce Law 2, 3.

18 However, the variation of content/services "against data" (as proposed in Recital 13 DCD-EP – Amendment 19) appears to be an acceptable alternative.

19 *See* Article 13(2)(b), Article 15(2)(b), Article 16(4)(a) DCD-COM.

20 *Cf.* Article 16(3) DCD-Council, Article 15(2) DCD-EP, Article 13a(2) DCD-Council/EP.

21 For example, in case of termination of the contract, a reference to the right to erasure (or, the "right to be forgotten") should be made in Article 13 DCD. This right is currently stipulated in Article 17 GDPR. Such reference would go beyond the suggestion of Article 13a(2) DCD-Council/-EP which provides for a general reference to the GDPR only. At the same time, there is no necessity to repeat or rephrase provisions from the GDPR within the DCD.

22 Recital 14, Article 3(1) DCD-COM.

application of the DCD to passively provided data.[23] Shifting to the opposite extreme, the EP would apply the DCD irrespective of the question whether data was actively "provided by the consumer or collected by the trader or a third party in the interest of the trader"[24] in order to avoid loopholes.[25]

26 As pointed out by the European Parliament's Committee on Civil Liberties, Justice and Home Affairs (LIBE), limiting the scope to actively provided data would create a perverse incentive for traders to not ask for the consumer's consent.[26] Simply assuming that data protection law will operate to protect a consumer whose data was passively collected in a manner that adversely affects his or her interests does not suffice. In case of passive collection that is unlawful under the GDPR, the protections of the DCD should apply all the more.

27 Excluding scenarios from the scope of the DCD where the counter-performance consists of passively provided data would in fact be counterproductive in terms of consumer protection. In case of non-personal data, neither the DCD nor the GDPR would apply. But also in case of personal data provided passively, where the GDPR does apply, the DCD can provide an additional layer of protection, e.g., a right to damages (Article 14 DCD-COM) if the digital content or service is not in conformity with the contract.

28 Moreover, it should be noted that the criteria set out in Recital 14 DCD-COM to distinguish between actively or passively provided data call for further clarification. This is especially true for the given example of cookies. There is no reason for consumers whose data is collected by the means of cookies to be less protected than consumers who actively consent to the collection of essentially the same data.[27] Passively collected data is neither less valuable than actively collected data, nor is it marginal in scope or importance.[28] In addition, the economic interests of both types of consumers are affected by the usage of the data in the same way. The conclusion is that discrimination between consumer groups on such basis would lack any plausible justification, and following a minimum harmonisation approach here would not suffice.

III. Recommendations

1. **The concept of counter-performance should be maintained. The wording "counter-performance" introduced by DCD-COM is preferable to the solutions proposed by the Council and EP in Article 3 (1) DCD. Alternatively, the wording of Recital 13 DCD-EP – Amendment 19, referring to content or services provided "against data" could be an acceptable alternative.**

2. **The DCD should apply to both personal and any other data. The phrase "or any other data" should therefore be maintained. Alternatively, Article 3(1) DCD should use the term "data" without differentiating between personal and any other data. In this case, Article 2 DCD and the relevant Recitals should clarify that the term "data" covers both personal and any other data.**

3. **The DCD should apply to data irrespective of the question whether it is provided actively or passively by the consumer. The term "actively" in Article 3(1) and Recital 14 DCD-COM should hence be deleted and the term "or collected by the trader or a third party in the interest of the trader" as stated in Article 3(1) DCD-EP should be maintained.**

23 *See* Article 3(1) Footnote 15 DCD-Council: "*However, Member States should remain free to extend the application of the rules of this Directive to such situations or to otherwise regulate such situations which are excluded from the scope of this Directive* [...]".

24 *See* Recital 14, Article 3(1) DCD-EP.

25 *See* Explanatory Statement within DCD-EP p. 90.

26 *See* Opinion of LIBE within DCD-EP p. 94.

27 *Cf.* European Law Institute (ELI), 'Statement on the European Commission's proposed directive on the supply of digital content to consumers' (ELI, 2016) <https://www.europeanlawinstitute.eu/fileadmin/user_upload/p_eli/Publications/ELI_Statement_on_DCD.pdf> accessed 23 March 2018, p. 15 f.; Axel Metzger (n 17) 3.

28 In fact, especially cookies in combination with applications such as Google Analytics are used to collect personal data and to create economic value on a large scale; *see* Gerald Spindler, 'Verträge über digitale Inhalte – Anwendungsbereich und Ansätze, Vorschlag der EU-Kommission zu einer Richtlinie über Verträge zur Bereitstellung digitaler Inhalte' (2016) MultiMedia und Recht 147, 149, with further references.

D. Embedded digital content and services

I. Relevant Provisions

European Commission (09.12.2015)	Council of the European Union (01.06.2017)	European Parliament (27.11.2017)
Recital 11	Article 2(12)[I]	Article 2(1)(1b)[II]
(...) this Directive should not apply to digital content which is embedded in goods in such a way that it operates as an integral part of the goods and its functions are subordinate to the main functionalities of the goods.	**'embedded digital content' means digital content present in a good, whose absence would render the good inoperable or would prevent the good from performing its main functions, irrespective of whether that digital content was pre-installed at the moment of the conclusion of the contract relating to the good or according to that contract installed subsequently.**	***'embedded digital content or digital service' means digital content or a digital service pre-installed in a good;***
Article 3(3)	Article 3(3)	Article 3(3)
With the exception of Articles 5 and 11, this Directive shall apply to any durable medium incorporating digital content where the durable medium has been used exclusively as carrier of digital content.	With the exception of Articles 5 and 11, this Directive shall apply **also** to any **tangible** medium **which incorporates** digital content **in such a way that the tangible medium serves** exclusively as carrier of digital content. Article 3(3a) **This Directive shall not apply to embedded digital content.**	With the exception of Articles 5 and 11, this Directive shall apply to ***embedded*** digital content ***or embedded digital services. Unless otherwise provided, references to*** digital content ***or digital services in this Directive also cover embedded digital content or embedded digital services. As regards goods with embedded digital content or embedded digital services, the trader shall be liable under this Directive to the consumer for meeting his obligations only in respect of the embedded digital content or digital service. The rules of this Directive are without prejudice to the protection granted to consumers by the applicable Union law with respect to other elements of such goods.*** Article 9(1) The ***trader*** shall be liable to the consumer for: [...] (c) ***any lack of conformity with the contract of embedded digital content or an embedded digital service which exists at the time of delivery of the goods in which the digital content or digital service is embedded and becomes apparent within two years from the time of delivery.*** Article 10(1) The burden of ***proving that a lack of conformity existed at the time specified in Article 9 shall be on the trader, when a lack of*** conformity with the contract ***becomes apparent during the following periods:*** [...] ***(b) within one year of the date of delivery of the embedded digital content or digital service;*** [...] Article 13b ***1. After termination of the contract*** [...] ***2. In the case of embedded digital content or an embedded digital service, the consumer shall, at the request of the trader, return, at the trader's expense, the good*** [...]

[I] Footnotes in the DCD-Council text omitted. Emphasis in original.

[II] Emphasis in original.

II. Comments

29 The proposal of the EP to include embedded digital content and services (EDCS) within the scope of the DCD is a welcome development. Having considered this a step in the right direction, we would recommend following through by removing the unnecessary differentiation between stand-alone and embedded digital content or services.

1. The importance of covering EDCS in general

30 Considering the regulative framework of the DCD alongside the current proposal for a directive of the European Parliament and of the Council on certain aspects concerning contracts for the online and other distance sales of goods (COM(2015) 635 final, hereinafter "OSD") reveals a notable misconception: the OSD is often mentioned as providing a "safety net" to consumers due to its capacity to close loopholes in the DCD protection scheme. It must be mentioned in this context that the OSD only applies if the physical good is bought and paid for with money. The DCD is therefore indispensable to rental or lending situations.

31 The same holds true for cases in which the good is given away for free or in exchange for data. With the sinking cost of electronic gadgets and the growing market for Big Data and targeted marketing, companies have an increased incentive to distribute consumer electronics without charge. Already now, some consumer electronics are sold at the manufacturing price or less. The main purpose of such "giveaways" is the collection of data generated through use and monetising that data. Under the current OSD draft, this entire market segment is not covered (since the OSD does not cover data as counter-performance transactions), opening a regulatory gap that the DCD is capable of closing.

2. The DCD should cover EDCS

32 Even in cases where the OSD could serve as a safety net for consumers, its anticipated protection scheme remains insufficient as it does not cover crucial questions such as (security) updates or modifications that are necessary irrespective of whether the content is provided as a stand-alone product or embedded in a physical good. Many observers as well as the EP have already recognised the importance of including EDCS within the framework of the DCD.[29] Despite the difficult challenge of reasonably delineating the scope of the DCD and its coexistence alongside the OSD, leaving EDCS outside of the domain covered under the DCD would be a resounding mistake.

33 The main arguments for such inclusion are identical with the arguments for enacting the DCD in general: smart goods (or for the DCD in general digital content) are becoming ever more relevant, and they differ from conventional goods in ways that call for specific regulation. A harmonized set of rules in this area is essential in order to bolster consumer rights and increase legal certainty, which might hinder transactions and thus the development of a Digital Single Market. As already noted by others, a failure to cover the digital aspect of physical smart goods on an EU-level would lead to confusion, inconsistencies and a "serious gap in consumer protection".[30]

34 The difficulties in implementing this approach, however, are rooted in the regulative perspective of the DCD, which focuses on the type of good (digital content), as opposed to the OSD that is tied to the legal consequences intended by the parties (transfer of ownership). The DCD's stance of focusing on the type of goods seems to contrast civil codes' regulative matrix in some Member States (including Germany). The resulting problems in aligning the DCD with the OSD are therefore likely to trickle down to future efforts of implementing the two instruments in national laws.[31] However, assuming that the basic structure of both directives will be upheld in the final versions, we submit that including EDCS products under the DCD is crucial.

3. The current proposals for the implementation of EDCS are insufficient

35 The current proposals to include EDCS in the DCD share one shortcoming: these proposals would expressly include "embedded" digital content and services and separately define the term "embedded".[32]

29 *E.g.* ELI (n 27) 10 ff.; Christiane Wendehorst, 'Sale of goods and supply of digital content – two worlds apart? Why the law on sale of goods needs to respond better to the challenges of the digital age' (European Parliament, PE 556.928, 2016) <http://www.europarl.europa.eu/cmsdata/98774/pe%20556%20928%20EN_final.pdf> accessed 23 March 2018 p. 4 ff.; Explanatory Statement within DCD-EP p. 90.

30 ELI (n 27) 2.

31 Those implementation difficulties already lead to the proposal that the member states should be obliged to introduce the DCD not integrated within contractual law, but as a sui generis regime, see Vanessa Mak, 'The new proposal for harmonised rules on certain aspects concerning contracts for the supply of digital content' (European Parliament, PE 536.494, 2016) <http://www.epgencms.europarl.europa.eu/cmsdata/upload/a6bdaf0a-d4cf-4c30-a7e8-31f33c72c0a8/pe__536.494_en.pdf> accessed 23 March 2018, p. 13.

32 The preparation of the EP-Report has produced a wide

Unfortunately, this approach tends to raise more questions than it can solve.[33] For instance, the latest EP proposal attempts to identify "embedded" digital content or services based on them being "pre-installed in a good". Alas, the definition is ambiguous and potentially too narrow. Ambiguous, as it is not clear whether it covers content that remains on a cloud and is accessed through the good in the course of use. In addition, the term is not really fitting for digital services: usually, they are not "installed" on the device. In such cases, merely a client or interface might be pre-installed to allow access to the service. Would the DCD in such cases only apply to the client or also cover errors on the remote server?

36 The definition is also too narrow, as it leaves the door open for common business models to escape its application. If EDCS are to be included separately, the logic of the proposal would be that digital content or services delivered through a physical good were not to be covered per se, but only if they were "pre-installed". Especially where content quickly runs out-of-date (e.g., maps on a Navisat), devices are only delivered with a basic environment and physical parts, while the majority of content must be downloaded after delivery. Such content would hence not be covered – a loophole the trader could take advantage of, even where there is no objective reason to deliver the digital content subsequent to providing the good.

4. Solution: No differentiation between stand-alone and embedded content / services

37 To avoid such definition-based problems, we recommend to remove the differentiation between stand-alone and embedded digital content or services and instead to conceive EDCS as a special way of supplying digital content or services. There is no need for distinction at the scope level, as the two forms of supplying digital content do not differ to an extent that requires two sets of specific rules. In both cases, there are similar consumer interests and market challenges at stake. Additionally, in case differentiation would still make sense in a limited context, this can be done ad hoc within the relevant provision.

38 The challenge of distinguishing between physical products distributed with an embedded digital content or service (hence, subject to EDCS regulation) and the distribution of digital content or services that are merely embodied in a physical article or otherwise connected with it (hence potentially subject to mixed/linked contract regulation) is only expected to grow in the future.[34] For this reason, we recommend to extend the scope of the DCD to all digital content and services irrespective of the way in which they are delivered. To implement this understanding into the DCD, one could for example clarify the scope by changing the definition of "supply" and omit any definitions and exclusions or inclusions of EDCS (e.g. "'supply' means providing access to digital content or services or making digital content or services available isolated or within or in connection with physical goods").[35]

39 If, however, the EU legislature nonetheless chooses to follow the definition-based approach, we join the recommendations of the ELI advocating for the amendment of the notions of "embedded" and "ancillary" content or services.[36] To broaden the directive's scope, these terms should then be defined in such a way that they cover all digital content and services delivered within or in connection with a physical good in fulfilment of a contractual obligation.[37]

5. Recommended deletion of EDCS-specific provisions

40 That there is no basic need for differentiation becomes clear if one is to investigate the rules that have been proposed by the EP specifically for EDCS. Those EDCS-specific provisions set out in Article 3(3), 9(1)(c), 10(1)(b), 13b(2) DCD-EP are superfluous and should hence be deleted:

- Article 3(3) DCD-EP should not exclude Article 5 and Article 11 DCD-EP (duty to supply the digital content and remedies for the failure to supply) for EDCS. Although Article 18 of Directive 2011/83/EG (Consumer Rights Directive,

variety of proposals: the content / service should be considered embedded, if its functions are subordinate to the main functionalities; or if its absence would render the good inoperable / prevent it from performing its main functions; or if it could not be easily de-installed by the consumer.

33 *See* Martin Schmidt-Kessel, 'Stellungnahme zu den Richtlinienvorschlägen der Kommissionzum Online-Handel und zu Digitalen Inhalten' (Bundestag, 2016), <https://www.bundestag.de/blob/422258/c3ecca9b7286f38bda7e060f7b420c06/schmidt_kessel-data.pdf>, accessed 23 March 2018, p. 2 ff. (also skeptical about the split approach and the possibility to find a suitable definition).

34 *See* Christiane Wendehorst (n 29) 7 ff. (*on the problems of distinction).*

35 If the definition of "supply" is to be deleted as proposed by the EP, the clarification could be amended within the recitals.

36 ELI (n 27) 13.

37 The requirement "in fulfilment of a contractual obligation" would exclude free extras that are not contractually owed, such as pre-installed MP3-Songs delivered with a smartphone.

hereinafter "CRD") already covers those rights with regard to physical goods, Articles 5 and 11 DCD-EP should additionally apply to the digital part of the good. As Article 18 CRD applies only to sales contracts, Articles 5 and 11 DCD-EP could guarantee a level of harmonization for all other cases. But even for sales contracts, the need for a designated rule concerning the digital component of the good persists – e.g., if the embedded service is to be unlocked after the delivery of the good.

- Article 9(1)(c) DCD-EP (relevant point in time for evaluation of conformity) is tailor-made for sale contracts, in which the goods are handed over at one single occasion. It does not fit for embedded content or services provided over a period of time. Article 9(1)(b) DCD-EP in its current form is capable of also covering EDCS-contracts. An additional rule as stated under Article 9(1)(c) DCD-EP is superfluous and in its current wording too narrow.

- Since both the newly redrafted Article 8(3) OSD-COM(2017)[38] as applicable to physical goods, and Article 10(1)(a) DCD-EP for digital content and services set a two-year time limit for the reversal of the burden of proof for the lack of conformity, there is no need to set a different limit of one year for EDCS in Article 10(1)(b) DCD-EP. But even if the time-limit for physical goods should remain less than two years, it is not necessary to also lower the time-limit for EDCS and treat them differently than stand-alone digital content or services.

- Article 13b(2) DCD-EP (duty to return the good) is overly complex. Beyond that, it interferes with other EU and domestic regulations governing sales contracts and touches upon national core contract law by regulating the effects of linked contracts despite the express intention of the DCD to avoid such impacts.

6. The troubled interface with other EU-Regulations (especially the OSD)

41 Another critical point is the interrelation between the DCD and the OSD in cases of sales contracts. In many instances, the separation is fairly easy to make, since most of the rules designed for digital goods logically do not apply to physical goods (a security update for the wristband of a smart watch would make no sense.) At the same time, serious issues begin to surface in the broader context of conformity with the contract and remedies for non-conformity as there are different regulations intended for conformity criteria, relevant time periods, or burden of proof. The EP's co-rapporteurs declared their intent to work together with the rapporteur responsible for the OSD in an attempt to align conformity criteria and thus minimize the impacts of the split approach.[39] However, already the on-going discussion about subjective and objective conformity criteria for digital goods[40] and the vastly different proposals on this matter by the COM, Council and EP show that a full alignment in that regard between DCD and OSD is unlikely. Besides, such an alignment would actually undermine to some extent the idea behind the DCD.[41] Although we embrace any approaches of aligning both directives, the following question will remain relevant: What set of rules should apply to situations, where both directives are applicable but provide different rules?

42 One possible solution would be applying the DCD to the digital component and the OSD to the physical component of a good. However, distinguishing between the digital and physical components can be rather tricky in real-life situations, making it difficult for the consumer to determine where the conformity deficiency lies. To solve this matter, the consumer could have the right to base the non-conformity claim on the DCD without the need to prove that the problem indeed relates to the digital part. To balance the picture, it was proposed by ELI that the supplier should have the opportunity to show that the problem lies within the physical part, and hence, the consumer would have no rights under the DCD.[42]

43 We welcome this approach but would recommend going one step further for achieving a higher level of clarity, certainty and consumer protection. In case of such a rebuttable assumption as proposed, the success of the consumer's claim under the DCD would depend on the trader not being able to prove that the deficiency falls within the physical component of the good. The problem here is that the consumer has limited possibilities to tell where the deficiency

38 COM, Amended proposal for a directive of the European Parliament and of the Council on certain aspects concerning contracts for the online and other distance sales of goods, COM(2017) 637 final – 2015/0288 (COD), 31.10.2017.

39 Explanatory Statement within DCD-EP p. 90.

40 Cf. Aurelia Colombi Ciacchi/Esther van Schagen, 'Conformity under the Draft Digital Content Directive: Regulatory Challenges and Gaps' in Reiner Schulze/Dirk Staudenmayer/Sebastian Lohsse (eds), *Contracts for the Supply of Digital Content: Regulatory Challenges and Gaps, Münster Colloquia on EU Law and the Digital Economy II* (Nomos, Baden-Baden, 2017) 99, 102 ff.

41 If one would assume that the same conformity criteria could be adopted to digital goods and physical goods, most rules concerning conformity for sales contracts in the DCD could be deleted and replaced by a referral to the OSD, making a large part of the regulation superfluous.

42 ELI (n 27) 12.

lies – and thus is not able to properly assess the risk of a lawsuit. In contrast, the 6-month reversal of the burden of proof under the CSGD[43] is more consumer-friendly: to answer the question whether the lack of conformity existed at the time of delivery, the consumer has some first-hand knowledge about facts such as when the fault initially occurred or whether he/she might have been responsible for the problem (e.g., by dropping the good or by not handling it properly in any other way). By comparison, when it comes to the question of where the defect lies, such or other facts will be mostly unavailable to the consumer. So, if the trader denies the consumer his rights under the DCD by simply alleging a physical fault, the consumer would be unlikely to challenge that claim, rendering this solution impracticable.

44 Therefore, an alternative approach to this dilemma appears more appropriate: The consumer should be mostly free to claim a defect in the physical part of the good or a fault in the digital component, a choice the trader should not be able to challenge by proving that the consumer's choice was incorrect. By performing that choice, the consumer should only be restricted by an "obviousness principle". Namely, relying on DCD remedies in a given case will be denied only if it is apparent without further investigation and expertise that the problem lies in the physical part of the good (and vice versa, if the consumer relies on OSD remedies).[44]

45 Following this consumer-friendly approach is a conscious policy decision that should guide the legislative process. In some cases, this indeed might lead to an extended liability of the trader; yet strong consumer protection and legal certainty are gained, and it should be easier for the trader to compensate for possible financial drawbacks resulting from the legal exposure, for example, by raising the price.

7. Multi-party scenarios

46 Multi-party scenarios that are typical to the supply of EDCS call for more discussion and analysis. The current DCD proposals seem to focus on bilateral contracts while overlooking more complex settings that involve multiple players in direct contact with the consumer.[45] Very often, the digital part of smart goods is supplied and maintained not by the vendor but by a third party (e.g., the product manufacturer). In such situations, the consumer's interests could be affected amongst others by: (1) the direct affiliate (i.e., the vendor); (2) the manufacturer of the physical good; (3) the (technical) supplier of the digital content / service; or (4) the data processor.

47 Multi-party scenarios are obviously not unique to EDCS, but there are several aspects of smart goods, especially the rights and duties connected to consumers' data, that call for enhanced attention.

48 For example, it has to be clarified (possibly within the Recitals), that the consumers' claims against the trader are not diminished by whether the trader does or does not also fulfil the functions of the manufacturer, digital content supplier, or data processor. Clarifications like these are crucial to close loopholes emerging from multiple parties being involved. For example, for "analogue" sale contracts, it is legally unambiguous that the contracting party is liable for the product sold. Even so, it can be observed that in many cases the consumer is redirected to the producer, often giving the impression that the consumer has no rights against the contracting party. It is foreseeable that the contracting parties liable under the DCD will use the same mechanisms to avoid requests to make available or delete user-generated content – a scenario that has to be avoided. The fact that a third party is responsible for data processing, for instance, should not automatically render Article 13a(4) DCD-EP (portability of user-generated content) inapplicable. Instead, the contracting party should have the obligation to support the consumer to the full extent possible, for instance by providing information about the data processor or the consumer's rights according to the GDPR.

49 We strongly recommend to further investigate this aspect and other possibilities to strengthen the consumers' position against third parties through full harmonisation in this field and to examine multi-party scenarios in general. If, however, the trilogue chooses not to harmonise such questions, the final draft should clarify that the DCD does not prejudice the ability of domestic law to regulate these questions.

43 Article 5(3) Directive 1999/44/EC of the European Parliament and of the Council of 25 May 1999 on certain aspects of the sale of consumer goods and associated guarantees [1999] OJ L 171/12.

44 For example, if there is a visible crack in the display of a smartphone, the consumer should not be able to make a claim under the DCD, but if the device keeps restarting with no explicable reason, the consumer should be free to choose between both instruments without risking having chosen the "wrong" set of remedies when it comes to litigation against the trader.

45 *Cf., focusing on the license holder:* Beale, 'Conclusion and Performance of Contracts: An Overview' in Reiner Schulze/Dirk Staudenmayer/Sebastian Lohsse (n 40) 33, 37; especially on data portability for smart goods: Janal, 'Data Portability – A Tale of Two Concepts' (2017) Journal of Intellectual Property, Information Technology and Electronic Commerce Law 59, 65 f.

III. Recommendations

1. **It is crucial to establish a harmonised level of consumer protection for embedded digital content and services (EDCS) covering the digital element of smart goods. EDCS must therefore be covered by the DCD.**

2. **The existing differentiations between stand-alone and embedded digital content/services at the level of the DCD's scope should be removed. EDCS should be understood as a subset of "supply of digital content or services" and be covered as such.**

3. **The EDCS-specific rules in Articles 10(1)(b), 9(1)(c), 13b(2), and 3(3) DCD-EP should be deleted. Only where absolutely necessary, EDCS-specific rules should be implemented.**

4. **For sales contracts, the OSD should generally apply to the physical component of the good and the DCD should generally apply to the digital component. Unless the non-conformity/defect obviously relates to either the digital or to the physical component, consumers should have the free choice on which set of norms to base their claim against the trader.**

5. **The consumer protection implications arising from multi-party scenarios must further be investigated and expressly addressed in the final text of the DCD.**

E. Portability

I. Relevant Provisions

European Commission (09.12.2015)	Council of the European Union (01.06.2017)	European Parliament (27.11.2017)
Article 13(2)(c)	Art 13a[I]	Art 13a[II]
When the consumer terminates the contract, the supplier shall provide the possibility to retrieve all content provided by the consumer and any other data produced or generated through the consumer's use of the digital content to the extent that data has been retained by the supplier. The consumer shall be entitled to retrieve the content free of charge, without significant inconvenience, in reasonable time and in a commonly used data format.	(2) **In respect of personal data of the consumer,** the supplier shall **comply with the obligations applicable under Regulation (EU) 2016/679 (...).** (3) **Furthermore,** the supplier **shall make available to the consumer any digital content** (...) **to the extent that it does not constitute personal data, which was uploaded or created by the consumer when using the digital content or digital service supplied by the supplier.** **The supplier shall not be required to make available such digital content created by the consumer when using the digital content or digital service to the extent that such digital content created by the consumer only has utility within the context of using the digital content or digital service supplied by the supplier, or which relates only to the consumer's activity when using the digital content or digital service supplied by the supplier or which has been aggregated with other data by the supplier and cannot be disaggregated or only with disproportionate efforts.**	**(2) In respect of personal data of the consumer, the trader shall comply with the obligations applicable under Regulation (EU) 2016/679.** **(4) The trader shall, upon request by the consumer, make available to the consumer any user-generated content to the extent that it does not constitute personal data, which was provided or created by the consumer when using the digital content or digital service supplied by the trader. The consumer shall be entitled to retrieve the content free of charge, without significant inconvenience, in reasonable time and in a commonly used and machine-readable data format.** **The obligation to make available such user-generated content shall not apply in case the user-generated content:** **(a) cannot be made available without disproportionate and unreasonable effort because it has no utility outside the context of the digital content or digital service supplied by the trader;**

	The consumer shall be entitled to **retrieve that digital** content free of charge, without **hindrance** from the supplier, in reasonable time and in a commonly used and **machine-readable** format. [...]	**(b) cannot be made available without disproportionate and unreasonable effort because it only relates to the consumer's activity when using the digital content or digital service supplied by the trader; or** **(c) has been aggregated with other data by the trader and cannot be disaggregated or only with disproportionate efforts.**

[I] Footnotes in the DCD-Council text omitted. Emphasis in original.

[II] Emphasis in original.

II. Comments

1. Purpose of portability provisions

50 The data portability rules of the DCD serve different purposes. The first and most obvious function is to safeguard the consumer's right of termination in order to avoid lock-in effects, see Recital 39 DCD-COM. A second purpose is to foster competition, see Recital 46 DCD-COM. The parallel provision in Article 20 GDPR underlines that portability provisions do also aim at the empowerment of the data subject. However, Article 13 DCD-COM is not restricted to personal data but covers also user generated content that is not personal data (in the following: UGC). The general tendency of the provision is to be welcomed. Without portability requirements, at least after termination of a contract, lock-in effects will prevent consumers from switching from one service to another. As a consequence, the consumers' freedom to make a choice and competition between services would be affected.

2. One coherent portability regime for personal data

51 The main difference between the COM's, Council's and EP's proposals concerns the applicable portability regime for personal data. In its proposals, the Council and EP suggest that for personal data, the portability provision of Article 20 GDPR should apply instead of the DCD. The clear advantage of a streamlined portability regime for personal data is its coherence. With one portability regime, codified in the directly applicable GDPR, it would be easier for both consumers and service providers in the EU to know their rights and duties and to adapt their conduct to the legal rules. Since Article 20 GDPR only covers personal data, it is vital that Article 13 DCD maintains additional rules on UGC. Even though most content created by consumers in the current business models meets the criteria of personal data, the provisions should be drafted in a forward-looking wording and cover as many different services as possible.

3. GDPR provides a higher level of protection

52 Abandoning the specific portability rules in the Directive should only be considered if the protection given by the GDPR arrives at the same level as Article 13 DCD-COM. Apparently, the most important advantage for consumers under Article 13 DCD is the broader field of application vis-a-vis UGC. By contrast, for personal data Article 20 GDPR provides a higher level of protection.

a) As Article 13 DCD, Article 20 GDPR:

- covers both personal data given with the consent of the data subject and data necessary for the performance of a contract;
- has a territorial scope according to Article 3 GDPR, which is comparable to the consumer contract rules of Article 6 Rome I Convention 593/2008;
- obliges the controller to provide the data in a structured, commonly used and machine-readable format and free of charge, see Article 12(5) GDPR.

b) Different from Article 13 DCD and more favourable for the data subject, Article 20 GDPR secures for the data subject the right:

- to receive the data at any moment, not only after termination of the contract;
- to ask for transmission of the data directly from one controller to another, where technically feasible;
- to retrieve personal data in case of embedded data processing devices assumed they are not covered by the DCD;
- to retrieve personal data from any controller and not just from the contracting

party of the consumer who might not even control the personal data if he acts as a mere reseller.

c) There are also aspects in which Article 13 DCD might provide a higher level of protection than Article 20 GDPR. However, closer examination shows that these differences concern few cases of a limited practical importance:

- Article 20 GDPR restricts the portability right to data for which the data subject has given its consent (Article 6(1)(a) GDPR) and to data necessary for the performance of the contract (Article 6(1)(b) GDPR). Article 20 GDPR does not cover data that has been processed unlawfully by the controller. Article 13 DCD may appear as more comprehensive. It covers all data "produced or generated through the consumer's use of the digital content". This may also cover any processing of data beyond the terms of the contract between consumer and trader. But given the fact that the trader determines the use of the data by its terms and conditions, cases of unlawful use in the framework of a contract are hard to imagine as long as the terms and conditions are valid. By contrast, cases of void contract terms should be solved by a sound interpretation of "consent" in the sense of Article 6(1)(a) GDPR.

53 All aspects considered, the Council's and EP's suggestion to replace the portability provision in Article 13 DCD by a reference to the GDPR is well-founded. However, the link to the GDPR should be clarified by an explicit reference to the portability regime enshrined in the GDPR. Also, the portability right for content that is not personal data should be maintained.

4. Remaining provisions on UGC (other than personal data)

54 With regard to UGC, the Council's and EP's proposals suggest a number of exceptions to the portability provision in Article 13a. These exceptions may seriously weaken the consumer's right to retrieve UGC provided to the supplier. For the application of the exceptions, it may not suffice that the suppliers assert "no utility outside the context of the digital content or digital service", that "only relates to the consumer's activity when using the digital content or digital service" or "has been aggregated with other data by the trader". Rather, it should suffice that the consumer claims that he sees utility outside the context of the service or that he wants to use the content outside of the service. With regard to the proportionality requirement, the provisions should explicitly oblige the supplier to configure its service in a way that allows UGC to be extracted separately for each consumer. Service providers should apply state-of-the-art technology to protect the consumers' interest in its UGC. If suppliers do not set up their services in such a way as to facilitate the retrieval of consumers' UGC to the maximum effect possible according to state-of-the-art technology, they should not be heard with the argument of disproportionality. The EU legislator should keep in mind that portability rules serve a competition-enhancing purpose. Consumers seeking to retrieve their personal data and UGC to change over to other traders are the key for a functioning digital single market.

55 Moreover, the provisions of portability of UGC should reflect that the trader may not always be the party who stores and processes the content generated by the consumer, e.g. in case of digital content supplied by a mere reseller which enables the consumer to access a service provided by a third party. In such a case, the consumer should have an additional direct right against this third party to retrieve its content.[46]

5. Long-term contracts

56 The right to retrieve personal data and other UGC must also be ensured in case of termination of long-term contracts according to Article 16 DCD. The COM's proposal suggests in Article 16(4)(b) to provide a rule which is in line with the termination rule in Article 13(2)(c). The Council proposes to implement a reference to Article 13a and to the GDPR into Article 16(3), which would streamline both sets of rules. Such a reference is missing in DCD-EP with regard to UGC. According to DCD-EP, in case of termination of a long-term contract, the consumer would have the right to retrieve personal data based on Article 20 GDPR. However, the drafters have obviously overlooked the necessity to provide a parallel rule for other UGC. The final text of the DCD should either stipulate explicit portability rules for UGC or contain a reference to Article 13a.

46 Whether such a right should be implemented as a direct action against the third party or as an obligation of the supplier to provide the consumer with enforceable rights against the third party (or to pay damages in case of breach of the obligation) is subject to further discussion.

III. Recommendations

1. **The portability of personal data should be governed exclusively by Article 20 GDPR. Article 13 DCD should refer explicitly to the GDPR.**

2. **The portability of UGC should not be undermined by too broadly defined exceptions, as proposed by the Council and EP. The criteria "no utility outside the context of the digital content or digital service" and "only relates to the consumer's activity when using the digital content or digital service" should be deleted. The right to retrieve UGC should only be excluded if it cannot be made available without disproportionate and unreasonable effort. It should be clarified that suppliers have a duty to apply state-of-the-art technology to guarantee that consumer's UGC can be extracted separately. If suppliers do not apply such technology, they should not be heard with the argument that portability is disproportionate.**

3. **The DCD must ensure that consumers have a right to retrieve UGC against the trader and any third party that stores and/or processes the content.**

4. **The DCD must ensure that portability of personal data and other UGC is ensured for long-term contracts under Article 16. The provision must either contain explicit rules or a reference to the GDPR and to Article 13 (or 13a) DCD.**

F. Conformity, Modifications, Termination

I. Relevant Provisions – Conformity

European Commission (09.12.2015)	Council of the European Union (01.06.2017)	European Parliament (27.11.2017)
Article 6(2)	**Article 6a[I]**	**Article 6a[II]**
To the extent that the contract does not stipulate, where relevant, in a clear and comprehensive manner, the requirements for the digital content under paragraph 1, the digital content shall be fit for the purposes for which digital content of the same description would normally be used including its functionality, interoperability and other performance features such as accessibility, continuity and security, taking into account: (a) whether the digital content is supplied in exchange for a price or other counter-performance than money; [...]	**Objective requirements for conformity of the digital content or digital service** **1. (...) In addition to complying with any conformity requirements stipula-ted in the contract the** digital content **or digital service** shall: **(a)** be fit for the purposes for which digital content **or a digital service** of the same **type** would normally be used, **taking into account**, where **applicable**, any existing **(...) national and Union laws**, technical standards or, in the absence of such technical standards, applicable **sector specific** industry codes of conduct [...] **2. There shall be no lack of conformity within the meaning of paragraph 1 if, at the time of the conclusion of the contract, the consumer was specifically informed that a particular characteristic of the digital content or digital service was deviating from the conformity requirements stipulated in paragraph 1 and the consumer has expressly and separately accepted this deviation when concluding the contract.**	**Objective requirements for conformity with the contract** **1. The digital content or digital service shall, where relevant:** **(a) possess qualities and performance features, including in relation to functionality, interoperability, accessibility, continuity and security, which are usually found in digital content or digital services of the same type and which the consumer may reasonably expect, given the nature of the digital content or digital service, and comply with, where relevant, any existing international or European technical standards or, in the absence of such technical standards, applicable industry codes of conduct and good practices, including on the security of information systems and digital environments [...]**

[I] Footnotes in the DCD-Council text omitted. Emphasis in original.

[II] Emphasis in original.

II. Comments – Conformity

57 The DCD aspires to harmonise a set of key rules, *inter alia,* in the areas of conformity of digital content with the contract, certain aspects concerning modification of the content, and termination (Recital 8 DCD-COM). As a result, Member States will not be permitted to provide more or less protection to consumers in the regulated area (Article 4 DCD-COM).

58 Once it has been decided to include data as counter-performance (DACP) transactions within the scope of the DCD, it appears advisable, as a matter of principle, not to discriminate between DACP-consumers and price-paying consumers, unless (1) discrimination is called upon due to the nature of counter-performance as data, or (2) discrimination is supported by an important public policy argument. It cannot be assumed that DACP-consumers *per se* are less worthy of (harmonised) protection both as a matter of equal treatment and as this premise does not appear to promote a better functioning Digital Single Market.

59 In the context of conformity, the COM's proposal prioritises subjective criteria (Article 6(1) DCD-COM) and would only consider objective criteria to the extent that important aspects of the transaction are not stipulated in the contract in a clear and comprehensive manner (Article 6(2) DCD-COM). Among other things, one of the elements that need to be taken into account while performing an objective conformity scrutiny is the question whether "the digital content is supplied in exchange for a price or other counter-performance than money."

60 It is not readily clear why this consideration is relevant, and if so, how the DACP-aspect of a contract should influence the application of conformity standards. Applying the non-discrimination principle described above suggests that discrimination between consumer groups on this basis is neither mandated by the nature of the counter-performance nor is it supported by an important public policy goal.

61 The EP proposed to apply objective conformity criteria *alongside* subjective criteria and not only where the contract is silent or unclear (Article 6a DCD-EP). The Council follows a similar approach in suggesting that objective criteria are applicable "[i]n addition to complying with any conformity requirements stipulated in the contract."

62 Objective conformity checks are important, and they might be especially important for DACP-consumers. The assumption that DACP-contracts usually involve small-value transactions, at least as typically perceived by consumers,[47] as they are less likely to insist on sufficiently clear or comprehensive provisions in the contract itself in such cases; they might not even bother to read it. It is therefore recommended to apply conformity provisions essentially in an equal manner regardless of the question whether the consumer is required to pay a price or to provide data.[48]

III. Relevant Provisions – Modification and Termination

European Commission (09.12.2015)	Council of the European Union (01.06.2017)[I]	European Parliament (27.11.2017)[II]
Article 15	Article 15	Article 15
Modification of the digital content 1. Where the contract provides that the digital content shall be supplied over the period of time stipulated in the contract, the supplier may alter functionality, interoperability and other main performance features of the digital content such as its accessibility, continuity and security, to the extent those alternations adversely affect access to or use of the digital content by the consumer, only if: (a) the contract so stipulates; (b) the consumer is notified reasonably in advance of the modification by an explicit notice on a durable medium;	**Modifications of the digital content or digital service** 1. Where the contract **specifies** that the digital content or digital service shall be **available to the consumer over a period of time (...)**, the supplier **shall be allowed to modify (...)** the digital content or **digital service supplied to** the consumer **(...), provided the following conditions are met:** (a) the contract **allows and gives a valid reason for such a modification,** and **(b) the modification is provided without additional costs for the consumer, and** (c) the consumer is **informed in a clear and comprehensible**	Modification of the digital content ***or digital service*** 1. Where the contract provides that the digital content ***or the digital service is to*** be supplied ***or made accessible over a*** period of time stipulated in the contract, the ***trader*** may ***only*** alter ***the*** functionality, interoperability and other main performance features of the digital content ***or digital service beyond what is necessary to maintain in conformity*** the digital content ***or digital service in accordance with Article 6a*** if: (a) the contract ***allows for, and gives a valid reason for, such a modification;***

47 *Cf.* Yoan Hermstrüwer, 'Contracting Around Privacy: The (Behavioral) Law and Economics of Consent and Big Data' (2017) Journal of Intellectual Property, Information Technology and Electronic Commerce Law 9 ("This currency [personal data] seems to be inherently inclusive and egalitarian, since there is no need to be wealthy in order to pay with data.").

48 *See also*, Vanessa Mak (n 31).

(c) the consumer is allowed to terminate the contract free of any charges within no less than 30 days from the receipt of the notice; and (d) upon termination of the contract in accordance with point (c), the consumer is provided with technical means to retrieve all content provided in accordance with Article 13(2)(c). 2. Where the consumer terminates the contract in accordance with paragraph 1, where relevant, (a) the supplier shall reimburse to the consumer the part of the price paid corresponding to the period of time after modification of the digital content; (b) the supplier shall refrain from the use of the counter-performance other than money which the consumer has provided in exchange for the digital content and any other data collected by the supplier in relation to the supply of the digital content including any content provided by the consumer.	**manner** of the modification, **provided that in the cases referred to in paragraph 2 the consumer is informed** reasonably in advance **on a durable medium of the features and time of** the modification, **and of his right to terminate the contract in accordance with paragraph 2 and 3, or, where applicable, about the possibility to maintain the digital content or digital service without modification in accordance with paragraph 5. (...)** 2. **The** consumer **shall be entitled** to terminate the contract **(...) if the modification negatively impacts the access to or use of the digital content or digital service by the consumer, unless such negative impact is only minor.** 3. **The consumer shall be entitled to exercise the right to terminate the contract in accordance with paragraph 2 without additional costs** and within no less than 30 days from **the day on which he is informed according to paragraph 1(c). The right to terminate the contract shall end not earlier than 14 days from the date of application of the modification. (...)** 4. **Where the consumer terminates the contract in accordance with paragraphs** 2 and 3 (...), **the supplier shall reimburse to the**	***(aa) such a modification can reasonably be expected by the consumer;*** ***(ab) the modification is provided without additional cost to the consumer; and*** (b) ***the trader notifies*** the consumer reasonably in advance ***in a clear and comprehensible manner and*** on a durable medium ***of the modification and, where applicable, of his right to terminate the contract under the conditions provided for in paragraph 1a;*** ***1a. The consumer shall be entitled to terminate the contract if the modification negatively impacts the access to or the use of the digital content or digital service by the consumer, unless such negative impact is only minor. In that case, the consumer shall be entitled to terminate the contract free of charge within 30 days after the receipt of the notice or from the time when the digital content or digital service is altered by the trader, whichever is later.*** 2. Where the consumer terminates the contract in accordance with paragraph ***1a of this Article, Articles 13, 13a and 13b shall apply*** ***Article 13a*** ***[...]*** ***2. In respect of personal data of the consumer, the trader shall comply with the obligations applicable under Regulation***
	consumer only **the proportionate part of the price paid corresponding to the period of time after the modification of the digital content** or digital service. **5. Paragraphs 2 to 4 shall not apply if the supplier has enabled the consumer and the consumer has accepted to maintain without additional costs the digital content or digital service without the modification, and the digital content or digital service remains in conformity.**	***(EU) 2016/679.*** ***3. The trader shall make every effort that he could be expected to make to refrain from the use of any user- generated content to the extent that it does not constitute personal data, which was provided or created by the consumer when using the digital content or digital service supplied by the trader, with the exception of:*** ***(a) the content that cannot be refrained from using without disproportionate and unreasonable effort because it has no utility outside the context of the digital content or digital service supplied by the trader;*** ***(b) the content that cannot be refrained from using without disproportionate and unreasonable effort because it only relates to the consumer`s activity when using the digital content or digital service supplied by the trader;*** ***(c) the content which has been generated jointly by the consumer and others, when other consumers can continue to make use of the content;*** ***(d) the content that has been aggregated with other data by the trader and cannot be disaggregated or only with disproportionate efforts.***

[I] Footnotes omitted. Emphasis in original.

[II] Emphasis in original.

IV. Comments – Modification and Termination

63 Article 15(1)(c) DCD-COM ff. stipulates the remedy of termination in case of a negative impact resulting from the supplier modifying the digital content/service. According to the COM's proposal, the consumer may terminate the contract without any charges within no less than 30 days from receipt of notice. In addition, Article 15(2)(b) DCD-COM stipulates the duty of the supplier to refrain from using data that has been provided as counter-performance after such termination. By comparison, the EP would maintain a similar rule of termination if the modification negatively impacts the access to or the use of the digital content or digital service by the consumer (Article 15(1)(a) DCD-EP). Regarding the consequences of termination, Article 15(2) DCD-EP refers to the general termination provisions stipulated in Articles 13, 13a and 13b DCD-EP.

64 In turn, Article 13a DCD-EP makes a distinction between personal data (subsection 2) and "user-generated content to the extent that it does not constitute personal data" (subsection 3). Regarding personal data, subsection 2 mandates a direct application of the GDPR, but regarding non-personal user-generated content, subsection 3 formulates an obligation to refrain from using that content after termination, while adding to it a fairly detailed scheme of exceptions.[49]

65 Interestingly, the result is a *de facto* discrimination *in favour* of consumers who extend personal data in return for content/services, since their ability to withdraw their consent under the GDPR – and thereby, effectively bring the contract to an end if their consent is a condition to the continuation of the relationship with the trader – is unqualified. In this case, however, the priority of the GDPR (specifically, Article 17(1)(b) GDPR)[50] over commercial regulation concerning contract termination mandates a differentiated treatment.

66 Such discrimination surely has practical implications. To name one example, under the EP proposal, termination in the case of modification with negative impacts is only effective 30 days from receipt of notice or from the time when the digital content or digital service is altered by the trader, whichever is later. By comparison, under Article 17(1) GDPR, once consent is withdrawn, with or without reason, the data subject has the right "to obtain from the controller the erasure of personal data concerning him or her *without undue delay* and the controller shall have the obligation to erase personal data *without undue delay*" (emphasis added). Viewed from this vantage point, the GDPR in fact creates an alternative termination regime that is comparatively insensitive to the commercial considerations underlying Article 13a(3) DCD-EP.

67 This interface point provides an example for a situation, in which the DCD *cannot* provide equal treatment to both consumer groups. This is a structural limitation of the DCD that cannot be undone by its drafter. Rather, an alternative route to prevent unreasonable results to the detriment of traders can be paved by national courts as they apply privacy regulations alongside consumer protection regulations. While doing so, national courts should be permitted to apply contract law remedies available to traders in case the withdrawal of consent is considered a (material) breach of contract under local doctrines, to the extent that such remedies do not collide with data protection law. For instance, if the domestic contract law in such case permits the immediate termination of the contract by the trader without notice, the DCD should not influence the effectiveness of such remedies.[51]

V. Recommendations

1. **A harmonised level of consumer protection under the DCD in the context of conformity should principally apply in an equal manner to DACP-consumers and paying consumers alike. The non-discrimination principle should guide the formulation of the DCD with the focus on avoiding unjustified differentiation between the two classes of consumers.**

2. **Objective conformity requirements play an important role within the harmonised consumer protection scheme. The type of counter-performance (data or price) should not result in lower requirements in the case of DACP-contracts, and by extension, a lower level of protection to DACP-consumers. Accordingly, subsection (a) under Article 6(2) DCD-COM should be either clarified or removed. In addition, removing the structural hierarchy between subjective and objective conformity criteria in line with**

49 Article 13a(3) DCD-EP.

50 Article 17(1)(b) GDPR: "The data subject shall have the right to obtain from the controller the erasure of personal data concerning him or her without undue delay and the controller shall have the obligation to erase personal data without undue delay [...]".

51 *See e.g.,* § 314 Abs. 1 BGB: "Each party may terminate a contract for the performance of a continuing obligation for a compelling reason without a notice period. [...]" (translation as available under <https://www.gesetze-im-internet.de/englisch_bgb/englisch_bgb.html#p1150> accessed 23 March 2018).

the approaches suggested by the EP and the Council would contribute to preventing an indirect discriminative effect.

3. **The application of data protection law to situations that are commercial in nature (such as the right to termination in general, or specifically, termination in the case of modification to the detriment of the consumer) marks the limits of the non-discrimination principle in favour of consumers that extend their personal data in exchange for commercial offers. Yet, the DCD should not intentionally inhibit the ability of domestic contract laws to provide remedies to traders in the appropriate case and to the extent that such remedies are in line with EU data protection law.**

List of Abbreviations

COM	European Commission
CRD	Consumer Rights Directive
DACP	Data as counter-performance
DCD	Digital Content Directive
EDCS	Embedded Digital Content and Services
EDPS	European Data Protection Supervisor
ELI	European Law Institute
EP	European Parliament
EU	European Union
GDPR	General Data Protection Regulation
LIBE	European Parliament Committee on Civil Liberties, Justice and Home Affairs
OSD	Online Sales Directive
TFEU	Treaty on the Functioning of the European Union
UGC	User-generated content

jipitec

Journal of
Intellectual Property,
Information Technology,
and Electronic Commerce
Law

www.jipitec.eu