jipitec

2 | 2016

Volume 7 (2016)
Issue 2 ISSN 2190-3387

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

Articles

www.jipitec.eu

jipitec

Journal of Intellectual Property, Information Technology and Electronic Commerce Law

Volume 7 Issue 2 September 2016
www.jipitec.eu
contact@jipitec.eu

A joint publication of:

Prof. Dr. Thomas Dreier, M. C. J.,
Karlsruhe Institute of Technology,
Vincenz-Prießnitz-Str. 3,
76131 Karlsruhe

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.

ISSN 2190-3387

Funded by

DGRI Deutsche Gesellschaft für Recht und Informatik e.V.

Table Of Contents

Articles

Proposals from Berlin and Paris – Intermediary Liability in European Copyright Law

by **Jonathan Griffiths***

Abstract: Two very different proposals on copyright policy – one a privately drafted document, the other a governmental report – are published in this edition of JIPITEC. There is an interesting point of intersection between them because they both consider the difficult question of the liability of online intermediaries for users' infringements. The first document is "The Berlin Gedankenexperiment on the Restructuring of Copyright Law and Authors Rights". This is a wide-ranging proposal for a complete recasting of the legal system that promotes the production of, and controls the use of, creative goods. The second policy document has a more limited focus. The French High Council for Literary and Artistic Property ("CSPLA")'s Mission to Link Directives 2000/31 and 2001/29 – Report and Proposals ("Mission Report") aims to provide a persuasive intervention in current policy discussions at European Union level concerning the liability or, more appropriately, the non-liability, of online intermediaries for copyright infringement. In this brief introduction, I outline the scope of both proposals and reflect briefly on their recommendations.

Keywords: Copyright Law, Copyright Reform, Intermediaries, Germany, France, French High Council for Literary and Artistic Property (CSPLA)

Recommended citation: Jonathan Griffiths, Proposals from Berlin and Paris – Intermediary Liability in European Copyright Law, 7 (2016) JIPITEC 70 para 1.

A. The Berlin Gedankenexperiment on the Restructuring of Copyright Law and Authors Rights

1 The Berlin *Gedankenexperiment* is the product of a project undertaken by a panel of German experts, predominantly from academic and "new media" backgrounds.[1] It develops an earlier set of guidelines on copyright policy issued by the "Internet & Gesellschaft Collaboratory",[2] which is supported by, amongst others, Creative Commons, Google and the Wikipedia Foundation. The project takes a "blank page approach", allowing the *Gedankenexperiment* to escape prevailing copyright norms where appropriate. However, it is clearly shaped by the view that the current system of copyright and author's rights is problematic in a number of respects.

2 The dominant problem at which the proposal takes aim is the transferability of legal entitlements in creations from authors to other categories of actor (generally to "exploiters", under the terminology employed). On this point, it is argued that, where one legal actor steps into the shoes of another in this way, there is a risk that the fundamental purpose of a legal regime designed to foster creativity will be subverted:

> *"Confusing and mixing authors' and exploiters' interests opens space for manipulative arguments, which may foster*

1 The German original is available at <https://irights.info/wp-content/uploads/2015/08/Gedankenexperiment.pdf>. - The English translation is reprinted on p. 76 of this volume.

2 <http://en.collaboratory.de/w/About_us>.

undesirable developments. These lead to a conflict of values, which may ultimately undermine the copyright law system as a whole."[3]

3 The *Gedankenexperiment* seeks to avoid this problem through the establishment of strict distinctions between the interests of different legal actors involved in the production and use of creative work. Under the proposed system, different categories of legal actors in the creative process ("authors", "exploiters", "non-commercial users" and "intermediaries") are each accorded their own entirely independent rights and duties. The legal position of each is balanced against that of others, without attribution of "structural superiority" to any amongst them. The underlying idea is that the separate contribution of each to the generation of creative products should be separately recognised and protected.

4 The authors of the *Gedankenexperiment* suggest that a whole-hearted commitment to this core idea would produce a legal structure differing from that which is currently applicable. For example, under the system proposed, a creator (such as a novelist) would be able to grant an "exploiter" (such as a publisher) (contractual) permission to exploit a protected work. However, as a matter of law, such permission could only be granted for a limited period of time.[4] At the same time, the publisher would itself acquire its own separate legal right in its published edition (recognising its own distinctive contribution to the dissemination of creativity). At the end of the limited period of permission, the publisher could continue to market its own published editions. However, from that point forward, it would potentially be subject to competition from other published editions permitted by the author.

5 The proposal is not intended to establish a fully codified body of rules. It is, after all, a *Gedankenexperiment* and it therefore raises, but does not come to a concluded view on, a number of features of the proposed system - including the precise duration of the various forms of protected interest and the specific treatment of complex works such as films and of works created by employees. In tracing the outline of a legal structure in this way, the project insulates itself from detailed critical analysis. Nevertheless, even against this avowedly sketchy background, it invites questions from the concretely-minded. For example, one of the categories of legal actor to which the proposal attributes rights and duties is described as "non-commercial users". Such users have a *right* to carry out acts which either (i) fall within a specified catalogue of use rights or (ii) are covered by an open fair use-type norm. However, the document makes no mention of *commercial* users in this context. Such users presumably fall within the category of "exploiters", who have their own designated duties and rights. However, while the proposal traces the entitlement of exploiters who have been granted contractual permission to use a copyright work, it does not appear to deal explicitly with the situations in which a commercial actor is typically entitled to use a copyright work without permission under current law (for example, for the purpose of quotation, news reporting or parody). This must surely simply be an omission. If the project team had intended to restrict the circumstances under which commercial users are entitled to commit otherwise infringing acts, it would surely have done so explicitly.[5]

6 The *Gedankenexperiment*'s 'blank page" approach undoubtedly brings a breath of fresh air to the sometimes poisonous debate on copyright reform. However, while it might appear to be based on a radical premise, its recommendations are relatively incremental in some respects. Many features of existing copyright and authors' rights systems – including creators' moral rights, copyright contract regulation and the special regimes applicable to film productions and to creations by employees – are retained. Indeed, even those elements of the proposal involving significant change to the existing legal order (with the possible exception of the elaborated "balance of independent rights" system outlined above) echo suggestions for reform made elsewhere in the recent past. Thus, for example, under the *Gedankenexperiment*, the terms of protection for authors' and exploiters' rights would be significantly shorter than those currently prevailing in European and international law. Many such calls to reduce the term of copyright so that it more closely reflects its underlying rationales have been made. Similarly, the proposal's suggestions that authors' promises of exclusivity should be

3 *Berlin Gedankenexperiment*, 3.

4 Ibid, 4. The document does not provide a final recommendation of an appropriate period for which permission may be given. – It should be noted that in its recent proposal for a law improving the claim of authors and performers to adequate remuneration, the German Federal Government proposes, albeit on the basis of a different legal construction, a somewhat similar result in providing that an author, who has granted an exploiter an exclusive exploitation right against payment of a lump sum fee, shall be free after ten years to exploit his work otherwise, with the initial exploiter retaining a non-exclusive right to continue his own exploitation; see § 40a (1), BT-Drucks. 18/8625.

5 Editor's note: The drafters of the *Gedankenexperiment* have explained that the term „non-commercial user" was only chosen to set a clear terminological distinction between users who are entitled to use protected material by statute and those who need a license for their uses (the latter are referred to as „exploiters"). This does not necessarily mean that there uses for a commercial purpose cannot fall under the user's statutory rights. Whether this is the case or not, depends on the particular balance of interests.

noted in a public registry, that the continuation of the exclusive protection for those that invest in the dissemination of works should be conditional on the payment of progressively increasing fees and that exceptions and limitations should be replaced with a set of "user's rights" for non-commercial users and a fair use-type clause all are not without precedent. Indeed, in the last case, recent judgments of the Court of Justice may already have delivered such an outcome in the European Union.[6]

7 Even the *Gedankenexperiment*'s most distinctive proposal, the "balance of independent rights" system may not have been breathed into life *ex nihilo*. To this common lawyer's untrained eye, the project's emphasis on the non-transferability of author's entitlements looks rather like a super-charged extension of the current German system for the protection of author's rights. It would appear that the page upon which this stimulating proposal has been drawn up may not have been entirely blank after all.

B. Intermediaries

8 The relatively reasonable, modest characteristics of the *Gedankenexperiment* are also apparent in its proposals concerning the potential liability of online intermediaries for copyright infringement. This is one of the most contested questions in current debates on copyright policy. Legal systems have struggled to develop appropriate theories to impose responsibilities on intermediaries without over-burdening them in a manner that would unreasonably hamper the functioning of new forms of communication technology. Considerable uncertainty on this question persists in many jurisdictions.[7] Within the European Union's legal order, online intermediaries benefit from the E-Commerce Directive's "safe harbours" for information service providers.[8] Under these general provisions, when information service providers function as "mere conduits"[9], "caches"[10] or as "hosts"[11] for information originating from others, they enjoy immunity from liability for damages for, *inter alia*, copyright infringement as long as certain conditions are satisfied.[12]

9 Thus, for example, under Art 14 of the E-Commerce Directive, where an information society service provider stores information provided by a recipient of its service (i.e. it functions as a "host"), it will not be liable for damages if it (i) has no actual knowledge of illegal activity and is not aware of facts or circumstances from which illegal activity is apparent and (ii) acts expeditiously to bring any illegal activity to an end on receiving such notice.[13] The scope of this provision is contested and some have argued that, while it might have been appropriate to grant such an immunity in the early years of development of networked electronic communications, online platforms, such as YouTube, now make vast profits through the hosting of unlicensed copyright materials posted by users and have no need for such shelters from liability. Critical concerns have been exacerbated by the Court of Justice's broad interpretation of Art 14 in cases such as *Google France*, in which the Court interpreted the hosting immunity as applying in circumstances in which a service provider lacks specific knowledge or control of stored data and fulfils a "merely technical, automatic and passive" role.[14] The European Commission is currently considering this issue within its review of the Union's copyright rules.[15]

10 The *Gedankenexperiment* advocates a nuanced approach to the legal responsibility of intermediaries:

"For an appropriate balance of interest, it seems necessary

6 Through its reliance on the Charter of Fundamental Rights in interpreting the copyright *acquis*. See, for example, (C-201/13) *Deckmyn v Vandersteen*, 3rd September 2014.

7 For discussion, see, for example, J Wang, "Not all ISP Conduct is Equally Active or Passive in Differing Jurisdictions: Content Liability and Safe Harbor Immunity for Hosting ISPs in Chinese, EU and US Case Law" [2015] EIPR 732; Federal Supreme Court (Bundesgerichtshof); 26 November 2015 – Case No. I ZR 174/14, "Germany: Disturber Liability of an Access Provider" [2016] IIC 481; A Gärtner & A Jauch, "*Gema v RapidShare*: German Federal Supreme Court Extends Monitoring Obligations for Online File Hosting Providers" [2014] EIPR 197; C Angelopoulos, "Beyond the Safe Harbours: Harmonising Substantive Intermediary Liability for Copyright Infringement in Europe" [2013] IPQ 253; M Leistner, "Structural Aspects of Secondary (Provider) Liability in Europe" [2014] Journal of Intellectual Property Law & Practice 75.

8 Directive 2000/31/EC of the European Parliament and of the Council of 8 June 2000 on certain legal aspects of information society services, in particular electronic commerce, in the Internal Market.

9 Art 12.

10 Art 13.

11 Art 14.

12 The provisions do not preclude the grant of injunctions against service providers conducting the specified activities (Arts 12(3), 13(2), 14(3). Information service providers also benefit from a prohibition on the imposition of a general obligation to monitor for infringement and/or a general obligation actively to seek facts or circumstances indicating illegal activity (Art 15).

13 Art 14(1)(a), (b).

14 (C-236/08 – 238/08) *Google France v Louis Vuitton* [2010] ECR I-2417. See also (C-324/09) *L'Oréal SA v eBay* [2011] ECR I-6011.

15 See European Commission, *Towards a Modern, more European Copyright Framework*, 9th December 2015, COM (2015) 626 final.

to differentiate between intermediaries whose offers tend to compete with goods and services from exploiters/creators or may even substitute them ("competing intermediary services"), and those whose offers complement the goods and services of rights owners or even make them possible in the first place ("complementary intermediary services")."[16]

11 Intermediaries that are "close to the content" (such as video and image hosting platforms) are considered to be more likely to compete with the offering of a creator/exploiter than intermediaries that are "far from the content" (such as, for example, technical internet access providers). Within the category of "competing" intermediaries, the *Gedankenexperiment* makes further distinctions. It recognises, for example, that the full range of legal remedies ought to be available against an intermediary that is concerned purely to freeride on the creative contribution of authors and exploiters. However, it is acknowledged that the situation of other online platforms is more ambiguous because, while they may cause prejudice to rightholders' distribution channels, they also promote public welfare in certain important respects.

12 In keeping with the approach that it adopts throughout, the *Gedankenexperiment* suggests that, in such circumstances, the interests of the various affected categories of actor must be balanced and that:

"A possible result of such a balancing act could be that providers of legitimate (potentially) competing offers... would be given an obligation to pay monetary compensation in lieu of their users, e.g. in the form of shares in revenue or an adequate compensation."[17]

13 Under such a system, which is acknowledged to bear similarities to some currently applicable mechanisms, an intermediary would have to decide itself whether to pass on the costs of such compensation to users or to finance the payment in other ways (presumably, for example, by advertising). In return for the assumption of an obligation to pay compensation, intermediaries would be completely relieved of monitoring obligations[18] and would be provided with immunity from liability for its users' infringements.

14 In truth, this analysis of the problem does not get us particularly far. On the face of it, the "close to the content"/"far from the content" distinction is more descriptive than analytical and the suggestion that the potential loss of revenue to rightholders could be made up through a balanced compensation system is not revolutionary. Nevertheless, the *Gedankenexperiment*'s strong commitment to the recognition and reconciliation of competing interests takes the notion of "balance" beyond rhetoric and, at least, establishes a foundation for the exploration of the problem of intermediary liability. By contrast, the second proposal published in this edition of JIPITEC addresses the same question but takes a very much less tentative and reflective position.

C. The Mission to Link Directives 2000/31 and 2001/29

15 The High Council for Literary and Artistic Property (*Conseil supérieur de la propriété littéraire et artistique*, CSPLA) is responsible for advising the Minister of Culture and Communications of the French Republic on matters relating to literary and artistic property. Created under legislative order, it has produced a number of reports on questions relating to authors' rights.[19] Its "Mission to Link Directives 2000/31 and 2001/29"[20] was presided over by Professor Pierre Sirinelli[21] and reported at the end of 2015.[22] The Mission, which consulted a number of stakeholders,[23] focused on two questions. First, it considered whether: "...[T]he regimes implemented by Articles 12 to15 of the E-Commerce Directive of 8 June 2000 (Directive 2000/31/EC) truly provide a full understanding of the activities of certain service providers (Web 2.0 in particular) who were barely in existence when this legislation was adopted?"

16 Secondly, should the answer to the first question prove to be negative, the Mission's role was to investigate potential solutions to the problem presented by the inappropriate application of the

16 *Berlin Gedankenexperiment*, 13 (footnote omitted).

17 Ibid, 14.

18 While Art 15 of the E-Commerce Directive precludes general monitoring obligations, some Member States have sometimes imposed more specific monitoring obligations on information service providers. See, for example, A Gärtner & A Jauch, "*Gema v RapidShare*: German Federal Supreme Court Extends Monitoring Obligations for Online File Hosting Providers" [2014] EIPR 197.

19 See <http://traduction.culturecommunication.gouv.fr/url/Result.aspx?to=en&url=http://www.culturecommunication.gouv.fr/Politiques-ministerielles/Propriete-litteraire-et-artistique/Conseil-superieur-de-la-propriete-litteraire-et-artistique/Travaux>.

20 The original site for the document reprinted on p. 88 of this volume is <http://traduction.culturecommunication.gouv.fr/url/Result.aspx?to=en&url=http://www.culturecommunication.gouv.fr/Politiques-ministerielles/Propriete-litteraire-et-artistique/Conseil-superieur-de-la-propriete-litteraire-et-artistique/Travaux/Missions/Mission-du-CSPLA-sur-l-articulation-des-directives-2000-31-et-2001-29>.

21 Université Paris-I (Panthéon-Sorbonne).

22 3 November 2015. Vice-Presidents of the Mission were Josée-Anne Benazeraf (lawyer at the Paris Bar) and Alexandra Bensamoun (Senior Lecturer, Université Paris-Sud).

23 Although "some technical service provider representatives opted not to respond to the mission's invitation" (*Mission Report*, 2).

E-Commerce Directive's safe harbors in the current technological context. The *Mission Report* builds on an earlier CSPLA report, likewise led by Professor Sirinelli, on proposals to revise the Information Society Directive.[24] That earlier report recommended that the E-Commerce Directive's immunities should be re-examined because of their negative effect on the holders of rights in literary and artistic property.

17 In these circumstances, it is perhaps not surprising that the first question ("Does something need to be done?") does not detain the authors of the *Mission Report* for very long. They note a near unanimous view among stakeholders that platforms' claims to immunity are problematic.[25] The Mission then goes on to consider the cause of, and potential solutions to, the problem that it has identified. It is highly critical of the Court of Justice interpretation of the scope of the Art 14 immunity as covering the activities of highly profitable platforms (described as "false hosting providers" in the report). The finding, in *Google France*, that information service providers fall within the safe harbour where they do not play an active role, so as to give them knowledge of, or control over, stored data, is characterized as an error of interpretation that should be remedied through European legislation. The Mission's preference in this regard is for the implementation of a copyright-specific solution rather than for a general revision of the E-Commerce Directive.

18 More specifically, it recommends the introduction of a new Article in the Information Society Directive (Art 9a):

> *"Without prejudice to Articles 12 and 13 of the Directive on electronic commerce, information society service providers that give access to the public to copyright works and/or subject-matter, including through the use of automated tools, do not benefit from the limitation of liability set out by Article 14 of said Directive.*
>
> *These service providers must obtain permission from the relevant rightholders as they, either alone or with the participation of users of their services, are implementing the rights set out by Articles 2 and 3.*
>
> *Such permission covers acts performed by users of their services when they send the copyright works and/or subject-matter to the aforementioned service providers in order to allow the access set out by sub-paragraph one, as long as these users are not acting in a professional capacity."* [26]

19 According to the *Mission Report*, this provision would restore a "better sharing of value" by distinguishing between service providers which purely "store" information for third parties (and would presumably still be covered by Art 14) and service providers which "give access to the public" to copyright works, and other protected material, and which would therefore be liable for infringement (with their infringing users).

20 Such a change would diminish the scope of the Art 14 immunity as it is understood today. Given that the Report's aim is to amend the Information Society Directive, it also seems possible that the introduction of the new Art 9a might risk expansion of the scope of liability well beyond the "YouTube"-type situation at which the *Mission* is ostensibly aimed. If the draft provision were introduced, the Court of Justice would have to wrestle with the relationship between the concepts of "giving access to the public", "making available to the public" and "communication to the public". As a result, it seems likely that the introduction of the proposed Art 9a would add complexity to an already confused area of jurisprudence.[27] As any shifts in the current situation would be likely to favour right-holders, this consequence might not be entirely unwelcome to the authors of the *Mission Report*. Nevertheless, they acknowledge that a diminution in the scope of Art 14 might cause difficulties for online intermediaries. Their proposed solution is, first, the introduction of transitional protection for intermediary business models developed on an expectation of immunity and, secondly, the implementation of a "duty of collaboration" between rightholders and service providers. It is perhaps rather ironic that, while the *Mission Report* is based on the assumption that the initially intended reach of the E-Commerce Directive's immunities is no longer appropriate in current technological conditions, the solution that it identifies is a reversion to legal orthodoxy, anchored by the authority of the Berne Convention,[28] a Treaty first agreed in 1886 and last revised in the 1970s.

21 By contrast with the *Gedankenexperiment*'s somewhat incomplete and speculative tracing of principle, the *Mission Report* is pragmatic and detailed. It is therefore not surprising that its faults are very different from those of the open-minded *Gedankenexperiment*. The disdainful rhetoric of the report leaves a reader with the strong impression that it might have been possible to predict the broad thrust of its conservative recommendations

24 Conseil supérieur de la propriété littéraire et artistique, *Report of the Mission on the Revision of Directive 2001/29/EC on the Harmonisation of Certain Aspects of Copyright and Related Rights in the Information Society*, December 2014.

25 "This affirmative response would have been unanimous but for the caution of certain technical service providers" (*Mission Report*, 3).

26 The proposed new Article is accompanied by two proposed new recitals (16a and 24a).

27 For recent interventions on this issue, see (C-160/15) *GS Media* (Opinion of AG Wathelet, 7th April 2016); (C-117/15) *Reha Training* (Court of Justice, Grand Chamber, 31st May 2016).

28 See *Mission Report*, 11, 12.

in advance of its consultation with stakeholders. It takes no account whatsoever of arguments that could be advanced in favour of a less conservative solution to the "value sharing" issue. Presumably, one of the underlying reasons for the Court of Justice's broad interpretation of Art 14 in *Google France* was its sense that some of the public benefits of technological advance might be lost if right-holders were granted unmitigated dominion over the activities of online platforms. The *Mission Report* does not engage with such concerns. Similarly, there is no mention of the fundamental rights framework upon which the Court of Justice has structured all its recent responses to intermediary liability. Through the application of the Charter, the Court has acknowledged the need to balance the right of property of copyright owners with the right to conduct a business of service providers and the right of freedom of expression, and access to information, of users.[29] In the *Mission Report*, no time is wasted on a discussion of this framework of competing rights or, indeed, on the due process rights of users under the "notice and take down" process facilitated by the E-Commerce Directive's safe harbours.

22 It is impossible to escape the conclusion that the Mission's predominant intention was to put down a marker for current discussions on copyright reform at European level. However, it seems unlikely that the European legislator will be entirely persuaded by its call to apply a right-maximalist form of regulation in the online environment. The "duty to collaborate" that the Mission envisages is surely too weak to offer adequate protection to online intermediaries (and, therefore, to the public interests that their activities support).[30] The strengthening of such a system based upon forced negotiations might, however, bring the proposal closer to the zone occupied by the *Gedankenexperiment*'s suggestion of an obligation to pay compensation or, indeed, to the current legal situation in which platforms pay a proportion of advertising revenues to creators and have negotiated licence agreements with bodies representing right-holders. I would argue that a half-appropriate solution to the problems presented by platforms' hosting activities is much more likely to be found through a difficult exploration of this contested zone than through any reassertion of doctrinaire copyright orthodoxy.

29 See, for example, (C-275/06) *Promusicae v Telefonica de Espana* [2008] ECR I-271; (C-70/10) *SABAM v Scarlet Extended* [2011] ECR I-11959; (C-314/12) *UPC Telekabel Wien GmbH v Constantin Film Verleih GmbH*, 27th March 2014. For discussion, see C Angelopoulos, "Tracing the Outline of a Ghost: the Fair Balance between Copyright and Fundamental Rights in Intermediary Liability" (2015-16) Info – the Journal of Policy, Regulation and Strategy for Telecommunications, Information & Media" 72.

30 The Mission Report has not been particularly well received by online intermediaries. See CSPLA, *Commentaires des organismes professionnels membres du CSPLA sur le rapport relatif à l'articulation des directives 2000/31 et 2001/29*, 2-5 (Response of ASIC, l'Association des Services Internet Communautaires). For an example of an interpretation of copyright rules in a manner that recognises to develop legal principles in the face of technological change, see the recent Opinion of AG Spuznar in (C-174/15) *Vereniging Openbare Bibliotheken*, 16th June 2016.

* *Jonathan Griffiths*, BA (Oxon) MA is a Reader in Intellectual Property Law at Queen Mary University of London.

The Berlin Gedankenexperiment on the restructuring of Copyright Law and Author's Rights

– Creators – Exploiters – Non-commercial Users – Intermediaries –

by **Till Kreutzer et al.**

Project Lead:

Till Kreutzer, PhD, iRights.Lab, iRights.Law, iRights.info

Expert Panel:

Per Christiansen, PhD, Professor at FOM – Hochschule für Ökonomie und Management

Dirk v. Gehlen, Director "Social Media/Innovation" at Süddeutsche Zeitung

Jeanette Hofmann, PhD, Science Centre Berlin, Alexander v. Humboldt Institute for Internet & Society

Paul Klimpel, PhD, iRights.Law

Kaya Köklü, PhD, Max-Planck-Institute for Innovation and Competition

Philipp Otto, iRights.Lab, iRights.info

Mathias Schindler, Assistant to Julia Reda, MEP, formerly Wikimedia Deutschland e.V.

Leander Wattig, Blogger, Independent Consultant for Publishers, Lecturer at UdK Berlin

Project Advisor:

Tim Renner, Secretary of State for Culture, Berlin, formerly Motor Music

Recommended citation: Till Kreutzer et al., The Berlin Gedankenexperiment on the restructuring of Copyright Law and Author's Rights, 7 (2016) JIPITEC 76, para 1.

Preface:[1]

1 The concept for a regulatory system concerning creative goods, as elaborated within the Berlin *Gedankenexperiment*, dates back to the third initiative of the "Internet & Gesellschaft Collaboratory" (2010-2011).

2 Back then, a panel of experts took the initiative and developed a set of *Guidelines* for a copyright law in the digital world, which took the form of a regulatory system for creative informational goods[2] (see final report of the 3rd. Initiative, p. 99 et. seq)[3]. The idea was to develop a new form of regulation that would withstand the challenges of the year 2035 (that means a time perspective of about 25 years).

3 From the outset the *Guidelines* were considered a working hypothesis, which should be discussed, reconsidered and redefined. This task was taken up by a newly composed working group that concluded its work in 2015. With the Berlin *Gedankenexperiment* this group submits a thoroughly reviewed version of the *Guidelines* of 2011.[4] The *Gedankenexperiment* at hand is thus a further stage of development of the first version of the *Guidelines*.[5]

4 The project´s work was based on a "blank page approach". The project team was asked to imagine that copyright law had never existed and that we could, in fact, create a new legal concept, fit for the digital world. We tried to prevent that the conceptual work would be influenced by the current legal situation, e.g. by the status quo and the boundaries arising from international treaties and other frameworks. Nevertheless, the considerate reader will notice that there are a number of references to present deficits and comparisons to the current copyright regime. In some cases these references might apply more to authors´ rights regimes than to copyright regimes, in others the opposite might be true. It should be noted, however, that the criticism of the present situation is not the focus of this paper. It is rather a means to illustrate and underpin the arguments for the newly suggested approach.

5 Akin to the *Guidelines*, the *Gedankenexperiment* constitutes a contribution to the discussion on the restructuring of copyright law in the digital age. However, the *Gedankenexperiment* does not come up with concrete legislative proposals; instead it outlines a regulatory concept. In particular, it is meant to determine actors and their particular interests and assign them roles and abstract rights and duties.

6 Within this framework, suggestions for essential, regulatory aspects of copyright law are made, for example regarding exclusive rights of creators and exploiters[6]. However, no specific evaluative decisions regarding points of detail are made. To answer these is not the aim of the *Gedankenexperiment* – but can and should be the task of follow-up projects.

A. Preamble

I. Regulatory Purpose

7 The *Gedankenexperiment* serves to conceptualize a framework of an institutionalized balance of interests in the form of a regulatory system that governs the creation, commercialization, use and mediation of creative goods. This regulatory system will foster creativity, art, culture and entertainment and thereby serve the public as a whole. To achieve this aim, individual interests will be guarded by protective, partly exclusive, rights. Guaranteeing such rights is meant to secure income opportunities for creative professionals and provide incentives to invest in creative, immaterial goods.

8 These protective rights, however, have to be implemented in a way that pays due regard to the public interest and to other conflicting interests. In other words, a balance needs to be struck, in which the respective interests receive the greatest possible consideration.

9 The regulatory system is thus more than a means to protect individual rights. It is an instrument that will allow the balancing of diverging interests, which may from time to time collide with one another. Akin

1 This paper was initially written in German. Translation into English by Sylvia Jakob and Till Kreutzer.

2 Original: „Leitlinien für ein Urheberrecht für die digitale Welt in Form eines Regelungssystems für kreative informationelle Güter".

3 <http://dl.collaboratory.de/reports/Ini3_Urheberrecht.pdf>. An English translation of the guidelines (not the whole report) can be found at: <http://www.collaboratory.de/w/Datei:Ini3_Copyright_Excerpt.pdf>.

4 The present project was sponsored by the Internet & Society Collaboratory e.V. (Co:Lab) with 10.000 Euros. The amount was spent on workshops, travel expenses and organization. No expert was paid a fee. We thank the steering committee of Co:Lab for its support. Co:Lab was not involved in the production of the contents and did not influence their development. The present publication reflects the personal opinion of the experts involved and is not necessarily consistent with the position of Co:Lab or of any of the respective institutions for which the experts work.

5 For a better understanding of the *Gedankenexperiment* it is recommended to read the original guidelines.

6 In this context the term „exploiter" is used for commercial users of protected creative goods. These are especially traditional producers such as publishers, labels or distributors and on top of that anybody who needs a license to use a protected work.

to the principle of balancing fundamental rights the *Gedankenexperiment* assumes that none of the involved interests enjoys a structural superiority. Whether one interest should be given priority over another needs to be determined in concrete evaluative decisions. In this context we assume that every legal position assigned by the regulatory system will have to be justified per se. This applies as much to the legal positions of non-commercial users, as to those of creators, exploiters and intermediaries[7]. A right will thus only be granted to the extent that it is justified in relation to the public interest and conflicting individual interests of the other parties involved.

II. Regulatory concept

10 The regulatory system for creative goods distinguishes four relevant groups of actors: creators, exploiters, non-commercial users and intermediaries. Their corresponding rights and duties are separate and separable, thus creating isolated spheres, within which the interests of the differing groups will be assessed. Thereby, distinct rights of creators, exploiters and non-commercial users accrue.

11 Due to the systematic separation, all actors will obtain distinct legal positions, which do not overlap and may not be re-assigned or licensed. The framing of each legal position is based on a thorough assessment of the individual situation of the respective actor. Thereby, one major flaw of current copyright regimes, as identified by the expert group, shall be prevented: If an actor assigns her legal position to another actor, the purpose of the protective right is undermined, as these rights were tailored to the needs of the first actor. This happens often in the relation between authors and exploiters who often have and pursue different interests. The balance of interests is thus jeopardized. A protective system based on these premises cannot establish a true balance of interests, as it is impossible to predict who ultimately owns the rights and who may exercise them.

12 The current copyright law systems, notably the European authors´ rights regime, suffer considerably from this flaw. Officially they are guided by the interests of the author. Political decisions are generally justified in the author's name. Practically, however, most legal positions accruing from the right of the author lie in the hands of exploiters. These legal positions, and in particular copyrights and exploitation rights, are in many cases to a greater or lesser extent contractually assigned or exclusively licensed. Exploiters, however, have other interests as to the exploitation of the work, i.e. in many cases they exercise the rights differently than the author herself might have exercised them. In addition, exploiters have their own legitimate interests for protection, which differ from those of the author and vice versa. Confusing and mixing authors' and exploiters' interests opens space for manipulative arguments, which may foster undesirable developments. These lead to a conflict of values, which may ultimately undermine the copyright law system as a whole.

13 To prevent such inconsistencies the *Gedankenexperiment* abandons the concept of derived rights and assignments. It proposes a strict distinction between authors´ and exploiters' rights. The authors´ rights will focus on the creator and will not be able to be licensed or assigned to exploiters. The latter, in turn, have rights of their own rights, which are tailored to their specific economic needs and which – opposed to authors´ rights – may be traded on the market.

14 The concept resembles the system of copy- and neighboring rights in the current authors´ rights systems. However, there is a fundamental difference due to the inability to re-assign or license legal positions of the author.

15 According to current law it is not only possible, but also necessary for a exploiter to obtain the rights to exploit a creative work from the author. In addition, exploiters have their own neighboring rights e.g. in a film production. Thereby different protective rights accumulate in the position of the exploiter.

16 An example: A composer licenses the copyrights in her work exclusively and for the whole duration of the copyright to a music label. The label thereby obtains the exclusive right to exploit the work, potentially for the coming 100 years or more, depending on how long the author lives. The work, however, is blocked for the exploitation by third parties, unless the copyrights owner (here the label) decides to allow it. In addition, the label has exclusive rights in the recordings that were produced from the composition.

17 Under the *Gedankenexperiment* such rights accumulation is neither possible, nor necessary.[8] Instead, its regulatory system suggests a different approach:

7 "Intermediaries" stands for actors who are not *using* protected material in terms of the law. They do not copy or distribute protected goods but provide the technical infrastructure that enables others (their users) to do so.

8 An exception applies only to those cases in which the creator is simultaneously the exploiter (e.g. in the case of self-publishers).

18 The creator can only grant the label a permission to exploit the work[9] for a limited period of time (e.g. five years[10]). This permission might also be exclusive. During this time the exploiter's initial investment is protected by two legal positions: (1) The contract with the author that prevents her from giving further permissions to publish her work to other exploiters. (2) The exploiter's own right in the published work, which enables them to take legal action against free riders. After the creators' permission has expired, she can allow another producer to use the work or exploit it herself. The exploiter, in turn, can continue marketing their own edition/production of the work, although they might have to cope with competition by other productions. The creator will not only benefit from the first, but also from any following publication by means of contractually agreed remunerations and statutory claims for monetary compensation. Society as a whole, in turn, profits from the resulting free competition.

B. Rights and duties of the creator

19 One of the main principles of the regulatory system is that the creator of a work should own the right in her creation (author's right)[11]. The right is bound to the individual and cannot be assigned nor licensed to third parties, neither exclusively nor non-exclusively, neither entirely, nor in parts. It can only be owned by a natural person.

I. Objective and scope of protection of the author's right

20 The work will be protected as an intangible commodity (as under current copyright law). The exclusive ownership of the work that follows from the author´s right includes both economical and moral rights. The author can therefore decide, if and who may use/exploit her work and under which conditions. The scope of protection ends where interests of third parties – in particular those of the general public – prevail.

II. Relationship between the author's rights and the exploiter's rights

21 In principle, both exploiters and creators have their own rights, which they can exercise independently. The creator cannot assign or license any usage/exploitation rights to the exploiter. Nonetheless, she may grant the exploiter an all-encompassing or limited permission to use/exploit the work for a limited period of time.

22 This is a contractually declared permission to first publish the work or use/exploit the work exclusively and/or non-exclusively for a limited period of time. Legally, the permission does not constitute a license/transfer of individual author's rights *in rem* but a consent under the law of obligations, i.e. under contract law. It can, if the creator agrees, be transferred to third parties. Should a creator grant a exploiter an exclusive permission to use or exploit the work (exclusivity agreement), she can only do so for a limited amount of time. The exclusivity agreement expires, depending on which point in time comes first, either after its maximum term of duration (e.g. 5 years) or with the expiration of the exploiter's right in her own edition/production. Should the exploiter's exploitation right be extended by means of registration (see below IV.4), the exclusivity agreement can also be extended. However, in order to prevent the creator from entering into promises of exclusivity without being able to predict their overall effect on her interests at a later stage, any extension should be accompanied by an additional agreement. For works *made-for-hire* or those which are part of more complex works such as films exceptions could be made from the aforementioned restrictions on the transferability of author´s rights.

23 If the exclusive permission to use/exploit is not extended, it may continue on a non-exclusive basis (depending on the agreement with the author). This would ensure that the exploiter can continue marketing the product, although without being the exclusive exploiter of the work.

24 Promises of exclusivity should be registered in a public registry in order to provide legal certainty. Promises of exclusivity not registered by the exploiter cannot be enforced against third parties.[12]

25 If the exclusivity agreement is registered, one can assume that it is publicly known and effective against any third party independent of their actual knowledge of the exclusivity agreement. *The effect* of

9 See the details of this concept below, sec. II.2.

10 As far as terms of protection or other terms are named in the *Gedankenexperiment* they need to be understood as variable. How long individual protective positions shall be granted or promises of exclusivity be valid, would have to be determined considering various factors, e.g. economic and demographic ones, etc. Such valuation is not part of the Gedankenexperiment.

11 Note the difference: "copyright" versus "author's right".

12 This means the exploiter may use or exploit the work. Without registration, however, she cannot prevent a third party from using or exploiting the work even though the creator promised her exclusivity.

a registered exclusivity agreement is thus similar to exclusive licenses under current copyright law but the negative ramifications of extensive licensing for the author are prevented.[13]

26 Should a competitor violate a registered exclusivity agreement, the exploiter can pursue them on the basis of either unfair competition law or the right accruing from their own production. In turn the creator who violates the exclusivity agreement by allowing a third party to use or exploit the work can be pursued under contract law.

27 **Examples of the effects of the permission to use or exploit a particular work:**

- **Example 1:** An author writes a novel. She grants a publisher the permission to first publish the work and promises him exclusivity for the maximum duration possible (e.g. 5 years). The publisher's exclusive right in his own edition (i.e. the proofread version of the book) might also last for five years since the date of first publication.[14] The exclusivity agreement thus expires simultaneously with the publisher's own right. After its termination the exclusivity agreement automatically transforms into a non-exclusive permission to use or exploit the work. From the date of termination, the author can allow a different publisher to re-publish her novel.

- **Example 2:** (variation): The author promises exclusivity but only for a period of three years after the first publication. Subsequently the creator can allow a different publisher to re-publish her work. This re-publication does not interfere with the first publisher's own exclusive right since it does not have the exclusive right to market the work (i.e. a novel), but only an exclusive right to its own edition of the novel (i.e. the version proofread and arranged by the first publisher).

- **Example 3:** A singer-songwriter allows a record label to produce, publish and market her song for a period of three years exclusively. The label's own exclusive right in the production has a duration of (assumingly) ten years. After the exclusivity agreement expires, the label's own right continues to exist for another seven years. During these seven years the label can exploit its production on the basis of its own exclusive right. After three years however the composer can allow a different label to create a new production, which will again acquire rights of its own. In the years after the expiry of the exclusivity agreement, both productions compete on the market.

- **Example 4:** A film director allows a film producer to exploit her work for a period of three years exclusively. The producer's own right in the production amounts to (assumingly) 10 years. Since the creative achievement of the film director will be inextricably interwoven with the contributions of the other participants in the production (e.g. cameramen) as well as being an inextricable part of the film production as a complex work, the terms of the exclusivity agreement and the producer's own rights should be synchronized. The director will in any case not be able to permit other producers to use her particular creative achievement. For cases like these special rules are needed.

III. Additional protection through copyright contract law

28 Although the power imbalance between creators and exploiters can be mitigated to a considerable degree by restricting extensive exclusive agreements, the need to protect creators against contractual overreaching still subsists.

29 There is, for instance, still a need to guarantee the author an appropriate remuneration and to protect her from entering into excessive promises of exclusivity. Hence there is a need for a strong, albeit balanced, copyright contract law.

30 Furthermore should creators be given the unlimited possibility to grant the general public far reaching, non-exclusive, indefinite and royalty-free permissions to use the work in order to ensure the functioning of open source and open content licenses.

IV. Protection of the alimentation interests through rights to a fair share and adequate compensation

31 In order to safeguard the interests of the general public, the term of protection for exclusive author´s rights should be restricted adequately. Excessively

13 For example the negative consequences of rights diffusion are avoided, *inter alia* because the exclusivity agreement only applies as long as the exploiter's own rights exist and because it can only be exercised vis á vis third parties if it has been registered. This approach prevents the erosion of values underlying the exclusivity right in copyright by avoiding unimpeded, unlimited transfer of rights or licenses to the exploiter.

14 See in relation to the proposition of time frames in this *Gedankenexperiment* the annotation in fn 10.

long exclusive rights obstruct the use of intellectual creations and generate legal uncertainty for re-publications or may even render them impossible since it is increasingly difficult to ascertain who owns the copyright.

32 This leads to the phenomenon of orphan works, excessive pricing and under- or non-usage, caused by the exclusive right´s artificially elevated transaction costs (licensing efforts and royalties). This applies first and foremost to digital exploitation, for which the marginal costs are so low that the likelihood of a re-use by third parties after the expiry of the protection period would normally be very high. However, if exclusive rights are granted excessively and their term of protection is determined too long, the likeliness of re-publication and further use will decrease significantly since there is no economic incentive to re-publish the work.

33 If the transaction costs to identify the rights owner or concluding individual license deals are too high for commercial disseminators or public institutions, many cultural works will disappear sooner or later although there might be a strong cultural interest in their availability and preservation. The protection of the public interest as the main goal of the legal system outlined in the *Gedankenexperiment* requires adequate restrictions of the exclusive rights.

34 As such a more reasonable duration for exclusive author´s rights than it is provided in current copyright systems is proposed. Instead of excessive exclusivity the *Gedankenexperiment* proposes a less invasive instrument of protection for authors. After the expiry of exclusive rights the creators should be legally entitled to profit sharing. Such claims particularly aim to protect the economical interests of the creators.

35 Any author should have a claim to a fair share of the commercial revenues derived from the exploitation of her work. Such a share would allow the creator to participate economically in her work's success, but prevents inflated licensing transaction costs. The difference to exclusive rights is obvious: whoever wants to use or exploit the work can do so without seeking prior permission and may not be prevented from doing so but has to pay the creator a monetary compensation. Hence the monetary compensation has only a narrow limiting effect on the use. Functioning structures to collect and distribute such remunerations provided, the burden for using the material would be manageable. A well-balanced system of exclusive rights and rights to monetary compensation enables an adequately refined protection of the creator and at the same time serves all other parties involved.

V. Protection of idealistic interests – the author's moral rights

36 The creator has moral rights, which protect in particular her interest in being recognized as the author, but also protect the work against distortions and non-authorized primary publications. These moral rights are based on other requirements than the economic exclusivity rights. Therefore they have an independent term of protection and may – similarly to claims for monetary compensation – under certain circumstances be bequeathed.

VI. Duration of the author's rights

37 When assessing the duration of author´s rights, it is important to consider their independent components. First and foremost a distinction should be made between economic protection rights and the author's moral rights.

38 The author's moral rights are inheritable; their duration can be determined (as under current copyright law) by the creator´s lifetime. Moral rights should have a uniform term that should be determined balancing the different interests.

39 For economic exclusive rights and claims for monetary compensation, different terms of protection are proposed. Above all, the exclusive rights allow the creator to control the first publication of her work and the negotiation of the best possible conditions. Their duration should not depend on the category of the work and be assessed according to the principle "as long as necessary, as short as possible". Since the purpose of the exclusive rights is first and foremost directed at the primary publication, their term of protection should be calculated from the moment of first publication.[15]

40 After a certain period of time exclusive rights are transformed into a claim for monetary compensation. This approach allows the creator to continue participating in the economic exploitation of her work. For employed creators exceptions can be made. The claims for monetary compensation should be calculated autonomously and objectively and be equal for all types of creators and works.

41 Since the author's exclusive exploitation rights (and in particular the claims for monetary compensation)

15 That does not mean that the rights only accrue with the first publication, but that the duration of the exclusive rights will only be counted from that point in time. That way it cannot occur that the exclusive rights have already expired, when a creator decides to publish her work long after its creation.

will also serve alimentation interests, they should be inheritable.[16] In order to prevent increasing legal insecurity their duration should not extend beyond a reasonable period. They could, for instance, be limited to 20 years after the creator´s death.

C. Rights and duties of exploiters

42 In the digital world the exploiter still assumes a leading role. He invests in the production and distribution of creative goods and services and is, in many cases, the first to make them available (e.g. film productions). Based on the underlying idea of investment protection for distinct achievements, he ought to obtain an exclusive right on his specific design, edition or production (i.e. his "edition of the respective work"). Being an exclusive right, the exploiter's right is meant to safeguard the amortization of investments and provide financial incentives to make them. The exploiter´s right accrues for the exploiter and is not a right derived from the creator. It can be assigned to third parties – as a whole or in parts. Exploiters obtain the necessary permission to use the work from the creator via promises of exclusivity or non-exclusive permissions (for details see above B.II).

I. Subject matter of the exploiter's right

43 It is the individual achievement of the exploiter which is being protected, i.e. the particular production, edition or issue of the respective work. The subject matter of the exploiter's right is thus the specific implementation of a creative product (e.g. a music recording or a movie).

II. Accrual of the exploiter's right

44 In order to prevent unjustifiable disadvantages for small exploiters, the exploiter's right should accrue automatically and without registration when their edition of the work has been produced. The right accrues, similar to today's neighboring rights, for the one who has made the essential investments and carries the financial risk. The publication of the product is, however, only allowed if the creator had previously granted the necessary permission to use and exploit his work.

III. Subject matter and scope of the exploiter's right

45 Exploiters have an exclusive right to their product, i.e. in their own specific implementation, edition, assembly, production. They can transfer that right and assign it fully or in parts to third parties or grant exclusive or non-exclusive licenses (as long as the creator agreed to grant a transferrable permission to use her work). Following therefrom, the concept of the exploiter's right corresponds to the concept of neighboring rights. Exploiters can enforce their right against piracy and other non-authorized uses of their goods and services based on their own right. Being a right owner they can claim injunctive relief, removal and damages. Should a competitor violate a registered exclusivity agreement, the exploiter can resort to competition law.

IV. Terms of protection of the exploiter's right

46 The exploiter´s right constitutes first and foremost a protection of investment. Its term should therefore be calculated based on economic facts and assessments. Given that the product of the exploiter enjoys a quasi monopoly that interferes extensively with free competition, the principle of "as long as necessary, as short as possible" should be applied in this context.

47 As to the specification of the term of protection two models can be considered, both of which come with advantages and disadvantages. On the one hand it might be possible to define product specific terms of protection, in order to account for the different diffusion curves on the markets.[17] The advantages of market-oriented differentiation are, however, offset by serious disadvantages as to legal security. The convergence of traditional and new types of works based on multi-media, such as computer games, would raise considerable problems in this model. In addition, in a model based on market-oriented differentiation the terms of protection would have to be continuously changed in order to be able to account for the changing market conditions.

48 In light of these disadvantages it seems that an approach applying uniform terms of protection – i.e. terms of protection which are independent of product and market – should be preferred. In order to calculate an ideal uniform term of protection, one should – just as in patent law – consider the average

16 Even though the duration of the exclusive rights is measured relatively shortly, they should be heritable. In any case it is possible that the creator dies shortly after the creation of the work.

17 This approach was preferred in the original guidelines, see final report, p. 117 (<http://dl.collaboratory.de/reports/Ini3_Urheberrecht.pdf>).

amortization period as a guideline. In other words, the term would have to be determined in line with the period of time in which investments into creative products usually pay off.

49 In order to prevent (in individual cases) inadequately short terms of protection, they should be able to be extended by registration. The registration should be subject to considerable yearly fees that should progress in amount the longer the protection lasts. The opportunity to extend the term of protection should be limited in time (e.g. 20 years after publication) in order to avoid the negative effects of excessively long exclusive rights and to reconcile the interests of the general public with the interests of investment protection on the side the exploiters. The registration fees can be spent on cultural purposes or similar.

D. Rights and duties of non-commercial users

50 The duties of non-commercial users[18] ensue from the protection rights of creators and exploiters. Within the scope of the exclusive rights of creators and/or exploiters, the work may not be used without permission. Should legislation guarantee the non-commercial user a right to use the work or the respective production, these rights prevail over the exclusive rights. In other words the exclusive rights are limited in scope.

I. Non-commercial users rights as enforceable personal rights

51 Different from the current system of exceptions and limitations especially in author's rights systems, the *Gedankenexperiment* protects the interests of non-commercial users by own rights. These user's rights are not derived from statutory limitations of the author's or exploiter's rights. They rather are separate and individual legal positions that belong to the non-commercial users.

52 The balance of interests between creators, exploiters and non-commercial users is realized by defining the scope of the exclusive rights on the one hand and the user´s rights on the other. The non-commercial user's interests are thus not subject to limitations of unlimited exclusivity rights, but constitute own subjective rights. They are outside the scope of protection.

53 All rights – i.e. those of the non-commercial users, the creators and the rights owners – are generally considered equal. Hence, the freedom to use is no exception to a general overall exclusivity. Consequently, the non-commercial users´ rights do not form part of the protective exclusivity rights. As such, they cannot be contractually excluded.[19] Being individual rights they can even be enforced, e.g. when citations of films are impossible, because the exploiter sells his copies with technical copy protection measures. Such a system would acknowledge the public interest in legally protected freedoms to use and affirm them with strong legal positions.

II. Relationship between non-commercial users', authors' and exploiters' rights

54 The creation of an independent statutory sphere for non-commercial users´ rights establishes a consistent systematic approach towards balancing of the opposing interests. In addition, this approach prevents inconsistencies by guaranteeing that non-commercial users´ rights are considered equal to the creator's and the exploiter's rights. Should a non-commercial user have a right of citation, she can enforce that right against both creators and exploiters.

III. The implementation of non-commercial user rights

55 As described in the original *Guidelines*[20] non-commercial users´ rights should be implemented by means of a regulatory system, which combines elements of both the continental European author´s right and US copyright law.

56 In practice, this means the establishment of a rule catalogue of typified usage rights (such as the limitation provisions under the present author's right system) in which permitted acts are specifically described, such as the quotation right or private copying. This catalogue however is – other than

18 Non-commercial users are referred to as users who use protected creative goods for their private or the public interest. Non-commercial users are individual members of the public and public institutions like museums, universities, archives and the like.

19 In this model an exclusion of usage rights would lead to the contractual creation of protective rights and the extension of existing protective rights that are not envisaged by the law. Such a "law-perverting" drafting of contracts should (under certain circumstances) be expressly prohibited by law.

20 See p. 112 et seq. of the final report (< http://dl.collaboratory.de/reports/Ini3_Urheberrecht.pdf>).

under the current European copyright law – not exhaustive. It will be complemented with an open norm, i.e. a general clause for constellations, which do not fall under the rule catalogue.

57 This regulatory system ensures on the one hand side legal certainty and on the other, the necessary flexibility in view of the ever-changing digital environment, which may easily upset the balance of interests. As far as it seems appropriate, monetary compensation through levies should be introduced to compensate certain forms of free uses.

58 The open norm only applies to uses which fulfill the criteria of a proportionality test. This could be guided by the four-step test set out in Art. 107 US Copyright Act (fair use).[21] In order to further substantiate the freedoms to use not explicitly mentioned in the legislation and to enhance legal certainty, complementary regulatory means could be established. It might, for instance, be possible to create a regulatory authority which defines permitted acts of usage under the open norm and lays down binding conditions for their applicability (e.g. the duty to pay adequate compensation). It could also assess whether existing freedoms to use are still necessary.

59 Alternatively or even cumulatively, certified user associations could be allowed to take representative action. Legitimized associations could apply for a generally binding decision in front of specialized courts determining whether, and if so, under which conditions an act of use falls under the open norm and thus constitutes a non-commercial users' right. A combination of regulatory instruments, courts, regulatory authorities and legislators would be involved in the further development of the balance of interests, which is so important for the effective regulation of creative goods. The involvement of all these powers in the development of the law can considerably accelerate the advancement of regulation and avoid protracted backlogs of reforms.

E. Rights and duties of intermediaries

60 Intermediaries are those which engage in the widest sense in the storing, making available and searchability of creative goods and services, without being commercial or non-commercial users or creators.[22] These include platform providers, web-hosting and web-sharing services, search engine providers, electronic program guides or similar services. Also telecommunication providers are in this broad sense "intermediaries".

61 Current copyright law does not govern the interests of Internet Service Providers. Since they are – according to jurisdiction and legislation – neither right owners nor do they engage in acts of use relevant under copyright law[23], their rights and duties are regulated outside the law of copyright. Instead they are regulated under Telecommunication Laws and especially under ISP liability rules found e.g. in the E-Commerce Directive of the European Union (2000/31/EC) or the US Digital Millennium Copyright Act.

62 In a new regulatory system for creative goods intermediaries are of systemic importance. That does not mean that intermediaries should be given protective rights of their own. It is a rather a question of balancing their interests against the interests of the other stakeholders involved.

63 As much as the label "intermediary" covers very different services and constellations, it may generally be said that intermediaries can exert an important influence on the commercial exploitation and use of creative goods. This becomes all the more obvious in the case of search engines and platforms. The activities of intermediaries can have a positive impact on creators, exploiters and non-commercial users, for instance when they enhance the findability and visibility of works and thus make them more accessible for a larger target group.

64 The activities of intermediaries, however, can also have a negative impact, for instance when free unlicensed services (such as user generated content platforms) start competing with fee-based offers provided by the right owners, or when an important intermediary obstructs access to creative goods.

21 According to Art. 107 of the US Copyright Act, four factors need to be taken into consideration when assessing whether an act of use constitutes fair use: (1) The purpose and character of the use, including whether such use is of a commercial nature or for non-profit educational purposes; (2) The nature of the copyrighted work; (3) The amount and the substantiality of the portion used in relation to the copyrighted work as a whole; and (4) The effect of the use upon the potential market for or value of the copyrighted work.

22 For the definition see also fn. 7. An intermediary is thus not somebody who uses works or creative products in terms of the regulatory system (or: copyright) and is thus not subject to licensing obligations. The scope of application of the intermediaries' regulations will therefore have to be defined according to the acts of use and not according to the identity of the user. Since in the online world content and infrastructure services are increasingly converging, it is very likely that providers are in some cases both users, intermediaries and under certain circumstances exploiters. The concrete regulations refer – as is the case under the current copyright law system – to the particular activity.

23 E.g.: A hosting provider does not copy protected works. It rather provides the technical facilities that are used (by end users) to copy.

65 The current legal system has difficulties in defining the role of intermediaries, since the economic principles of an intermediary´s role cannot be captured in a differentiated way under the current mechanism.

66 Under current copyright law, an act is either an act of use not relevant to copyright law – and thus neither requiring permission nor remuneration – or it is an act of use to which copyright applies in its entirety. Economically, intermediaries often lie somewhere in between. Although the intermediary does not "use" the provided goods in the legal sense, he profits nonetheless to a considerable extent of their being made available by third parties (his users), since his services would not be attractive without them.

67 On the other hand the commercial success of intermediaries is mostly based on their own achievements. They render a – for exploiters and authors – free service, which can also be beneficial for them. User generated content platforms such as video services can, for instance, significantly enhance the visibility and potential popularity of the uploaded creative content. The result is a marketing effect that comes for free for the right owners. In short: intermediaries are both beneficiaries and *apporteurs* of benefits in relation to creative goods. In a regulatory system for creative goods both roles need to be captured adequately.

I. The intermediary in the regulatory system for creative goods

68 At a glance, three regulatory areas for intermediaries can be considered in a regulatory system for creative goods:

- A precise definition of the term "intermediary";
- Whether authors and/or exploiters should have a direct claim to a share of the profit or other kinds of remuneration against (certain kinds of) intermediaries (primary liability);
- Whether intermediaries should be liable for right infringing acts of their users (secondary liability).

II. Categories of intermediaries

69 The potential for a conflict of interest between intermediaries, creators and exploiters depends most importantly on the kind of intermediation. Different types of intermediaries can be distinguished. Not all of them are relevant for the regulatory system for creative goods. Only those conflicts need to be solved which may have a negative impact on the interests of the other actors of the regulative system. As far as offers of intermediaries have a positive impact or can be rated neutrally in this regard, no legislative intervention is necessary.

70 The fact that an intermediary facilitates in some way the use of creative goods does not, on its own, point to a conflict of interest. In order to distinguish relevant conflicts of interests from irrelevant causal chains, criteria are necessary, which allow an abstract and general assessment of potential conflicts of interest which may ensue vis-à-vis particular types of intermediaries and which should carry different legal consequences.

71 For an appropriate balance of interest, it seems necessary to differentiate between intermediaries whose offers tend[24] to compete with goods and services from exploiters/creators or may even substitute them ("competing intermediary services"), and those whose offers complement the goods and services of rights owners or even make them possible in the first place ("complementary intermediary services").

72 Complementary services complement the goods and services of exploiters and foster the use and reception of works and are thus generally neutral or even beneficial for right owners. From the overall perspective of an information society they fulfill the important function of lowering information costs.

73 Competing services on the other hand can constitute a threat to the offers of the rights owners. In contrast to complementary services, they profit directly from the use of protected material. Although they do not *use* protected material themselves (in that case they would be non-commercial users or exploiters), they generate added value, e.g. by providing their users with services outside the marketing channels envisaged by creators and exploiters. That may even lead to the substitution of goods and services of the right holders. It thus appears generally necessary to consider intermediaries with potentially competing offers in the balance of interests which the regulatory system for creative goods envisages.

74 An important indicator for the division between competing and complementary offers is the

24 „Tend to", since in special cases every offer may have the contrary effect. A video platform, for instance, may be beneficial for the rights holder due to the advertising effects and the increase in publicity; for others on the other hand it might be detrimental due to its competitive impact. As always in the case of general assessments it is only possible to define general cases and to focus the regulations on them. For special cases exceptions to the general rule could be considered.

"proximity to the content". "Close to the content" are those services, which allow their users to place protected material online and to store and distribute it directly. By contrast services which only provide the means, especially the technical infrastructure, that enable the general use of the Internet are "far from the content". The same applies for services which merely systemize materials already available on the Internet and make them findable.

75 For instance, Internet access providers may be regarded as intermediaries in the sense of the definition, since their activities are a *sine qua non* for every Internet communication including the use of protected works. However, they only provide the technical means that make online communication generally possible. They are thus "far from the content". In addition, the offering of Internet access does not have a negative impact on the interests of creators or intermediaries. The offers of the right owners are thereby not substituted or impeded; instead they need access providers to enable the users to access their services. The fact that illegal acts of use are also being made possible does not change this fundamental assessment.

76 The same applies to other infrastructure and service providers or search engines. Neither do they compete with the offers of intermediaries nor do they replace them. They merely facilitate their searchability. They are therefore not competing against but complementing offers in the sense of the above categorization as long as they do not substitute or compete with the contents to which they point.

77 Counterexamples for services which are "close to the content" are, for instance, file hosters and video and image platforms. Such services can be operated in a way that competes with the offers of creators and exploiters or even replace them. Whether that happens depends on several other factors which cannot be examined in detail at this point, e.g. whether whole works or – as in the case of images – works in their full resolution are published on the Internet, whether access extends to everyone or only to a distinctive, limited user group, whether only own or also other people´s contents can be published, etc.

III. Legitimate and illegitimate intermediation and the consequences

78 Based on the proposed free differentiation the regulatory system for creative goods will be able to define appropriate legal consequences. Whereas intermediaries of complementary offers principally do not appear to require regulation, intermediaries of competing offers should invoke differently graded legal consequences.

79 Certain forms of competing offers may simply not be acceptable, for instance, when the service of intermediation is marginal and the true intention is to freeride and market the works circumventing exclusive rights. In this case the services should be subject to the full range of legal remedies, including injunctions and damages.

80 However, for most intermediaries the competing effect results from the attractiveness of the service, which represents an added value for the user. Such services are legitimate. There is – besides the own interest of the provider – a public interest in these services, so that they fall under the protection of the market and of the law. However, a balance of interest between intermediaries, the general public and the right owners, whose distribution channels may be prejudiced by such offers, needs to be struck.

81 A possible result of such a balancing act could be that providers of legitimate (potentially) competing offers – contrary to those offers providing complementary services – would be given an obligation to pay monetary compensation *in lieu* of their users, e.g. in the form of shares in revenue or an adequate compensation.

82 This could be justified, since they compete with the offers of the right owners or may even replace them, without having to invest in content or licensing. Their business model is based on the use of protected creative content by their users. Experience tells that especially non-commercial users do not only upload their own contents but also third party content – usually without having a license.

83 Unlicensed uses of non-commercial users provided through intermediaries are currently usually not economically compensated. This appears unacceptable in a regulatory system for creative goods. Experience has shown that copyright enforcement towards non-commercial users is not practical and leads to unwanted effects. Monetary compensation paid by intermediaries has the advantage that remunerations can be raised from a central actor and be passed on to the right holders. Decentralized licensing or royalty obligations towards end-users could also be prevented, avoiding mass enforcement against citizens.[25] Since the

25 The contemplated compensation paid in lieu by the intermediary does not necessarily lead to the conclusion that all acts of end-users have to be legalized. Both questions can in principle be assessed separately. Should it, however, emerge that compensatory payments are generally passed on to the users, but they are still being sued, legalization

intermediaries would only be called upon *in lieu* of the end-users, it would be in their discretion to ask users to reimburse them or to find other ways of refinancing[26] the payments made.

84 Such a solution resembles the systems of levies on blank media and storage devices, in which producers, traders and importers are liable for the payments *in lieu* of the end-users. In most European countries the system proved to be effective, at least in principle. In addition, the approach resembles industry specific solutions already established in some areas (such as YouTube's Content ID System)[27].

IV. Responsibility for legitimate intermediary offers with competitive potential

85 The inclusion of intermediaries into the remunerative relationship between right holders and non-commercial users is potentially contrary to their interests. Since their success is based on own achievements and self-created added value, and they do not engage in any act of use themselves, but only profit from the fact that their users do so, they should be granted advantages in return. This could be, for instance, the limitation of liability for use acts of their non-commercial users.

86 It might thus not be far-fetched to reduce the liability of intermediaries to purely reactive duties to act in a framework of notice and takedown procedures. This approach would most probably reduce the challenges currently faced by intermediaries in many parts of the world, where they sometimes face proactive duties to check uploaded content for its legitimacy and may even have to pay damages.

87 On the whole, such a system of monetary compensation and limitation of liability would be beneficial for all parties involved: instead of restricting the uses on platforms and hosting services, revenues would be generated for creators and exploiters. The intermediaries would be included in the remuneration system but not exposed to extended liability. Furthermore the approach would create more legal certainty, which is advantageous for all actors concerned.

would be inevitable. Otherwise the end-users would be charged twice for their uses and could still face legal consequences.

26 It would lie in the discretion of the intermediary to decide whether and, if so, in which form to make use of this opportunity. This would most likely depend on the business model and the technical feasibility.

27 The success of YouTube's Content ID System (see < https://en.wikipedia.org/wiki/YouTube#Content_ID>), which is voluntary for both sides, shows on the one hand that it can be more advantageous for intermediaries to make payments and share revenues than being exposed to legal remedies. In addition, they protect their users and maintain their business model. On the other hand, the active participation in the program (again voluntarily) by music labels and producers shows that such systems are regarded as advantageous also by the right owners. It would be interesting to examine, whether general lessons could be learned from this example. Should this be the case, it might be worth considering turning it into a regulatory approach.

Mission to Link Directives 2000/31 and 2001/29

Report and Proposals

by the **French High Council for Literary and Artistic Property**

Pierre Sirinelli, President

Josée-Anne Benazeraf, Vice-President

Alexandra Bensamoun, Vice-President

3 November 2015.

Recommended citation: French High Council for Literary and Artistic Property, Mission to Link Directives 2000/31 and 2001/29 – Report and Proposals, 7 (2016) JIPITEC 88 para 1.

1 In a mission letter dated 3 April 2015, the High Council for Literary and Artistic Property (CSPLA) expressed its wish for research to be carried out on "*proposing changes to current European Union legal provisions enabling the effective enforcement of copyright and related rights in the digital environment, particularly on platforms which disseminate protected content*".

2 The President of the mission will be Pierre Sirinelli, a Professor at the Université Paris-I (Panthéon-Sorbonne), while the Vice-President roles have been entrusted to Josée-Anne Benazeraf, a lawyer at the Paris Bar and Alexandra Bensamoun, Senior Lecturer HDR [accreditation to supervise research] at Université Paris-Sud.

Modus operandi

3 The mission began by setting up round tables in order to gauge the opinions of various professionals within the sector, both from the CSPLA and elsewhere. It continued by working on the proposals put forward by some of the contributors and has itself outlined some solutions. The draft legislation contained in this report is the result of these various discussions.

4 In addition the work of the mission – which met twice a week for several months – took many different forms including hearings, bilateral discussions, consultation meetings by sector or stakeholder category, cross-analysis of legislation and plenary meetings.

5 All sectors were heard, including the technical service providers, although some technical service provider representatives opted not to respond to the mission's invitation.

6 Work was performed alongside the work conducted in Brussels by the Commission and the Parliament, which recommended that the liability regime for some information society service providers[1] ought

1 The *harmonisation of certain aspects of copyright and related rights*, a European Parliament Resolution of 9 July 2015,

to be clarified in order to prevent these providers from capturing the value of the works which fuel their economy[2].

7 This mission's work has given rise to numerous discussions in this area with European institution representatives.

Questions

8 In a nutshell, the mission has been asked to resolve the following questions:

- Do the regimes implemented by Articles 12 to15 of the E-Commerce Directive of 8 June 2000 (Directive 2000/31/EC) truly provide a full understanding of the activities of certain service providers (Web 2.0 in particular) who were barely in existence when this legislation was adopted?

- If not, what solutions could be implemented in order to prevent some of the consequences of these statutes from being applied in the field of literary and artistic property?

9 In order to bring together initial impressions on the topic together with outlines of the answers to these two questions, the mission felt it necessary to ask each participant a series of simple questions:

1. Must we intervene in order to change the solutions adopted by certain courts in the absence of clear legislation providing a harmonised understanding of the new activities conducted by certain service providers?

2. If so, which activities need to be understood in order to be able to propose new solutions?

3. What form should this legislative change take?

4. What consequences should it have?

Positions expressed

10 The first question has been answered in the affirmative. This affirmative response would have been unanimous but for the caution of certain technical service providers.

11 The second question also gave rise to some main areas for consideration:

- A large majority of the respondents deemed that it would **not be appropriate to reform the legal regime for activities related to mere conduit, internet service provision or caching.**

- Moreover, there were no requests to revise the legal regime for hosting providers conducting activities which fully meet the definition proposed by Directive 2000/31 in the year 2000. This denotes a 'transparent' technical service provider which hosts content and remains **out of direct and intentional contact with the public.**

- However, **almost all of the contributors agreed that a clarification ought to be added to indicate that the regime proposed by Article 14 of the 8 June 2000 Directive should in no case be applied to what many professionals call 'false hosting providers'**, in other words, information society service providers whose role extends beyond that of a technical service provider as defined by the Directive. This includes certain **Web 2.0 platforms** (particularly **contribution-based or community sites**), **certain social networks**, and certain **services** that may be used by certain **search engines**. It should however be noted that although opinion was unanimously in favour of intervention for the former, more varied opinions were expressed on the latter. The proposed solutions were therefore considered in the light of the former parties, and are therefore not fully applicable to the other categories. Our thinking does not cover conduits, ISPs, cache providers or personal file storage services.

Basis of requests for change

12 The reasons put forward for the development of the solutions or clarification of the inadequacy of the manner in which the solutions set out by Article 14 of Directive 2000/31 have been applied for the above activities tend to be technical or economic rather than legal. Although all parties noted that the above activities do not match the assumptions made by the European authorities in 2000 when the legislation was drafted, it was also highlighted that **better sharing of value** may be gained through a

point 45, suggests "*a review of the liability of service providers and intermediaries in order to clarify their legal status and liability with regard to copyright (...)*"; *A Digital Single Market Strategy for Europe, European Commission Communication of 6 May 2015, p. 8: "In addition the rules applicable to activities of online intermediaries in relation to copyright protected works require clarification, given in particular the growing involvement of these intermediaries in content distribution.* "

2 *The harmonisation of certain aspects of copyright and related rights*, European Parliament Resolution as specified above, point O: "*whereas creative works are one of the main sources nourishing the digital economy and information technology players such as search engines, social media and platforms for user-generated content, but virtually all the value generated by creative works is transferred to those digital intermediaries, which refuse to pay authors or negotiate extremely low levels of remuneration.* "

clarification of the non-applicability of the European procedure to these activities.

13 This point does not require a long explanation:

1 - Claiming that they are covered by the conditional exemption from liability provided by Article14 enables the above service providers to make considerable profits due to the (unauthorised) existence of copyright protected works or items covered by related rights on their platforms. These services (whether charged or free-of-charge) generate considerable advertising revenues due to the existence of such works, yet consider that they do not need to seek authorisation and therefore redistribute a share of the profits made to the extent sought by rights holders. This is due to the fact that:

1.1. Either the courts wrongly apply this regime to information society service providers. It should be noted in this respect that in its *Google* judgments of 23 March 2010[3], the CJEU ruled that Article 14 of the E-Commerce Directive could be applied to a service provider that "*has not played an active role of such a kind as to give it knowledge of, or control over, the data stored*".

This condition, which is based on recital 42 of the E-Commerce Directive, is however a foreign concept to hosting providers. In addition it seems to have been distorted by the aforementioned assessment by the Court of Justice.

First and foremost, it is clear when reading recital 42 that it does not apply to hosting providers (which are covered by recital 46), nor to the search services in question in the case ruled on by the Court, but rather to conduits, ISPs and cache activities. The activity covered by the aforementioned recital 42 is indeed "*limited to the technical process of operating and giving access to a communication network over which information made available by third parties is transmitted or temporarily stored*".

Therefore, the condition (as expressed by the final sentence of the recital) under which one is covered by the exemption from liability based on a "*mere technical, automatic and passive nature*", applies to Articles 12 and 13 of the Directive, but not to Article 14.

Please also note that the vocabulary used by the Directive differs according to the service provider, i.e. "exemptions from liability" for conduits, ISPS and caches (recital 42), and "limitation of liability" for hosting providers (recital 46).

Secondly, the above-mentioned statement by the CJEU does not match recital 42, under which the "*mere technical, automatic and passive nature [...] implies that the information society service provider has neither knowledge of nor control over the information which is transmitted or stored*".

In recital 42, the Directive states that in order to occupy a passive role, the service provider must have neither knowledge of nor control over the information. But this does not mean that the service provider is necessarily passive just because it does not have knowledge and/or control of the information, or likewise that the service provider must have knowledge and control of the information in order to play an active role (as stated by the CJEU).

In other words, the condition set out by recital 42 is necessary but insufficient.

By turning the recital around, the CJEU has substantially changed its meaning, as its interpretation would mean that 'false hosting providers' would never actually be active, given that when users post content, service providers generally do not have knowledge of or always control over said content.

1.2. Or, rights holders refrain from enforcing their literary and artistic property rights on providers due to the cost of proceedings and the difficulty of implementing their exploitation monopoly in such circumstances.

⇨ It is therefore clear that the balance of power is not at all on the side of the rights holders, and that the economic and technical power as well as the high-profile nature of some of the providers makes it even more difficult to hold true negotiations, or even proper discussions between the two sides.

2 - The reasoning provided by rights holders is varied.

2.1. Some simply comment that the situation ought to be resolved purely for reasons of legal integrity, and that the application of Article 14 to the above activities is the result of twisted logic, a misunderstanding of circumstances or a manipulation of the legislation.

2.2. Others simply highlight the fact that the rejection of Article 14's conditional exemption system would enable **better negotiation of**

3 Case C-236/08 to C-238/08; also see CJEU, 12 July 2011, *L'Oréal et al. v. eBay*, case C-324/09; CJEU, 11 September 2014, *Sotiris Papasavvas*, case C-291/13.

the remuneration which is legally payable in order to provide access to copyright works or material. In other words, the objective of rights holders is not to prohibit their works from being posted online, but rather to ensure that they are in a position in which they can obtain improved compensation and sharing of value.

2.3. Other rights holders are less prepared to negotiate with service providers, and intend to recover the full scope of their copyright and related rights in order to ensure that they can continue with their own strategies without hindrance, and in order to be able to implement (at least initially) their own individual policies to make works available over networks.

14 To conclude, it is clear that there is a high demand for intervention in order to change current case law solutions, which are deemed to be both unfounded and unjust.

15 There are therefore two questions to consider:

1 - Which type of intervention should be carried out, and which means should be used (I)?

2 - What new solutions are sought (II)?

*****

I. TYPES OF INTERVENTION

16 First of all, one needs to look at the means to be used in order to develop a solution, and which reform instrument ought to be used.

Intervention at a European level

17 There are several possible routes, all of which involve a **European-level solution**. Of course, it would be easily possible to add numerous progressive changes to fill the voids of European Member State legislation, enabling national legislators to make their own clarifications or adjustments. Yet due to case law developments in many Member States, it was unanimously decided that it would be wiser to act at directive level in order to obtain a standard solution which would apply in all 28 Member States.

18 The question is whether this should be done by way of official drafted legislation or through a process of interpretation?

Preference for an official intervention setting out a clear standard

19 In a document[4] sent to the French Ministry of Culture (*Ministère de la Culture*), the audiovisual production sector initially expressed its preference for a **purely interpretation-based route**. Based on the statements that the existing legislation has advocated (recitals 40[5] and 48[6] of Directive 2000/31/CE), for the past 15 years, the existence of a **duty to act** for technical intermediaries in order to be able to block access to illegal content, some audiovisual professionals considered that "*the best way to address all of the areas that require improvement is to use a cross-disciplinary process of interpretation, the purpose of which would be to provide clarification to the concepts included in several European legislative texts on the subject of making works available online (2000/31/EC, 2001/29/EC and 2004/48/EC in particular). This could involve an interpretative communication by the European Commission*".

20 **This option failed to convince** a substantial majority of the contributors, who wish to see more clearly defined and effective action, so that the interpretation route could be used as a fallback solution should the adoption of specific legislation be unlikely to be implemented.

21 The **route of an official intervention setting out a clear standard** is justified both on a *positive* basis by the need to be able to rely on a provision which sets out the solution, in certain cases, of rejecting Article 14 of Directive 2000/31, and on a *negative* basis by the fear of the interpretation route being insufficient.

22 One of the fears expressed, apart from the risk of European judges retaining the disputed approach, lies in the fact that the interpretation route proposed to resolve the issue at stake here would involve clarifying the concept of a hosting provider as defined by Article 14 of the E-Commerce Directive.

4 *Propositions pour une meilleure efficience des droits* (Proposals to ensure that rights are more effective), 13 July 2015.

5 *"Service providers have a duty to act, under certain circumstances, with a view to preventing or stopping illegal activities. This Directive should constitute the appropriate basis for the development of rapid and reliable procedures for removing and disabling access to illegal information".*

6 *"This Directive does not affect the possibility for Member States of requiring service providers, who host information provided by recipients of their service, to apply duties of care, which can reasonably be expected from them and which are specified by national law, in order to detect and prevent certain types of illegal activities."*

In other words, it would therefore relate to all types of infringements which could be committed by a service with hosting provider status, and would not specifically deal with content protected by copyright and related rights.

23 Moreover, many of the participants believed that this approach may be counterproductive, as it would not provide a focus on the copyright-specific central issue of the **sharing of value**. They all however considered it essential to ensure that the issue of 'false hosting providers' remains at the heart of the issue of copyright and related rights, therefore justifying a change in legislation and the defining of the regime's boundaries.

24 This approach of enacting a specific standard does not only seem to be the most secure and targeted, but it also appears to offer more potential, in that it grants the option (where required) of adding a number of stipulations, including in cases in which it would be insufficient to simply reject Article14 for the service providers in question, but would also potentially be necessary to better define the applicable legal regime in such a scenario.

25 Of course, everyone is aware that the forthcoming solution is not necessarily a new concept, as the vast majority of contributors considered the case law solutions applying Article 14 to the above-mentioned scenarios to be flawed. Yet putting pen to paper offers certain advantages, such as being able to make a clear and simple statement of the most timely solution, taking into account the balance sought by the joint implementation of Directives 2000/31 and 2001/29, and providing a solution which is acceptable both to the political authorities and to professionals within the sector.

26 It remained only to decide which Directive the above-mentioned solution should be inserted into.

Intervention in copyright legislation

27 Some requests were made to change the actual wording of the **E-Commerce Directive**. Others expressed a preference for the legislative intervention to be restricted to clarifications in a **legislative text covering only to copyright and related rights.**

28 The decision between these two options was quickly made. Although a wider amendment may have seemed more apt to some, it soon became apparent that such an amendment would come up again strong objections and problems, and therefore a **limited intervention** was quickly deemed to be more realistic and apt in this case. Furthermore, the **specific nature of literary and artistic property rights** provides a solid basis for the rejection of Article 14 of the E-Commerce Directive when copyright and related rights are applied by certain service providers.

29 The vast majority of the contributors therefore ideally wanted the forthcoming standard to be inserted into the body of Directive 2001/29, should the latter be reopened. This report favours the latter route. Yet technically, the proposed provisions could be inserted into any copyright legislation.

30 Based on these considerations, what solution ought to be recommended?

II. THE CONTENT OF THE NEW LEGISLATION

31 Once the vehicle has been chosen, the boundaries of the new rule must be set out (A) and then a drafting proposal must be completed (B), followed by an explanation of the proposal (C).

A – The boundaries of the proposal

32 This consists of planning the scope of the legislation and its intended position.

33 Drafters are likely to have the option of several different constructions.

34 One might decide to create a whole **new status** (definition, regime etc.), but this option did not win the support of the contributors for two main reasons:

- Firstly, it has the disadvantage of having to propose new solutions, with complex boundaries to be set out, which themselves could quickly become obsolete due to future technical innovation and economic and social change.

- Secondly, it leads one to believe that new standards need to be drafted, where in fact it is simply a case of rejecting the consequences of case law which has failed to correctly interpret the pre-existing legislation. Such rejection is easily justified by the need for clear content due to the specific nature of copyright.

35 It was therefore deemed that this option ought to be rejected as long as it had not been proven that the renewed enforcement of copyright through the rejection of Article 14 was likely to prompt any problematic imbalances.

36 The route of shifting copyright and related rights back towards ordinary law was therefore deemed

to be more reasonable, especially as the option of (potentially and in certain cases) accompanying it by a simple duty of collaboration between rights holders and service providers seemed likely to maintain a balance between potentially opposing interests.

37 Based on these considerations, where should the new rule be inserted within **Directive 2001/29**?

- Does a new Article need to be created (9a)?
- Does the solution need to be written into an existing provision (Article 9)?
- Do several provisions need to be changed at the same time (Article 9 and Article 3)?

38 **Option one** has been chosen for the sake of simplicity and in order to avoid any adverse effects.

B -Drafting proposals

39 Insertion of a new recital 16a:

1. This Directive and the Directive on electronic commerce have been prepared in such a way as not to contradict one another, particularly insofar as the limitation of liability set out by Article 14 of the second Directive has been devised exclusively for hosting providers offering a mere technical service for storage of information. And yet their respective objectives, namely both the wish to provide a high level of protection for copyright and related rights, and that of ensuring immunity in order to allow hosting providers to develop their businesses, have been shown to be contradictory at the expense of rightholders when the aforementioned limitation of liability has begun to be applied to information society service providers whose intervention, beyond or besides the mere storage of information, consists of giving access to the public to copyright works and/or subject- matter. Such evolution in the application of the Directive on electronic commerce inhibits a high level of protection for copyright and related rights, and prevents rightholders from exercising the rights granted to them by this Directive.

2. It is therefore necessary to stipulate that these information society service providers whose intervention consists of giving access to the public to copyright works and/or subject-matter do not benefit from the limitation of liability set out for a different purpose by Article 14 of the Directive on electronic commerce.

In this respect, it is of no consequence whether the infrastructure or features used by these service providers to give such access to the public to copyright works and/or subject-matter are automated, as this does not provide an exemption from the implementation of the rights protected hereunder.

The provision of an access to the public to copyright works and/or subject-matter, which should not be confused with the mere provision of physical facilities as set out by recital 27 of this Directive, constitutes an act of communication to the public and/or making available to the public as defined by Article 3. This act is performed by the service provider giving such access, under its own liability. If the copyright work or subject-matter is sent to said service provider by a user of its services in order that an access to it is given to the public, the service provider and the aforementioned user together perform the act of communication to the public and/or making available to the public, and therefore hold their joint and several liability.

As they, alone or with the participation of users of their services, are implementing the rights set out by Article 3 and, where relevant, the right set out by Article 2, the information society service providers who give access to the public to copyright works and/or subject-matter must obtain permission from the relevant rightholders.

Such permission covers acts performed by users of their services in order that an access to copyright works and/or subject-matter is given to the public, as long as these users are not acting in a professional capacity.

40 Insertion of a recital 24a:

In accordance with the provisions of Article 11bis of the Berne Convention, these rights must apply whenever the copyright work or subject-matter is subject to an act of communication to the public and/or making available to the public by a third party to the initial act of communication to the public and/or making available to the public, whether this third party uses the same technical method or a different technical method to that used for the initial act.

41 Insertion of a new Article 9a:

Article 9a:

Linking of Directives 2000/31 and 2001/29

Without prejudice to Articles 12 and 13 of the Directive on electronic commerce, information society service providers that give access to the public to copyright works and/or subject-matter, including through the use of automated tools, do not benefit from the limitation of liability set out by Article 14 of said Directive.

These service providers must obtain permission from the relevant rightholders as they, either alone or with the participation of users of their services, are implementing the rights set out by Articles 2 and 3.

Such permission covers acts performed by users of their services when they send the copyright works and/or subject-matter to the aforementioned service providers in order to allow the access set out by sub-paragraph

one, as long as these users are not acting in a professional capacity.

C – Notes

42 **1.** The new provision states clearly that the activities performed by certain service providers do not match the definition provided by Article 14 of the *Directive on electronic commerce* (sub-paragraph one).

43 These service providers, whether alone or with the participation of users of their services, perform acts which apply copyright (sub-paragraph two), which does not necessarily need to be demonstrated within the body of the legislation. A technical and legal analysis of the performed acts provides sufficient proof, for example after a post has been made on a Web 2.0 content contribution site.

44 Such an analysis shows, however, that a single act of making available to the public can be attributed to two people or entities (the uploading web user and the website manager), while sub- paragraph three allows for the legitimacy of the two acts being technically performed simultaneously, as long as the uploading web user is not acting in a professional capacity.

45 **2.** The mission proposes that the service user's action of posting the protected content and that of the technical posting online by the information society service provider should be deemed to be **a single act** in the sense of copyright. Indeed, although a piecemeal understanding is possible in intellectual terms, it would not be logical here as autonomous actions alone have no interest as such[7]. The service provider is dependent on the user who provides it with the content, and the web user must use the service of the provider which, through its intervention, grants access to the work. There is only one final result.

46 This **access criterion** is essential in order to constitute the act of making available. Indeed, the service providers in question enable the public to access protected content. Without their intervention, the public would not have access to this content. Their role is therefore 'indisputable'[8]. The concept of access is moreover central to Article 3.1 of Directive 2001/29: "(...) including the making available to the public of their works in such a way that members of the public may access them from a place and at a time individually chosen by them." Thus, the service provider's intervention to give access to the work prevents Article 14 of Directive 2000/31 from being applied, insofar as the activity in question cannot be summarised as mere storage.

47 Furthermore, the fact that the intervention has been made using automated tools has no bearing on the qualification of the act in question. Indeed, even a technical act does not prevent copyright and related rights from being applied (see for example the transient or incidental copy which required an exception).

48 The wording used in sub-paragraph one may seem broad, but this shouldn't be a cause for concern as it in fact only pertains to those parties that claim to be covered by Article 14, even though they are not simply storing but also giving access to the protected content. ISPs and conduits continue to be covered by the exemptions set out by Articles 12 and 13 (mere conduit and caching), as specified by the chosen provisions.

49 Furthermore and in order to reinforce legal certainty, it is proposed that permission granted by rights holders to service providers will ensure the legitimacy of the act in question as a whole, as long as the service users are not acting in a professional capacity. The latter would therefore no longer be threatened with legal action.

50 **3.** The reference to the Berne Convention made in recital 24a, is crucial at a time at which the Court of Justice of the European Union is interpreting legislation (particularly the right of communication to the public) in a manner which seems to be far removed from a strict legal orthodox approach. This critical change in approach is demonstrated by many of the Court's global case law specialists, particularly but not exclusively in the area of hyperlinks[9]. Two robustly-argued resolutions[10] adopted by the International Literary and Artistic Association (ALAI) are of relevance here. It is of note that this 'learned society' – which was at the source of the Berne Convention – has criticized the Court of Justice for deviating from the meaning that ought to be taken from the international legislation by adding a legal assumption of the requirement of a 'new public' as a basis for the enforceability of copyright.

51 This is an important point, given that some service providers may in the future decide to provide link databases rather than storing files of copyright works.

7 Comp. CJEU, 13 October 2011, *Airfield*, caseC-431/09 and C-432/09.

8 See, on these issues, not. CJEU, 7 December 2006, *SGAE*, case C-306/05, para. 42; CJEU, 13 February 2014, *Svensson*, case C-466/12, para. 18.

9 See CJEU, 13 February 2014, case C - 466/12, Nils Svensson et al. vs. Retriever Sverige AB and C - 348/13, 21 October 2014, BestWater International GmbH vs. Mebes et al.

10 <http://www.alai.org/assets/files/resolutions/2014-avis-public-nouveau.pdf> and <http://www.alai.org/assets/files/resolutions/201503-rapport-et-avis-hyperliens-3.pdf>.

52 **4. So what will be the consequences** of the service providers described by the proposed legislation no longer being covered by Article 14 of Directive 2000/31?

53 The first logical effect, as mentioned earlier, will be **the enforceability of copyright and related rights** on these service providers where they have made copyright works or material accessible. The issue raised is therefore that of a **harmonised and balanced application** of copyright and related rights **in order to enable fair sharing of value without hindering the launch of new services that might be offered to the public.**

54 There are therefore two potential areas for concern.

55 The first is linked to the issue of **the implementation of the new rules over time,** given that, as far as many are concerned, the solutions in question should have been applied as soon as Directive 2000/31 was adopted. It may be wise to consider the circumstances of service providers which have rightly or wrongly relied on the solutions provided by case law, by setting out a time period for application of the procedure, with a view to enabling service providers to adapt to it and find solutions in consultation with rights holders.

56 The second area for concern is linked to the setting up of a system to prevent the '**backlash**' of copyright and related rights which might be considered **overly drastic**, in a scenario in which certain rights holders refuse to grant user licences to service providers. The exercising of the right to prohibit may in this case be accompanied by **a duty of collaboration between the aforementioned rights holders and service providers.**

57 This duty of collaboration is likely to occupy **various forms** in the light of the **state of the art technology** and the **virtuous uses** likely to be developed in this domain. One might decide to set up the negotiation of charters, standard contracts or fingerprint recognition in order to screen works and prevent acts of infringement, potentially from the outset.

58 This approach forms part of a move to **create** an **environment of participation and respect** for the various interests represented. It is not a question of forcing consent, but rather of creating a **virtuous circle and above all a positive spiral enabling the development of new markets to benefit all.**

59 Moreover, this forms part of a greater movement which is already beginning to take hold, and which has led some service providers to enter into discussions with rights holders, as well as being part of the extension of initiatives that some advertising professionals have agreed to in order to clean up the sector and remove the economic dominance of those providers offering works unlawfully by attempting to position themselves out of reach of intellectual property rights.

60 In addition, it is clear that the generalisation of virtuous systems will benefit not only rights holders but also those service providers that are very keen to develop new lawful methods of distributing intellectual works. The latter are however apprehensive of being forced into a situation of 'unfair competition' with those that do not, for the time being, share these concerns, and prefer to maximise their profits by using illegal set-ups rather than considering ways of improving the sharing of value and complying with intellectual property rights.

61 The adoption of a specific legislative text would provide the necessary boost in order to trigger and maintain a virtuous spiral. It would offer rights holders and service providers a 'win-win' situation and provide a basis for the development of new consumer services.

62 In the light of the sheer diversity of circumstances to be taken into account, the detail of each party's duties and rights is of course a sensitive issue, and therefore the reference to the general notions of 'cooperation' and 'collaboration' would appear to be sufficient as a starting point and for cases in which they are necessary.

The Feasibility of Applying EU Data Protection Law to Biological Materials

Challenging 'Data' as Exclusively Informational

by **Worku Gedefa Urgessa***

Abstract: Though controversial the question of applying data protection laws to biological materials has only gotten a little attention in data privacy discourse. This article aims to contribute to this dearth by arguing that despite absence of positive intention from the architects to apply the EU Data privacy law to biological materials, a range of developments in Molecular Biology and nano-technology—usually mediated by advances in ICT—may provide persuasive grounds to do so. In addition, paucity of sufficient explication of key terms like 'data/information' in these legislations may fuel such tendency whereby laws originally intended for the informational world may end up applying to the biological world. The article also analyzes various predicaments that may arise from applying data privacy laws to biological materials. A focus is made on legislative sources at the EU level though national laws are relied on when pertinent.

Keywords: Biological materials; DNA; interpretive framework; data privacy laws; data/information

Recommended citation: Worku Gedefa Urgessa, The Feasibility of Applying EU Data Protection Law to Biological Materials: Challenging 'Data' as Exclusively Informational, 7 (2016) JIPITEC 96 para 1.

A. Introduction

1 There were numerous reasons for enacting the first data protection laws in the 1970s. Among the most important factors was a public fear and disempowerment engendered by greater dissemination, use, and re-use of personal data across organizational boundaries facilitated by new technology in the form of electronic data processing. The latter has also created a sense of loss of control over technology and automation of societal processes.[1] In addition to rapidly increasing capacity to store data, computers permitted information to be searched and organized by multiple attributes, rather than through a single index (for example, first and last name only). This capacity changed the way information could be linked to an individual[2] which led to data protection laws focused on protecting "personal data" in the EU and "Personally Identifiable Information (PII)" in the United States of America.[3] The definitions of these key concepts delimit the scope of application of data protection laws. Since those early days one of the major changes in the EU has been the recognition of data protection as a fundamental right in itself, independent from the right to respect for private life.[4]

2 Today, more than 40 years since the early data

1 Lee Bygrave (2014), *Data Privacy Law, an International Perspective*, Oxford University Press, Oxford p. 8-15; See also, Article 29 Working Party, "Opinion 4/2007 on the Concept of Personal Data," Adopted on 20th June, 2007, 01248/07/ENWP 136, p.5. Recital 4 in the preamble to the DPD makes a similar assertion.

2 Paul Schwartz & Solove Daniel, "The PII Problem: Privacy and a new Concept of Personally Identifiable Information," *New York University Law Review*, Vol. 86, (2011), p. 1820.

3 The U.S., however, lacks a comprehensive set of data protection rules as is available in Europe and relies instead on sector specific rules. (See, Bygrave (2014), p. 110-12).

4 See Article 16 of the Treaty of the Functioning of the European Union and Article 8 of the Charter of Fundamental Rights of the European Union.

protection laws[5] and two decades after the EU Data Protection Directive was adopted, the technological landscape has dramatically changed. Computer power continues to grow[6] with additional capacities and data processing capabilities. The growth in computer power has aided a significant transformation in many fields of study including molecular biology and nano-technology. Consequently, there is strong criticism on sustainability of the definition of personal data[7] maintained in the EU data protection laws. According to this definition personal data is, in essence, *information* which is capable of *identifying* living human data subjects. Other elements[8] of the definition have gotten a fairly detailed analysis except the phrase '*any information*.' I seek to analyze and challenge the conceptual predispositions behind this criterion: the notion that data protection laws apply to '*data* and/or *information*.'

3 The propriety of data as exclusively 'informational' is being put to test as advancements in bio-technology and ICT continue to blur the distinction between the human biological materials[9] on the one side and information derived from them on the other.[10] The fear is that such distinction may arbitrarily undermine the protections offered under the right to privacy in general.

B. Data/Information Defined

4 The terms 'data' and 'information,' though key legal terms, are often taken for granted and insufficiently, if at all, defined in data protection discourse.[11] Data is habitually used as synonymous with information. Scholars attribute this dearth in clarity, specifically in laws directly dealing with information concepts, to various factors and contestable assumptions ranging from a simple oversight, to an assumption of obviousness, and to pessimism that the terms are incapable of definition, at least a legally workable one.[12]

5 While it might have worked reasonably well in the past, the paucity in clearly defining[13] the two terms appears to have reached an unsustainable stage. The most germane reason for the purpose of this study is the challenge scientific and technological developments[14] introduce to the boundary between information and biological materials — and, in effect, traditional distinction between the message and the medium — which can also trigger application of laws that employ information concepts to biological material.[15]

6 Outside of the legal world, the day-to-day usages of the two terms seem to draw no clear line of distinction; neither is there a need to make a major differentiation between the two. In their normal parlance, Oxford English Dictionary defines 'data' as 'facts and statistics collected together for reference or analyses'[16] and 'information' as 'facts provided or learned about something or someone.'[17] Even though a first glimpse at these definitions tells us that information is a result of analysis carried out on data, one can also see the usage of the word 'facts' in both definitions which suggests that no serious distinction is aimed to be made. Besides, the thesaurus[18] section

5 The first national Data Protection law was enacted by Sweden in 1973 (Sweden's Data Act); repealed and replaced by Personal Data Act of 1998; the first data protection law ever enacted was the *Data Protection Act* passed by the German *Land* Hassen in 1970. (See, Bygrave, 2002, p. 179, 187).

6 According to the notorious *'Moore's Law'* (an observation named after Gordon E. Moore of Intel) computer power (i.e. transistor count on an integrated circuit) continues to double every two years at least for another decade.

7 Article 2(a) of the DPD reads: 'personal data' shall mean any information relating to an identified or identifiable natural person ('data subject'); an identifiable person is one who can be identified , directly or indirectly, in particular by reference to an identification number or to one or more factors specific to his physical, physiological, mental, economic, cultural or social identity. The definition remains essentially the same in the new General Data Protection Regulation.

8 For an overview of the elements of the definition, see, Article 29 Working Party's Opinion no. 4/2007 on 'the concept of personal data.'

9 In this article the phrase 'biological materials' is used to describe a natural substance taken from a living human — such as blood or tissue — where information contained in the material can be traced back to the individual.

10 See Lee Bygrave, "Information Concepts in Law: Generic Dreams and Definitional Daylight," *Oxford Journal of Legal Studies*, Vol. 35, No.1 (2015): 1-30; Lee Bygrave, "The Body as Data? Bio-bank Regulation via the "Back Door" of Data Protection Law," *Law, Innovation and Technology* Vol. 2, issue1 (2010): 1-25; Irma van der Ploeg, "Genetics, biometrics and the informatization of the body," *Ann 1st Super Sanità*, Vol. 43, No. 1, (2007): 44-50, and Mark Taylor, *Genetic Data and the Law: A Critical Perspective on Privacy Protection*, Cambridge University Press, 2012, Chapter 7.

11 The A29WP as well, in its opinion 4/2007 where it defined the concept of 'personal data', took the term 'data' for grated and had never even asked the question.

12 Bygrave (2015), p. 107-111.

13 By clear definition it is not meant here to necessarily create a distinction between the two terms; clarifying them to be synonyms works well.

14 As will be discussed further below, these technological developments include: the advancement in ICT and Biotechnology which enabled an ever greater generation of information from biological materials, and making them core constitutive elements of information systems. (Bygrave, 2015, p. 93) In addition, developments in nano-technology and neurology are also blurring the boundaries between technology and human body.

15 Bygrave (2015), p. 94.

16 Available at: <http://www.oxforddictionaries.com/definition/english/data>, last accessed 23 May 2016.

17 Available at: <http://www.oxforddictionaries.com/definition/english/information>, last accessed 23 May 2016.

18 The thesaurus also lists other related words like facts, figures, input, documentation and file as synonyms to data/

of the same dictionary puts 'information' and 'data' as synonyms.[19]

7 In the fields of Informatics and Computer Science, however, a more systematic distinction is drawn between data and information. In these fields, the notion of 'data' usually denotes signs, patterns, characters or symbols which potentially represent something (a process or object) from the 'real world' and, through this representation, may communicate 'information' about that thing.[20]

8 Expectedly, compared to the nebulous day-to-day and, even, legal usage the distinction made in Informatics appears to be more logical and coherent. The question, however, is would these conceptual walls built in the fields of Informatics and Computer Science be sustainable in the face of the current development in ITC and bio-technology? And, even if they continue to work, should the same distinction be made in legislating new or interpreting the existing laws dealing with information concepts? By focusing on data protection law among the latter types of laws, the following sections will strive to address these questions.

C. Are Biological Materials Personal Data in the EU Data Protection Regime? (lex lata)

I. The Existing Legal Regime

1. The Data Protection Directive

9 A brief glimpse at the EU Data Protection Directive (DPD) not only fails to answer whether biological materials are considered to be personal data but makes the answer even fuzzier by its interchanging usage of the words 'data' and 'information.'[21] However, a closer look at the provisions of the DPD indicates absence of intention by its architects to consider biological materials to be personal data. Though absence of intention to cover biological materials appears clear, for reasons discussed below, one cannot, at the same time, plausibly argue that that was an intentional exclusion either.

10 First, nonexistence of a clear intention to consider biological materials as personal data is rooted on how the law and policy in this area generally operates. Professor Bygrave observes:

> *"[T]he law and policy on data protection have generally tended to operate on the assumption that a distinction exists between data/information on the one hand, and, on the other, the person(s) to which the data/information can be linked."*[22]

11 We see this in the definitions of 'personal data' and/or 'personal information' given in data protection laws.[23] Therefore, paucity of a good indication to treat biological materials as personal data begins from the very definition under Article 2(a) of the DPD. The definition portrays '*humans*' as data subjects to which *information* relates; not humans, or a sample taken from them, as information by themselves. It is worth noting, though, that when it tries to further define 'an identifiable person' the directive employs terminologies that relate to the human body. It provides that, in addition to information like a person's identity number, a person can be identified by his physical, physiological or mental identity. Yet, a reference to, say, physical identity of a person to identify him, quickly winds up being an information about his physique, like his appearance, and not the physical self as such. The same is implied by the preparatory materials towards adoption to the directive.[24] The then EC Commission's commentary[25] to this part of Article 2(a) of the directive, after indicating the typical numerical information[26] as identifying factors, mentions that the definition would also cover data such as appearance, voice, fingerprints and genetic characteristics.[27]

12 Secondly, other key provisions of the DPD are also indicative of the absence of a positive intention[28] by the legislature to treat biological materials as

information.

19 Available at: <http://www.oxforddictionaries.com/definition/english-thesaurus/information>, last accessed 23 May 2016.

20 Paolo Atzeni et al, D*atabase Systems: Concepts, Languages and Architectures* (McGraw-Hill, 1999) p. 2; Chrisanthi Avgerou and Tony Cornford, *Developing Information Systems: Concepts, Issues and Practice* (Macmillan, 2nd eds 1998) p. 115.

21 For instance, recital 26 in the preamble to the DPD uses both 'information' and 'data' in the same context when it tries to delimit the application of data protection principles. This is problematic because, even when human biological materials may be considered as 'data', along the lines of the conceptual distinction between information on one side and data on the other, the directive does not make sense of such distinction.

22 Bygrave (2010), p. 13.

23 Ibid.

24 Commentary of the Commission, October 1992: COM (92) 422 final—SYN 287, p. 9.

25 Ibid.

26 A person can be identified....indirectly by a telephone number, a car registration number, a social security number, a passport number or by a combination of significant criteria.

27 Commentary of the Commission, October 1992: COM (92) 422 final—SYN 287, p. 9.

28 By positive intention I mean a deliberate and calculated move from the architects to consider biological materials as personal data.

personal data. Some vital words and phrases used throughout the directive cannot semantically accommodate human biological materials. Words like 'recording' and 'alteration' as set of operations to be performed on personal data under Article 2(b) of the DPD epitomize such inhospitable accommodation. Other instances are under Article 6 whereby personal data is required to be 'accurate' and 'up to date' which presupposes that data could be 'inaccurate' and/or 'out of date', which a biological material cannot be. Similarly, the right to 'rectify' under Articles 10 and 11 presuppose some form of error in recording.

13 Thirdly, and perhaps more importantly, the crafting of the scope of application of rules of the DPD, under Article 3, cannot comfortably accommodate application of rules of the directive to human biological materials. The directive applies to the processing of personal data in two scenarios: wholly or partly by automatic means, and to manual processing of personal data which form/intended to form part of a filing system. At least partly-automatic processing of data, which the directive requires under the first scenario, has in mind the use of a device, computers mostly, to process information electronically, i.e. when data is computerized. This is exactly what is referred to by the Commission's commentary on this provision.[29] As far as biological materials are concerned, one may not, right away, use computers to process blood samples or a swab of specimen of a person. An exposure to a different interpretative framework may be required. The same holds true for the second scenario, i.e., filing system: a file literally presupposes recorded information.

2. The General Data Protection Regulation

14 Having been invited by the European Council to evaluate the functioning of EU instruments on data protection, as part of the Council's Stockholm Program Notices[30], the EU Commission came up with a proposal for the GDPR in December, 2012.[31] On 12 March, 2014, European Parliament made its formal First Reading vote confirming the text of the draft Regulation.[32] EU Justice and Home Affairs ministers reached a general approach on the Regulation at their Council meeting on 15 June, 2015.[33] After months of "trilogue" negotiations, the EU Commission, Parliament and Council of Ministers reached agreement on the GDPR on 15th December, 2015.[34] Following political agreement reached in the "trilogue" the official texts of the Regulation was published in the EU Official Journal on 4 May, 2016. While the regulation will enter into force on 24 May, 2016, it shall be applicable from 25 May, 2018 onward.[35]

15 To examine the position taken by the GDPR on the issue of human biological material, I will analyze, mainly, the official text (of 4 May, 2016). However, in order to trace the developments on this issue, I will also make references to the Commission Proposal (of January 2012), Parliament's first reading (of March, 2014), the Council's general approach (of June, 2015) and the compromise text that resulted from the final trilogue.

16 The Commission's proposal explicitly mentions the term 'biological samples'[36] in recital 26 of the preamble to the proposed regulation. The mention is made as part of enumerating the constituents of personal data relating to health. It reads:

> *"Personal data relating to health should include... information derived from the testing or examination of a body part or bodily substance, including biological samples..."*[37]

17 Whilst a bold step in separately and explicitly bringing up 'biological samples' which creates a tempting syntax to consider 'biological samples' as personal data relating to health, a closer examination of the recital as a whole shows that it is dealing with information derived from testing or examination of biological samples, not biological samples in and of themselves. In other words, the recital conveys the following meaning: personal data relating to health should not be limited to the information derived from testing/ examination of body part or bodily substance (which require the physical presence of the examinee) but should also include the result of examination of samples when it is taken from examinees, the presence of whom is no longer required for examination.

18 While the same ambiguous syntax is employed in other language versions such as Danish, Swedish and French, Professor Bygrave observes that the German

29 Commentary of the Commission, October 1992: COM (92) 422 final—SYN 287, p. 12.

30 The Stockholm Programme — An open and secure Europe serving and protecting citizens, OJ C 115, 4.5.2010, p.1.

31 COM(2012) 11 final.

32 Bird & Bird, EU Framework Revision: Overview, at: <http://www.twobirds.com/en/practice-areas/privacy-and-data-protection/eu-framework-revision>, last accessed 23 May 2016.

33 Ibid.

34 Ibid.

35 European Commission, Personal Data Protection, available at: <http://ec.europa.eu/justice/data-protection/>, last accessed 23 May 2016.

36 The word is mentioned for the first time in EU instruments on data protection.

37 Recital 26 of the preamble to the Proposed General Data Protection Regulation.

version rules out such ambiguity.[38] In that case, it comes down to a question of interpretation: which language version takes precedence? Recourse to the jurisprudence of the Court of Justice of the EU tells us that the different language versions are all equally authentic and that the interpretation of a provision of Community law involves a comparison of the different language versions.[39] The court further notes that every provision of Community law must be placed in its context and interpreted in the light of the provisions of Community law as a whole, regard being had to the objectives thereof and to its state of evolution at the date on which the provision in question is to be applied.[40] Therefore, the task of ascertaining the true meaning of deferring language versions is not simply mechanical, i.e. it does not depend on the comparison of the number of versions that avoid the problematic syntax against those which contain such syntax. It should be rooted in the context in which the words are placed, its evolution and the objective of the law as a whole. Seen from this angle, it is difficult to claim that the proposed Regulation, indeed, considers biological materials as personal data related to health.

19 The European Parliament's first reading did not introduce changes to the Commission's proposal in this regard. A small alteration with additional mentions[41] of 'biological samples' came with, first, consolidated text of the Council and the Commission and, latter, with the compromise text. In these versions, recital 26 to the preamble of the regulation reads:

> *"Personal data concerning health should include... information derived from the testing or examination of a body part or bodily substance, including genetic data and biological samples...."*[42]

20 As can be discerned, in this version of the regulation the phrase 'genetic data' is added to the original script. The overall reading of this part of recital 26 would not offer the exact same meaning that the corresponding sentence in the Commission's version did. In that version, the phrase 'biological samples' can be meaningfully read back to the phrase: 'Information derived from testing or examination of...' That makes sense because like body parts or body substances, biological samples can also be subjects of said testing/examination, and, thus, be carriers of personal information to be derived from them. In addition, referring the phrase 'biological samples' to the 'information derived from testing or examination of...' would be repeating oneself as 'examination of a body part or bodily substance' is already mentioned and biological samples can be considered to be body parts/ bodily substance.

21 However, the same interpretation wouldn't be logical with the addition of 'genetic data' in the later versions of the regulation. That is mainly because genetic data is already a result of analysis of biological materials.[43] Genetic data is generally understood to be information by itself, and while possible, it is usually not a subject of testing or examination to derive information, as we frequently do from body parts/ bodily substances. Therefore, it creates a temptation to read 'genetic data' and 'biological samples' back to the phrase with which the recital begins: 'personal data concerning health should include...' Otherwise, referring it back to the inner phrase which reads: 'Information derived from testing/examination of...' would end up being, 'information derived from testing/examination of information about heritable characteristics of individuals. That, in turn, ends up being 'Information derived from testing/examination of information.'

22 While not particularly strong, this can be taken as a reasonable interpretation of the wordings of the compromise text. But, it still remains ambiguous at this point. This interpretation also advances the attainment of the general objectives[44] of the regulation set out by the Commission, particularly the first objective: helping citizens to be in control of their data.[45] After all, the very conception of privacy is ingrained in the protection of personal integrity, which, at some level, requires extending protection to our biological materials.

23 However, towards the end of writing this study, the official text of the Regulation was published in the EU Official Journal on 4 May, 2016.[46] Recital 35 in

38 "Informationen, die von der Prüfung oder Untersuchung eines Körperteils oder einer körpereigenen Substanz, darunter biologischer Proben, abgeleitet wurden" (See, Bygrave (2015), p. 6).

39 Case-283/81, *CILFIT v Ministry of Health* [1982], Para. 18.

40 Ibid, Para. 20.

41 The compromise text mentions the phrase 'biological samples' at three different instances in the regulation. The first being in recital 26, the other two are made in relation to elaborating and defining 'genetic data' under recital 25(a) and Article 4(10) respectively.

42 See recital 26 in the preamble to the GDPR (the compromise text).

43 Recital 25(a) and Article 4(10) of the compromise text of the regulation clearly testify to the fact that genetic data results from the analysis of biological samples.

44 The Commission sets out three general objectives for the regulation, See The Proposal for GDPR, P. 102.

45 Some commentators, though, have argued these objectives are based on fallacious assumptions, thus, unattainable. See, Koops, B.J. (2014) "the trouble with European data protection law," *International Data Privacy Law*, Vol. 4, No. 4.

46 REGULATION (EU) 2016/679 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation).

the preamble to the official text of the regulation clarifies some of the issues raised with in recital 26 of the previous versions. The relevant part of the recital reads:

> *"Personal data concerning health should include ... information derived from the testing or examination of a body part or bodily substance, including from genetic data and biological samples..."*[47] *(emphasis added)*

24 The addition of the preposition '*from*' now makes it difficult to read 'biological samples' back to the beginning of the recital. It should be read with the phrase 'information derived from testing or examination of...' This implies the absence of positive intention by the architects of the regulation to consider biological samples to be personal data. The previous version can, therefore, be considered a result of poor draftsman-ship.

25 Having said this much about the DPD and the GDPR, I will now briefly turn to the status of biological materials under European case laws, and national legislations. The focus of the study being on the legal regime at the European level, the coverage of national legislation will be brief. As far as national laws are concerned, they appear to be divided along geographic lines. Many western European countries tend to adopt the view that biological materials are not personal data while some eastern European countries have taken the opposite stance. Bulgaria, Estonia, Latvia and Romania are among eastern European countries that recognize body samples as data in contrast with other western European countries like Spain, Portugal and Germany.[48] Outside Europe, the Australian state of South New Wales's privacy and information legislations clearly include bodily samples in their definition of personal information.[49]

26 As was the case for data protection in general, case law on the issue of 'biological materials as data' has not been abundant. While there is a considerable number of case law relating to data protection today, many of them have hardly shed any light on the issue of bio-materials as data. That could be attributed, at least in part, to the level of awareness of the European population regarding the systemic accumulation and use of biological materials in general. For instance, it is not only unclear what bio-banks are used for or how their use may affect the status of fundamental rights but it also is not widely-known that they even exist. One study of the European Commission found that more than two-thirds (67%) of Europeans have never even heard of the term itself.[50] Only 2% of the population has actively inquired into and searched for bio-banks.[51] As awareness rises on what bio-banks are, how they are used, and their adverse effects on privacy, it can be expected to lead to privacy litigations which would involve biological materials.

27 Among the few instances in which courts dealt with this issue are the cases of *S and Marper v United Kingdom*[52] handed down by the European Court of Human Rights and the decision of Norwegian Data Inspectorate.

28 In *Marper* the European Court of Human Rights essentially ruled that the retention of fingerprints, cellular samples and DNA profiles of individuals arrested but who are later acquitted or have charges against them dropped is a disproportionate interference to their right to privacy under Article 8 of the European Convention on Human Rights. That being the chief finding of the court in this judgment, the court has also directly, though scarcely, addressed the issue of human tissue samples. It found that cellular samples constitute personal data within the meaning of Data Protection Convention:

> *"The Court notes at the outset that all three categories of the personal information retained by the authorities in the present cases, namely fingerprints, DNA profiles and cellular samples, constitute personal data within the meaning of the Data Protection Convention as they relate to identified or identifiable individuals. The government [UK] accepted that all three categories are "personal data" within the meaning of the Data Protection Act 1998 in the hands of those who are able to identify the individual."*[53]

29 While a remarkable judicial activism, the effect of this view in the judgment is limited in a number of ways. It figured only marginally in the judgment because the court did not need to delve in to the issue of biological material as application of Article 8 ECHR, on which the judgment is based, does not turn upon whether 'data' or 'information' are/is processed but on whether or not there is interference with the right privacy. Also, the court does not have a legal mandate of interpreting the Data Protection Convention.[54]

30 It is also worth mentioning here that in prior litigation of the case in the UK by the House of Lords,

47 Recital 35 in the preamble to the GDPR (EU Council's Position with the view of adoption, 6 April, 2016).

48 See, Bygrave (2010), p. 16-17 for references.

49 Section 4(2) of Privacy and Personal Information Protection Act 1998, section 5(2) of the Health Records and Information Privacy Act 2002 and the Government Information (Open Access) Act 2009, Schedule 4, clause 4(2).

50 EU Commission(2012), *Bio-banks for Europe: A challenge for governance*, P. 24.

51 Ibid, p. 25.

52 *S and Marper v United Kingdom*, European Court of Human Rights, (App no 30562/04 and 30566/04), 4 December 2008.

53 *S and Marper Vs UK*, para. 68.

54 For detailed analysis of this decision, see Bygrave (2010) p. 7-13.

the issue of bio-samples as data is directly touched upon by Baroness Hale. She argued that the same privacy principles should apply to all the three (fingerprints, DNA profiles and cellular samples), essentially, because they are all kept for and as 'information.' Those are her words:

> *"But the only reason that they [samples] are taken or kept is for the information which they contain. They are not kept for their intrinsic value as mouth swabs, hairs or whatever. They are kept because they contain the individual's unique genetic code within them. They are kept as information about that person and nothing else. Fingerprints and profiles are undoubtedly information. The same privacy principles should apply to all three."* [55]

31 As will be discussed in the next section, Hale's point forms one of the basic arguments put forth in favor of considering bio-samples to be data/information.

D. Should Biological Materials be treated as Personal Data (lexferenda)?

32 There is no consensus on the issue of whether human biological materials should be treated as personal data. Some scholars, commentators and agencies enforcing data protection laws have taken the view that personal data should not be seen to include biological materials for the purposes of data protection laws. The Article 29 Working Party[56] and the UK's Information Commissioner's Office (ICO)[57] are cases in point. In its opinion where it clarifies the concept of personal data under the DPD, the Working party makes a clear distinction between biometric data — which it rightly considers as personal data — and human tissue samples from which biometric data is extracted, which it is opined not to constitute personal data. In the Working Party's words:

> *"Human tissue samples (like a blood sample) are themselves sources out of which biometric data are extracted, but they are not biometric data themselves (as for instance a pattern for fingerprints is biometric data, but the finger itself is not). Therefore the extraction of information from the samples is collection of personal data, to which the rules of the Directive apply."*[58]

33 In a similar way, the official view from the UK's Information Commissioner is reported to be analogous: a sample is not treated as personal data, 'because it is physical material'.[59]

34 On the other hand, even though much of the data protection law and policy have been operating on such distinction, scholars[60] have questioned the logic underlying the distinction between human biological materials on the one hand and personal data on the other. Those pushing the view that biological material may be personal data or information tend to pay more regard to pragmatic considerations, such as the need to fill lacunae in bio-bank regulation, the growing ease with which persons can be identified from biological material, and the fact that such material is often only stored for generating information.[61] Others who take the view that biological material does not constitute personal data depend on conceptual logic claiming that "data is a formalized representation of objects or processes, while information comprises a cognitive element involving comprehension of the representation."[62] In the following sections I will analyze whether such conceptual distinction still makes sense, at least as far as (human) biological materials are concerned, in relation to recent developments in the field of bio-technology.

I. The Conceptual Framework: Does it still make Sense?

1. DNA: the Game Changer

35 The discovery of the structure and basic nature of DNA (deoxyribonucleic acid) as carrier of human genetic information around mid-20th century[63] brought about significant development of how we understand the code of life. It has been argued that the discovery of DNA, as well as our understanding of its structure and functioning, may well be the most important discovery of the last century.[64] The effect

55 *S, Regina (on application of) v South Yorkshire Police*, [2004], Para.70.

56 The Article 29 Working party (A29WP) is an independent advisory body established by the Article 29 of the EU Data Protection Directive.

57 The ICO is the UK's independent body set up to uphold information rights in general, including those under the UK Data Protection Act.

58 A29WP, Opinion 4/2007, p. 9.

59 Beyleveld, Deryck et al., "The UK's Implementation of Directive 95/46/EC" in, Deryck Beyleveld *et al (eds.) Implementation of the Data Protection Directive in Relation to Medical Research in Europe*. Ashgate, 2004 P. 428.

60 Bygrave (2010); Bygrave (2014); Taylor (2012), Chapter 7; Ploeg (2007).

61 Bygrave (2015) p. 7, Bygrave (2010) p. 8-9.

62 Ibid, p. 6-7.

63 The chemical DNA was first discovered in 1869, but its role in genetic inheritance was not demonstrated until 1943. In 1953 James Watson and Francis Crick determined that the structure of DNA is a double-helix polymer, a spiral consisting of two DNA strands wound around each other. (*Encyclopedia Britannica: Science and Technology*).

64 Murnaghan (2016), available at Explore DNA, <http://www.

of the discovery of DNA on scientific and medical progress has been enormous, whether it involves the identification of our genes that trigger major diseases or the creation and manufacturing of drugs to treat these devastating diseases.[65]

36 Among the noteworthy effects of this discovery (reinforced later by the genome project[66]) is the characterization of DNA as a recipe of life; a carrier of information based on which our cells make the necessary protein. That means the very essence of all living cells which make up a human person are the products of those information. But before that analysis, it will be important to say few words on the nature and meaning of DNA to put the discussion in context.

37 Our bodies are made from billions of individual cells, and DNA is the control center of each and every cell.[67] DNA is the hereditary material in humans and almost all other organisms. Nearly every cell in a person's body has the same DNA.[68] Therefore, almost every cell in our body houses a complete set of our hereditary materials, i.e., the genome.

38 On a deeper level, DNA consists of a strand of four nucleotides called adenine, guanine, cytosine, and thymine, commonly abbreviated to A, G, C, and T, respectively.[69] A particular arrangement of these nucleotides forms a gene. Genes specify the kinds of proteins that are made by cells.[70] That means, the sequence of the nucleotides are read to make a particular type of protein that our body needs. It is from that information that proteins are made.

39 Almost everything in the body, from hair to hormones, is either made of proteins or made by them.[71] Therefore, as a protein forms the building blocks of our body, it literally means that we are made up of information read from our DNA, the arrangement of nucleotides. That is why Matt Ridley wrote "the idea of the genome as a book is not, strictly speaking, even a metaphor. It is literally true."[72]

40 This striking scientific discovery about our body is at odds with the traditional conception of distinguishing data as (medium representing reality) opposed to information (comprehension of the representation), at least as far as the body is concerned. The human body itself is a construct of information; information which instructed the formation of proteins, which, in turn, make up our body.[73] The conceptual rigor, thus, begins to crumble when we closely scrutinize the human DNA.

41 In addition to being a source of our genetic code, it is now understood that DNA also possesses a capacity to carry external information; a scientific breakthrough has discovered that it can carry external large size information for a long time.[74] But, that development still remains nascent.

2. Other Developments in Biotechnology and Beyond

42 In addition to the scientific facts revealed about our DNA, the conceptual distinction between data and information is also challenged by multiple other developments that blur the clear boundary between biology and technology.

43 First, after the Human Genome Project, another initiative labelled 'America's next big thing'[75] in neuroscience research, called the 'BRAIN' (Brain Research through Advancing Innovative Neuroethologies) was announced by President Obama in his State of the Union address of January 2013.[76] The BRAIN initiative aims to decode the tens of thousands of connections between each of the ~86 billion neurons[77] that form the basis of human

exploredna.co.uk/the-importance-dna.html>, last accessed 23 May 2016.

65 Ibid.

66 The Human Genome Project (HGP), undertaken from 1990 - 2003 with billions of dollars involving multiple continents, was an international scientific research project with the goal of determining the sequence of chemical base pairs which make up human DNA, and of identifying and mapping all of the genes of the human genome from both a physical and functional standpoint.

67 Calladine, Chris, Horace Drew, Ben Luisi and Andrew Travers, *Understanding DNA: The Molecule and how it Works*. London: Elsevier Academic Press, 2004, p.3.

68 Some cells, like the red blood cell, do not have nucleus, thus, a DNA (Ridley, Matt. *Genome: The Autobiography of a Species in 23 Chapters*, New York: HarperCollins Publishers, 1999, P.6).

69 Amos (2005), p. 6.

70 Jeremy Berg, John Tymoczko and Lubert Stryer, *Biochemistry*. New York: W H Freeman, 2002, Chapter 5.

71 Ridley, (1999) P.7.

72 Ridley (1999) p.6.

73 It may be important to note here that my argument is only limited to biological materials. The conceptual distinction, otherwise, still makes full sense elsewhere.

74 See, Independent, *Single DNA molecule could store information for a million years following scientific breakthrough*, 17th August, 2015.

75 Such project, though, is not of interest only in the United States of America; the European Commission has almost simultaneously announced the Human Brain Project with an award of 1.19 billion Euros. (See, Kaku (2014), p. 250).

76 See, Isabelle Abbey, News and Views: The Brain Activity Mapping Project – What's the plan? April 24, 2013. Available at: <http://thebrainbank.scienceblog.com/2013/04/24/news-and-views-the-brain-activity-mapping-project-whats-the-plan/>, last accessed 23 May 2016.

77 Neurons are nerve cells that carry information between the brain and other parts of the body (Cambridge Dictionaries

brain.[78] That means, as the Human Genome Project sequenced all our genes, the BRAIN initiative will map all of our neurons. That can be said to be the general goal of the initiative.

44 The unstated goal of this initiative, the part directly germane to this study, is eloquently described by Dr. Michio Kaku, Professor of Theoretical Physics at City University of New York in his 2014 book titled 'The Future of the Mind.'[79] The ambitiously expected main output of this project is what scientists call a *connectome*: a comprehensive map of neural connections in the brain which encodes all our memories, dreams, hopes and desires, perhaps, on a CD. This raises very important questions: by putting together a CD of a person's connectome with their genome, are scientists creating, in some sense, immortality?[80] Because even after people are dead, their body could be revived from their genome and their consciousness can be restored from their connectome. That means that we can continue to live, even after we are dead, as information. That possibility that we can still continue to live as information tempts us to conclude that we are nothing but information.

45 Secondly, the undergoing various forms of 'human enhancement projects[81]' are clouding the boundary between human body and technology. Our body may no longer be limited to what it is today;[82] its shape, composition and, as a result, its capabilities are radically changing. It is now clear that "human enhancement" is a reality and not just a product of science fiction.[83] Even more so as technological advances will imminently provide various devices that will interface with the human body in various ways.[84]

46 Thirdly, the steadily growing accumulation of human biological samples in bio-banks[85], and the increased deployment of biometric technologies in every sector are also 'informationalizing' the human body by converting features of it in to processable digital data.The upsurge in the proliferation, coverage, sophistication and uses of bio-banks is spurred in large part due to the advances in genetic science.[86] The need for identification/verification of persons in both public (like in forensic investigations) and private (such as private security) is largely the reason for the expansion in deployment of biometric technologies. Regardless of the reasons for their upsurge they have a clear common effect: conversion of particular aspects of physical existence into electronic data and digitally processable information.

47 All of these developments — from the sequencing of our genome, to the future mapping of our neurons, to the various human enhancement initiatives, and to our continued existence in the form of biometric information—undoubtedly challenge the conceptual separation between the human body, on the one hand, and information about it, on the other.

48 Dr. Irma Ploeg convincingly suggests that this should be seen as something more profound than constituting yet one more instance of the collection of "personal information", as is more commonly done. Rather, the human body is implicated in a process of co-evolution with technology, information technologies in particular.[87] A new conceptualization of bodily existence; an emergence of new body ontology: body as information.[88]

II. Pragmatic and Other Considerations

49 In the previous section it is argued that the conceptual distinction between biological material and information can no longer be logically defended for all the reasons discussed therein. In this section, I will turn to the more pragmatic, and more importantly persuasive, reasons for extending the definition of 'personal data' to have a room for biological materials.

Online).

78 Ibid.

79 Kaku, Michio. *The Future of the Mind: The Scientific Quest to Understand, Enhance and Empower the Mind*. New York: Doubleday, 2014, p. 252.

80 Ibid.

81 In the context of engineering, human enhancement can be defined as the application of technology to overcome physical or mental limitations of the body, resulting in the temporary or permanent augmentation of a person's abilities and features (See, *Human Enhancement*, Dartmouth Journal of Undergraduate Science, In Fall 2013).

82 As a naturally (biologically) constituted being with natural organs, muscles, bones and bodily fluids.

83 The Guardian: *Yes, nano science can enhance humans – but ethical guidelines must be agreed*, Monday 3 June 2013.

84 Ibid; an article in Science magazine exemplified how machines can interact with living brains to allow wireless changes in behavior by the implantation of devices directly into the brains of mice. These devices could then be remotely controlled to activate different parts of the brain using light. (*Science Magazine, Injectable, Cellular-Scale Optoelectronics with Applications for Wireless Optogenetics*, 12 Apr 2013 (www.science.sciencemag.org)).

85 Bio banks may exist in any forms; be it, tissue, blood, cell material, skin, gamete, or embryo banks.

86 Bygrave (2010), p.3.

87 Irma (2007), p. 47.

88 Over the past century developments in the medical Sciences have resulted in various body ontologies like 'the endocrinological body' (in the early twentieth century) whereby the body is viewed as just biochemical entity. (Irma 2002).

1. Indistinguishable Interpretive Potential[89]

50 If one is concerned about practically preventing adverse effects on the right to privacy, what matters most is the interpretive potential of data/source i.e. the ability to generate information that can be linked, not just the assumed availability of identifiable information. If any concerning, from a privacy-related viewpoint, identifiable information can easily and readily be generated from a given source — which more often is the case for biological samples — then that raises as much privacy concerns as the information derived from them would. We can consider two important, but related, reasons to substantiate this sameness in interpretation potential between the two.

51 First, if interpretation[90] is the reason for the distinction, even recorded information will undergo an interpretation before it informs. Taylor observes that: it remains the case that data (as recorded information) must always be interpreted before its meaning can be understood: records must be read. If the privacy protection established by the Directive extends to include the physical record of information, then the viability of any division between (biological) sample and information built upon the former's need for subsequent interpretation crumbles.[91]

52 Secondly, even if recorded information might be said to have an imminent and easy potential to inform than a biological material before it is interpreted, this would not lead to the conclusion that the relative ease in accessibility of recorded information puts right to privacy any more vulnerable than biological materials. It all depends on the availability of the necessary interpretive framework to derive readily accessible information from the samples. A western person, born and raised in the west, may not be able to be informed by having access to 'information' written in an eastern script — say Mandarin. But that does not, in any way, mean that the 'Mandarin text' is not recorded information. It just means that, for that text to inform, the necessary framework should be in place: the skills to read and understand Mandarin.

53 Thus, recorded information and biological samples have an indistinguishable potential of putting right to privacy in jeopardy. In some situations, however, a concern from biological samples could be much worse. Interpreted information may be manipulated, if necessary, to meet certain privacy standards while biological materials will always be available to give away any information in the open. While the manipulation of data may seek to make certain information more accessible, it might also seek to obscure it (e.g. through coding), and the source data may remain interpretable in any event.[92] In this regard, Taylor argues that even information, not just samples, can be subjected to new interpretation, thus, sharp distinction should not be drawn between recorded information and bio-samples.[93]

54 While Taylor's argument is valid, it should be noted, however, that bio samples are more susceptible to a new form of interpretation, as they are often kept for interpretation and only for interpretation. That makes, in some situations, biological materials even more worrisome in terms of privacy than information derived from analysis of such materials.

55 Similarly, the interchangeable usage of the words 'information' and 'data' both in the law and policy circles — including in the DPD — and in our day-to-day usage is yet another tribute to similar effects that they produce implying absence of a real reason to distinguish the two. Two reasons are worth mentioning for such interchangeability. The first one explains why we, hitherto, use the two words interchangeably, and the second pertains to why we will, perhaps, continue to do so even more in the future.

56 First, information derived from interpretation of data can then be recast and used as data for another interpretation in a way that we are tempted to use the two words interchangeability.[94] From a given national census, for instance, sex and age 'data' can be used to derive 'information' about the percentage of the youth in a relevant population which can, in turn, be used as 'data' for youth centered policy making. In the same token, information derived from biological materials can be used for another analysis as data.

57 Secondly, pervasive, repeated and systematic extraction of information from human biological materials would eventually end up making the bio-samples themselves 'information' mainly because the extraction is of such extensive nature and the sole reason they are stored is for information. This trend can be paralleled with the gradual change in meaning of the search engine 'Google'. Because of large scale usage of this service, 'searching' on the web by authoring some key words came to be analogous as 'Googling.' This development came from the repeated and extensive use of 'Google'

89 By 'interpretive potential' I am referring to the ability to generate (potentially) identifiable information.

90 By 'interpretation' I mean mechanisms and processes that may be employed to derive information from biological materials.

91 Taylor (2012), p. 162.

92 Ibid, p. 163.

93 Ibid, p. 164.

94 Taylor (2012), p. 42.

for indexation even if Google still remains just one search engine provider and the term does not have any semantics indicating 'search.' In a similar way, continuous and pervasive derivation of information from biological materials means that it is more and more tempting to use the two words interchangeably. Thus, a time may come when we could call 'bio-sample' as information and not just 'data.' It all depends on how easily-accessible the interpretative frameworks are and how frequently we use them.

2. Enhancing Bio-bank Regulation

58 The other major benefit expected from the inclusion of biological materials in to the concept of personal data is the anticipation of filling the regulatory vacuum in bio-banks. What makes this regulatory vacuum all the more germane to data protection discourse is the fact that it is manifested in the incapacity to effectively preserve the fundamental rights of privacy and data protection of participants, even though such is one of the primary objectives of bio-bank regulations. In this regard, an EU Commission's study on Bio-bank governance notes 'one of the main challenges has been, and still is, to identify ways to protect the autonomy and dignity of patients and research participants and their fundamental rights (e.g. private life and data protection, especially in case of loss of control on personal data/data misuse, discrimination) with fostering the public interest in carrying out medical research to address the central public health challenges (such as cancer, cardiovascular and metabolic diseases.)'[95] The same study reiterates absence of clear legal framework governing bio-banks as one of the major problems for the imbalance against protection of fundamental rights.[96] With relatively comprehensive rules and well-established enforcement mechanisms, data protection laws can serve as a better mechanism, even though the latter also have their own limitations.[97]

3. Just 'About Us' or but not 'Us' (Moral Plea)

59 As it stands today, the existing data protection regime in the EU protects information that *relates to* us but does not, strictly speaking, protect us. Even by layman standards, leaving out bio-materials may not be considered as the right thing to do. To make full sense of how morally questionable the current system is, one needs only to consider two facts against which this moral claim should be assessed. One is the fact that the starting point of discussions on the right to privacy has usually been a concern for bodily integrity. The division between informational privacy and bodily privacy are made fictitious by technological development, especially since the past decade. In this regard, the Australian Office of Federal Privacy Commissioner, back in 2002, rightly noted:

> *"... an attempt to maintain a clear demarcation between different types of privacy protection may be problematic in light of new technologies which involve the merging of biology, mathematics and computer science, namely, biometrics and bioinformatics. Such developments give rise to new forms of body templates or records which further blur the distinction between personal information and its source in individual humans, rendering the concepts of information privacy and bodily privacy inherently interrelated."*[98]

60 Secondly, in the face of such division, the regulatory landscape pertaining to bio-banks has largely been uncoordinated and ineffective, as noted above. Therefore, not only does this fact stand in contrast to the original conception of privacy, thus failing the very essence of its inception, but the human body is also failed by the disarray in the regulation of bio-banks.

61 Against these two backgrounds alone, is it morally indefensible to protect information about individuals but not individuals themselves, or a sample taken from them. The human body or a sample taken from it is one of the most sacred representations of one self. To argue that a fingerprint represents the finger while a sample doesn't represent the person is not only morally questionable but also logically weak. Distinction should also be made between the human body/sample as source of data/medium and other sources of data as integrity and privacy is often an issue when human body is involved.

E. The Consequences of Treating Biological Materials as Personal Data

62 Despite crumbling conceptual rigor that distinguishes human biological materials from data/information, and various pragmatic considerations that increasingly challenge such distinction, collapsing differences that were maintained in the

95 EU Commission(2012), *Bio-banks for Europe: A challenge for governance*, P. 45.

96 Ibid, p.46-48.

97 See, Bygrave (2010), p. 21-22, for details and references on similar problems of some European national bio-banks regulations.

98 ALRC and AHEC, (2003), *Essentially Yours*, p.280.

regulatory discourse for such a long time is not without its own drawbacks.

I. Over Stretching the Scope of Data Protection Laws

63 The inclusion of a new subject matter in to the scope of application of data protection law, to the least, demands a closer look at the existing subjects of the law to see whether it properly fits with the law's regulatory apparatus. Data protection law already suffers from regulatory overreaching in the sense that its rules tend to apply *prima facie* to a wide range of activities with relatively scant chance of being respected, let alone enforced.[99] The Data Protection Directive is, for instance, said to have a long arm with application to multiple actors based outside the European Union.[100]

64 Article 4(1) (c) of the data protection Directive epitomizes one such long arm. This provision subjects any controller located anywhere in the world to European data privacy regime when it utilizes an equipment situated in any member state for the purpose of processing personal data.[101] The General Data Protection Regulation, perhaps, does more than the directive in this regard.[102]

II. Centrality of Consent

65 The other problem in the inclusion of biological materials in to the scope of data protection regime comes from the inadequacy of the current rules to meet the normative position of consent in the laws currently concerned with regulation of biological materials. The fundamental principle that underpins the governance framework of human biological materials in general is the need to obtain voluntary and informed consent of participants. The history of how biological materials were governed — such as by the European Convention on Human Rights and Biomedicine, and Declaration of Helsinki[103] show that consent is unequivocally important as it occupies a central normative position. The Convention on Human Rights and Biomedicine stipulates that an intervention in the health field may *only* be carried out after the person concerned has given *free and informed consent* to it. This person shall beforehand be given appropriate information as to the purpose and nature of the intervention as well as on its consequences and risks.[104] The interests and welfare of the human being shall prevail over the sole interest of society or science.[105] In addition to securing free and informed consent for the purposes of medical research the convention requires other safe guards like making sure that there is no alternative of comparable effectiveness to research on humans.[106]

66 In this regard, the Data Protection Directive or the Regulation are too liberal to accommodate what is customarily and legally expected if biological materials were to be governed by these regimes. That requires the role of consent under the directive and the General Data Protection Regulation to be seen more closely.

▪ Does Consent Play Central Role under the Current EU Data Protection Regime?

67 Broadly speaking, data subject's consent is one of many control mechanisms[107] in which data subjects, as active actors in data protection laws[108], influence the data processing operations of controllers. Though there are some non-negligible reasons, in particular for sensitive personal data, more convincing evidences suggest that consent does not play any central role in the existing data protection regime. There are, however, more stringent requirements for consent of the data subject with regard to processing sensitive data. In principle, processing sensitive personal data is prohibited. In addition, the jurisprudence of the European Court of Human Rights (ECtHR) in some of the cases — such as *Z v Finland* and *MS v Sweden* — suggest normative importance of data subjects' consent regarding sensitive data, particularly, medical information.

99 Bygrave(2010), p. 22.

100 See Lokke Moerel, "The long arm of EU data protection law: Does the Data Protection Directive apply to processing of personal data of EU citizens by websites worldwide?" *International Data Privacy Law*, Vol. 1, No. 1 (2011): 28-46.

101 Bygrave (2014), p. 202.

102 The Regulation applies to controllers not established in the Union when they process personal data of European residents in relation to the offering of goods and services to them and monitoring of their behaviour (Article 3(2)). The Parliament's version of the regulation, which has also made to the compromise text, even goes on saying that the goods and services need not be offered for consideration (The Parliament's reading and the Compromise text of the GDPR, Article 3(2)).

103 World Medical Association (WMA), World Medical Association Declaration of Helsinki: ethical principles for medical research involving human subjects, 2008.

104 Council of Europe, Convention for the Protection of Human Rights and Biomedicine, Article 5.

105 Ibid, Article 2.

106 Ibid, Article 16.

107 Other control mechanisms in which data subjects can influence processing of personal data can be: opposing a particular processing or withdrawing consent.

108 We have two additional main actors in the operative sphere: DPAs and controllers (Bygrave 2014, p. 18-19).

Thus, the problem can, somehow, be mitigated by the fact that consent enjoys relative central role under the directive with regard to sensitive data. That is because biological materials would most probably belong to the category of sensitive data as data concerning health under article 8(1) of the directive.

68 Generally, however, under articles 7 & 8 of the DPD, consent is not only just one precondition among the alternatives for legitimate processing, member states are also allowed to introduce new grounds for reasons of substantial public interest.[109] Similarly, the EU Charter of Fundamental Rights provides: personal data can be processed "*on the basis of the consent of the person concerned or some other legitimate basis laid down by law.*"[110] While consent is expressly mentioned, the Charter makes it clear that personal data can be processed on the basis of other legitimate grounds laid by law.

69 In addition, from a pragmatic viewpoint, the DPD incentivizes data controllers to first utilize other preconditions — such as the one under article 7(f) — and employ consent when a processing exercise can't be justified under those grounds. This flows from the cost and delay involved from securing consent and, the desire to avoid the possibility of refusal by the data subject.

70 Though all these facts demonstrate absence of normative priority, a closer look at at-least some of the preconditions tells us that they are framed on the assumption that 'if the data subjects were asked to consent, they would have agreed to the processing.' The preconditions like 'necessary to protect vital interests of the data subject' and 'necessary for performance of contract in which the data subject is a party' are examples in point. Therefore, I would argue, that the other preconditions also aren't completely devoid of an element of consent. Consent can still be read in to them in its broadest and indirect/implied sense.

71 However, what is problematic is not just that consent does not play a central role under the existing regime; there are also convincing arguments against a central role of consent as a precondition for data processing. First, there are legal problems in properly delineating the requirements of consent, for instance, how informed should consent be under article 2(h) of the DPD. Secondly, the degree of choice presupposed by consent mechanisms will often not be present for certain services or products, particularly those offered by data controllers in a monopoly (or near-monopoly) position.[111] Thirdly, despite the requirements of informed consent and notification (for instance articles 10&11 of DPD) controllers will typically have greater knowledge about their data processing operations than will the subjects.[112] The asymmetry will further weaken the 'informed' nature of data subject's consent. Finally, problems of consensual exhaustion, laxity and apathy – in addition to ignorance and myopia – can reduce the amount of care that data subjects invest in their decisions of whether or not to consent.[113]

72 Therefore, not only is it doubtful that consent plays a central role in the processing of personal data — including sensitive data — but it is also, arguably, not desirable that it plays such a central role. Yet, it remains central in other laws traditionally concerned with human biological materials. Thus, the extension of the DPD or the GDPR[114] to biological materials only poorly meets the central normative position of 'consent' in laws currently governing biological materials. As indicated earlier, this problem can, somehow, be mitigated by the fact that consent enjoys relative central role under data protection laws when it comes to sensitive data, the category to which biological materials would most probably belong.

III. Enforcement

73 Yet another major concern in trying to extend the scope of data protection regime is the fear that the enforcement of the law, that includes biological materials, would require strong data protection authorities with additional competence to handle the particularities of biological materials. This problem gets even more alarming because the ability of data protection authorities to ensure effective compliance of the law is already under pressure as they are chronically under-resourced.[115] The addition of biological materials in their task sheet, thus, fuels the difficulty. Not only will the authorities need additional material resources, but they may also want personnel with broad and interdisciplinary professional background.

109 DPD, Article 8(4).

110 EU Charter of Fundamental Rights, Article 8(2).

111 Lee Bygrave, & Dag Schartum, (2009), "Consent, Proportionality and Collective Power," In Serge Gutwirth *et al.* (eds.), *Reinventing Data Protection?* Springer, p.160.

112 Ibid. p.160-161.

113 Ibid. p.161.

114 With some clarifications on the requirement of 'consent' the Regulation remains structurally the same with regard to the normative position of consent as a ground of processing personal data.

115 Bygrave (2010), p. 22.

F. Conclusion

74 The analysis in this article is made in an endeavor to challenge the conceptual predispositions behind one of the building blocks of the definition of personal data under the current and *en route* EU data protection rules: the terms *'information/data.'* Despite their importance, these terms are often taken for granted and insufficiently, if at all, defined in data protection discourse. As technology, particularly in the field of bio technology develops, however, a workable definition is increasingly needed because the blurring of the boundary between human body and technology may trigger application of laws intended for the informational world — such as data protection — to the biological world.

75 A close look at the Data Protection Directive, in this regard, reveals the absence of a positive intention by the architects of the directive to consider biological materials as data/information. While it makes mention of 'biological materials,' it does not appear that the General Data Protection Regulation is intended to be applicable to such materials. The DPD and its preparatory materials indicate that the architects did not have the issue of biological materials on the table. The same assumption, however, can't be made about the General Data Protection Regulation as it introduces numerous tempting terminologies. By introducing proper terminologies such as — biological materials and genetic data — the architects of the regulation tried to create an appearance that the regulation applies to biological materials without providing any real substance in this regard.

76 The question of whether biological materials *should be* treated as personal data is far from consensus. Scholars who pay more attention to pragmatic considerations have forwarded the view that biological materials should be regarded as personal data/information. Other scholars, commentators and data protection enforcement authorities have opposed this view mainly based on conceptual logic, arguing that data is a formalized representation of objects while information comprises cognitive elements involving comprehension of that representation.[116]

77 However, a range of developments in molecular biology and nano-technology, largely mediated by advances in ICT, are at odds with the conceptual distinction between data and information. First, proteins — which make up the basis for almost everything in the human body — are made as per 'the information' obtained by reading the order of strands of nucleotides in our DNA. Thus, information lies at the very origin of life. Secondly, ambitious scientific initiatives such as the BRAIN (Brain Research through Advancing Innovative Neuroethologies) — which intends to decode neurons in our brain much like the Human Genome Project did for our genome — may lead to our continued existence as information. Thirdly, the ongoing human enhancement projects (HEP) are clouding the distinction between the human body and technology. Moreover, proliferation of bio-banks and the increasing deployment of biometric technologies are converting aspects of our bodies in to processable digital data.

78 In addition, multiple pragmatic considerations beseech the collapse of the distinction between data, as carrier, and information, as a result of processing data. First, it is difficult to find distinguishable interpretive potential between data and information; it all turns on availability of the right interpretive framework. Secondly, the lacunae in the regime governing bio-banks might be assisted by the more comprehensive rules under data protection, which also possesses better enforcement mechanisms. And finally, considering biological materials only as a medium may, sometimes jeopardize our fundamental rights even more, thus, making maintenance of the distinction morally indefensible.

79 Despite a crumbling conceptual rigor that distinguishes human biological materials from data/information and various pragmatic considerations that increasingly challenge such distinction, collapsing differences that have been maintained in the regulatory discourse for such a long time is not without its own drawbacks. First, it will overstretch the rules that are already said to have a long arm which may be counterproductive for their effective enforcement. Secondly, while 'consent' enjoys a relatively central role under the directive when with regard to sensitive data—the category to which biological materials would most probably belong — it is doubtful that consent plays or would play a central role in the processing of personal data in general. As consent remains central in other laws traditionally concerned with human biological materials the extension of the DPD or the GDPR to biological materials only poorly meets the normative position of consent maintained by these laws. Finally, extending biological materials to the data protection regime would demand DPAs to have more financial and human resources with the requisite skills to handle the peculiarities of biological materials.

116 Bygrave (2015), p. 6-7.

* *Worku Gedefa Urgessa*, University of Oslo — Norwegian Research Center for Computers and Law, Oslo, Norway. The author would like to thank Professor Lee A. Bygrave for his valuable support and constructive comments on the earlier versions of this work.

Ten Questions for Future Regulation of Big Data

A Comparative and Empirical Legal Study

by **Bart van der Sloot and Sascha van Schendel***

Abstract: Much has been written about Big Data from a technical, economical, juridical and ethical perspective. Still, very little empirical and comparative data is available on how Big Data is approached and regulated in Europe and beyond. This contribution makes a first effort to fill that gap by presenting the reactions to a survey on Big Data from the Data Protection Authorities of fourteen European countries and a comparative legal research of eleven countries. This contribution presents those results, addressing 10 challenges for the regulation of Big Data.

Keywords: Big Data; Empirical; Comparative; Survey; Data Protection Authorities; Comparative Legal Research

Recommended citation: Bart van der Sloot and Sascha van Schendel, Ten Questions for Future Regulation of Big Data: A Comparative and Empirical Legal Study, 7 (2016) JIPITEC 110 para 1.

A. Introduction

1 Big Data is a buzzword used frequently in both the private and the public sector, the press, and online media. Large amounts of money are being invested to make companies Big Data-proof, and governmental institutions are eager to experiment with Big Data applications in the fields of crime prevention, intelligence, and fraud, to name but a few areas. Though the exact nature and delineation of Big Data is still unclear, it seems likely that Big Data will have an enormous impact on our daily lives. Positively, undoubtedly, but there are also inherent risks to Big Data applications, as it might result in discrimination, privacy violations, and chilling effects. The ideal situation would be to have an adequate framework in place that will ensure that the beneficial uses of Big Data are promoted and facilitated, while the negative effects are mitigated or sanctioned. This contribution provides building blocks for developing such a framework, by giving an overview of the experience in the use and regulation of Big Data in 23 countries, aiming in particular at the use of Big Data by governments.

2 The research presented in this article was conducted in two phases. The first phase involved desk research and looked at Big Data policies, legislation and regulation in a number of countries. Second, a questionnaire was sent to several European DPAs. The desk research examined eleven countries. These countries were selected on the basis of three criteria. The first was global coverage – the research sought to be as representative as possible to provide a full picture of global developments in relation to Big Data, which is by nature an international phenomenon. Therefore, at least one country from each continent (with the exception of Antarctica) was examined. The second criterion was an estimation of the potential value of the expected outcomes of the research – some countries are more innovative and ambitious

than others in terms of technological developments such as Big Data. Thirdly, the role a country plays in international politics was taken into account; on that basis, China rather than South Korea was studied, even though the latter country is often in the forefront of technological developments. Based on these three criteria Australia, Brazil, China, France, Germany, India, Israel, Japan, South Africa, the United Kingdom and the United States were selected. The desk research focused on two issues in particular. First, government policy decisions were analyzed, as were initiatives related to this topic, such as governments using Big Data themselves or stimulating the use of Big Data in the private sector, either through financial support or by engaging in partnerships. Second, research was carried out on legislation and case law revolving around Big Data in the selected countries. It should, again, be noted that this study is not exhaustive – there is, undoubtedly, a myriad of relevant laws, court cases and DPA reports that are not discussed here.

3 In studying the eleven countries, almost exclusive use was made of official sources, especially government websites. The reason for this is that it is often difficult to establish the reliability of foreign sources. This choice does, however, imply that this article mainly presents a picture of the governmental view of Big Data and of governmental regulation. Criticism of those initiatives and autonomous processes in the private sector remain largely undiscussed. This bias was accepted as a tradeoff in order to guarantee the reliability of the sources studied. When discussing Israel, however, use was made of online newspaper articles from Israeli news sources and a published online interview, because this provided vital information and because the news-source was regarded as reliable. The information from these sources was not available on government websites, but was nonetheless considered essential.

4 Publications on government websites and in press releases about new initiatives were selected by using terms related to Big Data, both in the official language of the country concerned and in English, such as 'data mining', 'data analytics', 'data projects', 'Big Data initiatives', etc. Several countries have a Ministry of Science and Technology, or a similar ministry. Those ministries were taken as the starting point of the research in those countries. General search engines were also used to scan government initiatives related to Big Data, by limiting the search to the national public domain of the country concerned. For case law and legislation, the official national search engines and general search engines were used. The search terms entered here were related to Big Data, privacy and data protection, such as 'data protection', 'privacy', 'surveillance', etc. This process yielded a list of government initiatives, legislation and relevant jurisprudence. The sources consulted and the full list of references used for this article are listed in a working paper published earlier.[1]

5 The results of the comparative desk research can be found in Appendix I and the results of the survey in Appendix II to this contribution. It has to be stressed that not all governments and governmental agencies use the term Big Data when creating, operating on, or using large scale data bases. That is why this study primarily identifies those initiatives that have been identified as Big Data by the government itself, or when it has used terms that are related to it. This means that many uses of large scale databases by governmental agencies are not included in this study. When analyzing the countries, six questions were kept in mind: 'Is a specific definition of Big Data used?', 'Is Big Data used within the government?', 'Is there a public-private partnership?', 'To what goal is Big Data used by the government?', 'Which laws are especially relevant for Big Data?' and 'Are there judicial decisions relating to Big Data?'

6 A relatively short and simple questionnaire was designed for the survey, so as to increase the potential response of the DPAs. The accompanying email, as well as the introduction to the survey, briefly explained the goal of the survey. The survey comprised six questions: 1. Are you familiar with the debate on Big Data? If so, how would you define Big Data? (max. 500 words) 2. Are there prominent examples of the use of Big Data in your country, especially in the law enforcement sector, by the police or by intelligence services? (max. 500 words) 3. Have you issued any decisions/reports/opinions on the use of Big Data? If so, could you provide us with a reference and your main argument? (max. 500 words) 4. Are there any legal cases/judgements by a court with regard to (privacy/data protection) violations following from Big Data practices in your country? If so, could you provide us with a reference and the main consideration of the court? (max. 500 words) 5. Which legal regimes are applied to Big Data/ is there a special regime for Big Data in your country? Are there any discussions/plans in parliament to introduce new legislation to regulate Big Data practices? (max. 500 words) 6. Are there any final remarks you want to make/suggestions you have for further research? (max. 500 words)

7 The reason for choosing these questions for the desk research and the survey is that the background of this study is a project by the Netherlands Scientific Council for Government Policy (WRR). The WRR

1 <http://www.wrr.nl/fileadmin/en/publicaties/PDF-Working_Papers/WP_20_International_and_Comparative_Legal_Study_on_Big_Data.pdf>. The literature studied for this article can be found here. <http://www.wrr.nl/fileadmin/nl/publicaties/PDF-Rapporten/rapport_95_Big_Data_in_een_vrije_en_veilige_samenleving.pdf>.

is an independent advisory body for the Dutch government. The task of the WRR is to advise the government on issues that are of great importance for society in the intermediate and longer term. The reports of the WRR are not tied to one policy sector but rather touch on various terrains and policy sectors; they are concerned with the direction of government policy for the longer term. The members of the WRR are established university professors who have often worked on policy related subjects and/or have made tracks in public administration themselves. The Dutch government had requested the WRR to advise on the regulation of Big Data, taking into account how privacy and security should be assessed in the deployment of big data analytics in security related policies. Questions that were suggested to be addressed include whether a distinction needs to be made between access to and use of data, how transparency and individual rights can be guaranteed in Big Data practices and what the likely impact of the emergence of quantum computing will be. In addition to the policy advice, published in the form of a report for the Dutch government,[2] a scientific book was delivered[3] and a number of working papers were written to do indicative research,[4] which were used as building blocks for the report to the government. This article is based on one of those working papers.[5]

8 The DPAs in all 28 EU Member States were emailed with a request to complete the survey. Requests were also sent to the DPAs in three non-EU countries, namely Norway, Serbia and Switzerland, because a short preliminary study had shown that they might have specific expertise in relation to Big Data. DPAs that did not respond within the period specified in the initial request were sent a reminder; those that did not respond to this mail either were sent a final reminder. In most cases, the questionnaire was sent to the general contact address as posted on DPA's website. However, since the French website lists no general email address, personal contacts were used to email two specific employees of the CNIL. For three other DPAs (Germany, the Netherlands and Norway), in addition to an email to the general email address, an email was also sent to a specific individual employee. For other DPAs, either no such personal contacts existed or they existed but it was not necessary to use them because a response had been received. Eventually, of the 31 DPAs included in the survey, 18 responded: Austria, Belgium, Croatia, Denmark, Estonia, Finland, France, Hungary, Ireland, Latvia, Lithuania, Luxembourg, the Netherlands, Norway, Slovakia, Slovenia, Sweden and the United Kingdom. Four of these (Austria, Denmark, Finland and Ireland) were negative responses, stating that the DPA in question would not participate in the study. Consequently, about half of the DPAs invited to join the survey have actually responded. The results found in this study can, therefore, not be seen as determinative but as indicative of possible trends, feelings and attitudes towards Big Data. It should be taken into account that those DPAs that have already dealt with Big Data projects would be more likely to respond to such a survey than those that haven't.

9 Rather than presenting the bare facts, listing the regulatory initiatives in the various countries studied and the answers from the DPAs, this article uses the insights gained from those results to shine light on some of the most difficult questions regulators have to answer when deciding on future regulation of Big Data. These questions are partly based on those asked in the survey and partly follow from the desk research. Additional questions have been added in order to present the most interesting findings from both the desk research and the survey in an orderly fashion. Ten issues/questions are discussed in more detail: (1) What is the definition of Big Data? (2) Is Big Data an independent phenomenon? (3) Big Data: fact or fiction? (4) What is the scope of Big Data? (5) What are the opportunities for Big Data? (6) What are the dangers of Big Data? (7) Are the current laws and regulations applicable to Big Data? (8) Is there a need for new legislation for Big Data? (9) What concept should be central to Big Data regulation? (10) How should the responsibilities be distributed? These questions will be discussed in the subsequent sections. The article will conclude with a short summary of the main findings.

B. What is the definition of Big Data?

10 The first choice when it comes to regulating Big Data is to determine a definition and delineation of Big Data. Three definitions were encountered a number of times in both the desk research and in the survey. First, the Article 29 Working Party holds that Big Data refers to the exponential growth, both in the availability and in the automated use of information. It refers to gigantic digital datasets held by corporations, governments and other large organizations, which are then extensively analyzed using computer algorithms. Big Data can, according to the Working Party, be used to identify more general trends and correlations, but it can also be

2 <http://www.wrr.nl/fileadmin/nl/publicaties/PDF-Rapporten/rapport_95_Big_Data_in_een_vrije_en_veilige_samenleving.pdf>.

3 <http://www.wrr.nl/fileadmin/en/publicaties/PDF-Verkenningen/Verkenning_32_Exploring_the_Boundaries_of_Big_Data.pdf>.

4 <http://www.wrr.nl/publicaties/working-papers/>.

5 <http://www.wrr.nl/fileadmin/en/publicaties/PDF-Working_Papers/WP_20_International_and_Comparative_Legal_Study_on_Big_Data.pdf>.

processed in order to directly affect individuals.[6] Second, the European Data Protection Supervisor (EDPS) suggests that Big Data means large amounts of different types of data produced at high speed from multiple sources, whose handling and analysis require new and more powerful processors and algorithms. Not all of these data, the EDPS points out, are personal, but many players in the digital economy increasingly rely on the large scale collection of and trade in personal information. As well as benefits, these growing markets pose specific risks to individual's rights to privacy and to data protection, the EDPS warns.[7] Third, and perhaps most well-known, the Gartner Report focusses on three matters when describing Big Data: increasing volume (amount of data), velocity (speed of data processing), and variety (range of data types and sources). This is also called the 3V model or 3V theory.[8]

11 The desk research also showed that a number of countries apply their own definition of Big Data. For example, in Germany, Big Data is defined as 'das Synonym für den intelligenten Umgang mit solchen großen oder auch heterogenen Datenmengen' (synonymous with the intelligent use of large or heterogeneous datasets).[9] The Podesta Report (United States) builds on the Gartner definition and suggests that there are "many definitions of 'Big Data' which may differ depending on whether you are a computer scientist, a financial analyst, or an entrepreneur pitching an idea to a venture capitalist. Most definitions reflect the growing technological ability to capture, aggregate, and process an ever-greater volume, velocity, and variety of data. In other words, 'data is now available faster, has greater coverage and scope, and includes new types of observations and measurements that previously were not available.' More precisely, Big Datasets are 'large, diverse, complex, longitudinal, and/or distributed datasets generated from instruments, sensors, Internet transactions, email, video, click streams, and/or all other digital sources available today and in the future."[10]

12 Finally, several DPAs also gave their own definition of Big Data when completing the survey, or referred to specific definitions used in their country. For example, the Estonian DPA describes Big Data as collected and processed open datasets, which are defined by quantity, plurality of data formats, and data origination and processing speed.[11] The French DPA refers to a definition adopted by the French General Commission on terminology and neology (Commission générale de terminologie et de néologie). The official translation of Big Data in French is 'mégadonnées', which stands for data, structured or otherwise, whose very large volume require appropriate analytical tools. The DPA of Luxembourg suggests that Big Data stems from the collection of large structured or unstructured datasets, the possible merger of such datasets, as well as the analysis of these data through computer algorithms. These datasets can usually not be stored, managed and analyzed with average technical means due to their size, it also points out. The Dutch DPA primarily points to the 'volume' aspect of Big Data and argues in particular that Big Data is all about collecting as much information as possible, storing it in ever-larger databases, combining data that is collected for different purposes and applying algorithms to find correlations and unexpected new information. The DPA from Slovenia not only refers to the use of different types of data, acquired from multiple sources in various formats, but also to predictive analytics used in Big Data. Finally, the Swedish DPA suggests the concept is particularly used for situations where large amounts of data are gathered in order to be made available for different purposes, not always precisely determined in advance.

13 It can be seen from this list of definitions that a number of components are regularly mentioned. Broadly, they relate to three states of Big Data processing, namely the collection, analysis and use of data. When it comes to collecting data, Big Data is about collecting large amounts of data (volume) from varied (variety) and often unstructured data sources. With regard to analyzing the collected data, Big Data revolves around the speed (velocity) of the analyses and the use of certain instruments such as algorithms, machine learning and statistic correlations. The results are often predictive in nature (predictive analytics) and are formulated at a general or group level. The results are usually applied by means of profiling. Many of the definitions contain some of these components; none of the definitions used mention all of these components. Consequently, none of these elements should be seen

6 <http://ec.europa.eu/justice/data-protection/article-29/documentation/opinion-recommendation/files/2013/wp203_en.pdf>.

7 <https://secure.edps.europa.eu/EDPSWEB/edps/Consultation/big_data>. See also: <https://secure.edps.europa.eu/EDPSWEB/webdav/site/mySite/shared/Documents/Consultation/ <Big%20data/14-07-11_EDPS_Report_Workshop_Big_data_EN.pdf>.

8 <http://blogs.gartner.com/doug-laney/files/2012/01/ad949-3D-Data-Management-Controlling-Data-Volume-Velocity-and-Variety.pdf>.

9 References to the situation in the different countries studies might be found in Appendix I or at: <http://www.wrr.nl/fileadmin/en/publicaties/PDF-Working_Papers/WP_20_International_and_Comparative_Legal_Study_on_Big_Data.pdf>.

10 <https://www.whitehouse.gov/sites/default/files/docs/big_data_privacy_report_5.1.14_final_print.pdf>.

11 References to the answers to the survey might be found in Appendix II or at: <http://www.wrr.nl/fileadmin/en/publicaties/PDF-Working_Papers/WP_20_International_and_Comparative_Legal_Study_on_Big_Data.pdf>.

as essential – that is, if one or more of these elements do not apply, it does not follow that the phenomenon being studied is not Big Data. Rather, these elements should be seen as parameters; if none of the elements apply, the phenomenon is definitely not Big Data; if all the elements apply, the phenomenon being studied definitely is Big Data. Mostly, however, it will somewhere in between. It is impossible to say, for example, how big a dataset must be in order to qualify as Big Data; although Big Data usually works with combined datasets, it is conceivable that one enormous dataset could qualify as Big Data; although Big Data usually (partially) works with unstructured data, this is not a condition sine qua non; etc.

C. Is Big Data an independent phenomenon?

14 The overview of definitions already shows that Big Data should not be seen as an isolated phenomenon. It is a new phenomenon which by its nature is strongly connected to a number of technical, social and legal developments. This conclusion is supported by the desk research, which also found that Big Data is intertwined with several other terms. For example, lots of Big Data initiatives are linked to Open Data. As the name suggests, Open Data is the idea that (government) data should be placed in the public domain. Traditionally, it has been linked to efforts to increase transparency in the public sector and give more control over government power to media and/or citizens. The Estonian DPA is in particular very explicit about the relationship between Open Data and Big Data, as it defines Big Data as "collected and processed open datasets, which are defined by quantity, plurality of data formats and data origination and processing speed". The desk research also shows a clear link between the two concepts in countries such as Australia, France, Japan and the United Kingdom.

15 Linked to Open Data is the idea of re-use of data. Yet, there is one important difference. While Open Data has traditionally been concerned with transparency of and control over government power, the re-use of (government) data is specifically intended to promote the commercial exploitation of the data by businesses and private parties. The re-use of Public Sector Information is fostered through the PSI Directive of the European Union. More generally, re-use refers to the idea that data can be used for a purpose other than that for which they were originally collected. Obviously, the link between Big Data and re-use is often made, as appears both from the desk research and from the survey. The Norwegian DPA, for example, uses the definition of Big Data of the Working Group 29, 'but also add what in our opinion is the key aspect of Big Data, namely that it is about the compilation of data from several different sources. In other words, it is not just the volume in itself that is of interest, but the fact that secondary value is derived from the data through reuse and analysis.' The desk research also showed a link between the two concepts. In France, for example, Big Data is primarily seen as a phenomenon based on the re-use of data for new purposes and on the combination of different data and datasets.

16 The term 'Internet of Things' refers to the idea that more and more things are connected to the Internet – cars, lampposts, refrigerators, clothing, or any kind of object. This opens the way for the development of smart devices – for example, a refrigerator that records when the milk has run out and automatically reorders. By fitting all objects with a sensor, large quantities of data can be collected. As a consequence, Big Data and the Internet of Things are often mentioned in the same breath. An example is the DPA of the United Kingdom, which notes 'that Big Data may involve not only data that has been consciously provided by data subjects but also personal data that has been observed (e.g. from Internet of Things devices), derived from other data or inferred through analytics and profiling.'

17 Because of the applications of the Internet of Things and the constantly communicating devices and computers, the development of smart products and services has spiraled. Examples of such developments are smart cities, smart devices and smart robots. The desk research indicates that a number of countries – for example, the United States, China and the United Kingdom – make a link between such developments and Big Data systems. The Luxembourg DPA also emphasizes the relationship with smart systems, such as smart metering. 'At a national level, a system of smart metering for electricity and gas has been launched. The project is, however, still in a testing phase. - The CNDP has not issued any decisions, reports or opinions that are directly dealing with Big Data. The Commission has, however, issued an opinion in a related matter, namely with regard to the problematic raised by smart metering. In 2013, the CNDP issued an opinion on smart metering. The main argument of the opinion highlights the necessity to clearly define the purposes of the data processing, as well as the retention periods of the data related to smart metering.'

18 A term that is often associated with Big Data and is sometimes included as part of the definition of Big Data is 'profiling'. As increasingly large datasets are collected and analyzed, the conclusions and correlations are mostly formulated at a general or group level. This mainly involves statistical correlations, sometimes of a predictive nature. Germany is developing new laws on profiling and a number of DPAs emphasize the relationship between

Big Data and profiling; for example, the DPAs of the Netherlands, Slovenia, the UK and Belgium. The latter argues that 'we expect that de new data protection regulation will be able to provide a partial answer (profiling) to Big Data issues (legal interpretation of the EU legal framework).'

19 Similar to the term profiling, 'algorithms' is used in many definitions of Big Data. This applies to the definition by Article 29 Working Party, the EPDS and a number of DPAs responding to the survey, such as those of Luxembourg, the Netherlands and the UK. A number of countries also have a special focus on algorithms. To provide an example, in Australia, a 'Program Protocol' has been developed – a report may be issued which contains the following elements: a description of the data; a specification of each matching algorithm; the anticipated risks and how they will be addressed; the means of checking the integrity of the data; and the security measures used.

20 To provide a final example, cloud computing is also often associated with Big Data processes. In China and Israel, especially, the two terms are often connected to each other. For example, the Chinese vice-premier stressed that the government wants to make better use of technologies such as Big Data and cloud computing to support innovation; according to the Prime Minister, mobile Internet, cloud computing, Big Data and the Internet of Things are integrated with production processes, and will thus be an important engine for economic growth. In Israel, the plan is for the army to have a cloud where all data is stored in 2015 – there is even talk of a 'combat computing cloud', a data center that will make different tools available to forces on the ground. Some DPAs also suggest a relationship between cloud computing and Big Data; the Slovenian DPA, for example, states that 'new concepts and paradigms, such as cloud computing or Big Data should not lower or undermine the current levels of data protection as a fundamental human right.'

21 There are other terms that are often mentioned in connection with Big Data, such as machine learning, commodification of data, datafication, securitization and risk society. It goes beyond the scope of this article to discuss all these terms in depth. What is important to note is that Big Data should be primarily viewed in its interrelationship and in conjunction with other phenomena. Big Data is a part of and, in a certain sense, the umbrella term for many of the technological and societal developments that are already taking place. This needs to be taken into account when regulating Big Data. It seems advisable for regulators to take a holistic approach to the regulation of Big Data and related phenomena.

D. Big Data: fact or fiction?

22 There is still no clarity about the extent to which Big Data processes are already being used in practice. The reactions of a number of DPAs seem to suggest that Big Data is not yet an established practice. For example, the Austrian DPA declined to participate in the survey because it had encountered few if any Big Data processes; cautious reactions were also received from the DPAs of Latvia, Lithuania and Slovakia. The Belgian DPA suggests that there is currently a lack of clarity about Big Data and refers to Gartner's hype cycle.[12] It also adds: "Most Belgian projects seem to still be in a pilot phase and the visibility of Big Data in practice is still low." However, other DPA responses show a different picture – they confirm that Big Data is a major trend, and that Big Data is playing an increasingly significant role. Some DPAs, such as Norway, have written a special report on the regulation of Big Data practices. The United Kingdom DPA has also issued a discussion paper on this topic. Furthermore, it emerged from the desk research that projects are under way in most countries that are connected to Big Data, although it should be noted that a fairly broad approach was taken in the desk research to what qualified as 'Big Data'.

23 The picture that emerges from all of the foregoing is one in which Big Data plays a minor role in most countries at present but is set to become increasingly important. Big Data should, therefore, not be seen as either an actual practice or as a fiction, a hype that will blow over, but rather as a trend that will play a major role in five years' time and will have a significant impact on the government sector, on business, and on citizens' everyday life in the future. What is clear from the desk research is that in most countries the government feels it is missing out on this important trend. While industry is investing billions in Big Data projects, many governments are – or feel they are – lagging behind. This is why many governments are now beginning to invest heavily in Big Data projects.

24 To give a few examples, the desk research showed that in the United States, more than $200 million was reserved for a research and development initiative for Big Data, which was to be spent by six federal government departments; the army invested the most in Big Data projects, namely $250 million; $160 million was invested in a smart cities initiative, investing in 25 collaborative ventures focused on data usage. In the United Kingdom, £159 million was spent on high-quality computer and network infrastructure, there was £189 million in investments to support Big Data and to develop the UK's data infrastructure, and £10.7 million will be spent on

12 <www.gartner.com/technology/research/methodologies/hype-cycle.jsp>.

a center for Big Data and space technologies. In addition, £42 million will be spent on the Alan Turing Institute for the analysis and application of Big Data, £50 million will be set aside for the 'Digital Catapult', where researchers and industry are brought together to come up with innovative products; and lastly, in February 2014 the Minister of Universities and Science announced a new investment of £73 million in Big Data. This money will be used for bioinformatics, open data projects, research and the use of environmental data. In South Africa, the government has invested 2 billion South African Rand, approximately €126.8 million, in the Square Kilometre Array (SKA) project, which revolves around very large datasets. In France, seven research projects related to Big Data were awarded a total of €11.5 million. In Germany, the Ministry of Education and Research invested €10 million in Big Data research institutes and €20 million in Big Data research; this Ministry will also invest approximately €6.4 million in ABIDA, a four-year interdisciplinary research project focusing on the social and economic impact of large data sets.

25 These are just a few examples of what is being spent by the governmental sector. In the private sector, a multiple of these sums is being spent on Big Data projects. The expectation is that these Big Data projects will develop over the next five or ten years. Only then will many of the effects of Big Data become apparent. Consequently, when designing Big Data regulations, it seems advisable for governments to develop future-proof policies that follow and, where possible, anticipate this trend. If regulators only begin to regulate this phenomenon five or ten years from now, many of the projects will have already started. The negative impact may already have materialized, and it will be difficult to adjust and alter projects and developments that have already flourished. It should also be remembered that good, clear regulation can contribute to innovation and the use of Big Data. Since the current framework applying to new Big Data projects is not always clear, some government agencies and private companies are reluctant to use new technologies for fear of violating the law. New regulation could provide more clarity

E. What is the scope of Big Data?

26 This study, and especially the desk research, shows that Big Data projects are initiated for very different purposes. In Brazil, for example, the so called Data Viva system was initially used mainly for the formulation of economic policy. In addition, the police in Sao Paulo use a system (Detecta) that is based on Big Data technology. Detecta is an intelligent system for monitoring crime. In the United Kingdom, too, Big Data is used to fight crime. The POSTnote about Big Data and crime and safety provides an example of the use of Big Data by the police. Software has been developed as part of a pilot to predict the location of burglaries, and two British police forces use software developed for predictive policing to predict the locations of crimes. The British tax and customs authority, HMRC, also uses a Big Data system, 'Connect', in which all the data held is aggregated and analyzed. This Big Data system is used to detect tax fraud and tax evasion, and is said to have led to the recovery of £2.6 billion since April 2013. The system displays relevant information in searches that is otherwise difficult to find, allows complex analyses to be performed on the development of multiple datasets simultaneously, and enables profiles to be constructed which can help uncover patterns that may indicate particular crimes.

27 In some countries, Big Data is primarily seen as a means for the government to increase its own service to citizens; prominent examples are Australia and China. Reference can also be made in this connection to the Aadhaar project that has been developed and carried out by the 'Unique Identification Authority' of India and which involves the collection of biometric and demographic data on residents of India. One of the uses of Aadhaar is 'micropayments', a means of identification which should help improve access to financial services for people living in rural areas. The identification number makes it possible to identify people in remote regions from a long distance and also reduces costs through economies of scale, making it easier for poorer people to obtain financial services. Other sectors where Aadhaar provides solutions include demographic planning, paying security social benefits and improving the identification of beneficiaries by eliminating duplicate identities. Government administrative processes should become more efficient because the authorities now have access to all relevant information at a glance.

28 Several countries see Big Data mainly as a phenomenon that can help the private economy. Germany, for example, has launched a funding initiative to support the competitiveness of it companies, and France also feels that Big Data is set to take off, especially in the private sector, through the growth of it companies and startups which help to stimulate the economy and create jobs. There are also countries, such as Japan, Germany and the United Kingdom, where Big Data is approached primarily in relation to scientific research and innovation. Israel, finally, is unique in that it also uses new technological systems for facilitating the activities of the army. It also has to be borne in mind that many intelligence services are involved with Big Data-like projects; however, often little is

known about these projects, other than what has been leaked by whistleblowers.

29 The picture that emerges from this research is that Big Data could be used in almost every sector and for almost any task. Generally, the use of Big Data can be divided into three types. Firstly, the use of Big Data for specific government tasks – examples include the use of Big Data by intelligence services, the police, tax authorities and other public bodies, for example in the context of formulating economic policies. Second, the use of Big Data by the private or semi-public sector, helping or facilitating them in achieving their specific tasks and/or goals. Examples include the use of Big Data by companies to create risk profiles, to find statistical correlations and to personalize services and advertisements, and the use of Big Data by universities and research institutes for research-related purposes. Big Data is also widely used in the medical sector; for instance, the United Kingdom has heavily promoted the use of Big Data in the healthcare sector, and the Israeli Ministry of Health has a large dataset containing medical data on the citizens of Israel and on the healthcare system. According to the Ministry, the potential benefits lie in the facilitation of a variety of healthcare functions (including assisting in the clinical decision-making process, in monitoring diseases and in proactive healthcare). Thirdly, Big Data is used by both governments and private sector companies to improve their service to citizens or customers; this might, for example, involve increasing the transparency of their activities, strengthening the control of citizens over data processing, etc.

30 These three categories should lead to different approaches to regulation. The last category is relatively unproblematic because it serves the interests of the citizen. Here, the current legislation on aspects such as the use of personal data should suffice. The situation is different when Big Data is used by governmental agencies to support their goals. It is important to distinguish between the different fields in which Big Data is used by the government. If Big Data is used for the development of economic policies, for routinely inspecting fire installations or for epidemiological research, this should be relatively unproblematic. In these instances, general patterns and statistical correlations are used to promote the efficiency and effectiveness of public policy. However, if Big Data is used by the police, a different picture emerges – while Big Data is about processing large amounts of data and detecting general patterns, the police need to investigate and possibly arrest specific individuals on the basis of concrete facts. There is a particular danger of mismatches when general profiles are applied to specific individuals. When regulating Big Data, the potential impact on citizens must be taken into account; that impact will be greater when Big Data is used by the police, intelligence services and the army than when it is used for the development of general economic policies. It also appears from the survey that several DPAs are skeptical about the use of Big Data by the police, both because of the possible impact on the citizen and because of the potential for mismatches between general profiles and specific individuals.

31 Finally, the use of Big Data in the private sector can also be problematic. It emerged from this study that two things in particular need to be taken into account. First, use can be made of data or profiles that are based on sensitive information, such a data about race, medical conditions or religious beliefs; use can also be made of categories that appear neutral but are, in fact, based on these types of information – a practice known as redlining. Second, the consequences of the use of Big Data in the private sector may also be substantial, irrespective of whether or not sensitive information is used. Where advertisements are personalized through the use of Big Data-like applications, the impact will, of course, be relatively small; however, when Big Data is used to develop risk profiles on the basis of which banks decide who may be eligible for a loan and on what terms, or by health insurers to decide who they are prepared to insure and on what terms, the consequences can be significant. Factors that could be taken into account when regulating Big Data are the impact of its use on the individual, the types of data and data analysis that are used and the potential danger of a mismatch between general profiles and specific individuals. A distinction could also be made between the type of organization that uses Big Data and the specific purpose for which it is used. The general interest that is served by the use of Big Data naturally also has an impact on what should be considered legally admissible.

F. What are the opportunities for Big Data?

32 From both the desk research and the results from the survey it appears that Big Data represents both significant opportunities and significant risks. For example, in 2013, 'France Stratégie', an advisory body to the French Prime Minister, performed an analysis of the advantages and disadvantages of Big Data. It emphasized that, on the one hand, Big Data provides for more knowledge and opportunities, but that, on the other, it may cause problems in relation to the protection of privacy and confidentiality. John Podesta also stressed this duality. He published a blog on 1 May, 2014, which discussed the results of the Working Group Review. In his blog, Podesta describes Big Data as a vital technology. He refers to the devastation and suffering caused by tornadoes

and, more implicitly, to the predictive powers of Big Data in preventing these adverse events. Big Data could provide opportunities for virtually every sector of the economy, Podesta suggests, and could make the government more efficient. The report of the US Working Group recognized in addition that Big Data carries risks, noting the fact that 'how we protect our privacy and other values in a world where data collection is increasingly ubiquitous and where analysis is conducted at speeds approaching real time.'

33 The opportunities for Big Data can be discussed relatively briefly; they follow from the field of application as discussed earlier. The first opportunity that Big Data offers lies in improving the service to the citizen or customer, improving transparency in the public or private sector and giving more control to individuals. Second, particularly in the private sector, it is expected that Big Data will lead to substantial growth in the number of companies, especially start-ups, the number of jobs and the profits generated by those companies. For example, according to the roadmap developed by the Comité de Pilotage de la Nouvelle France Industrielle (Steering Committee of the New Industrial France) headed by the French Minister for Industry, Big Data activities in France represented €1.5 billion in 2014 and would reach approximately €9 billion in 2020, with Big Data activities also generating an additional 137.000 jobs. The EDPS report on Big Data also stresses the economic potential of Big Data. 'According to the OECD, 'Big Data related' mergers and acquisitions rose from 55 in 2008 to 134 in 2012. The internet sector is hugely successful with revenue per employee in 2011, among the top 250 companies, of over $900 – over twice as high as for the ICT industry overall (OECD). Internet companies could enjoy 'economies of scope', network effects of more data attracting more users attracting more data, culminating in winner-takes-all markets and near monopolies which enjoy increasing returns of scale due to the absolute 'permanence' of their digital assets.'[13]

34 Finally, Big Data can also be used for achieving the specific objectives of organizations, institutions and government departments. Yet, the question is to what extent Big Data is actually used within the public sector. The underlying research for this article seems to indicate that most countries and DPAs mainly recognize the opportunities for Big Data in the private sector, in relation to economic growth, stimulating businesses and increasing the number of jobs. The use of Big Data by the government, and especially by governmental institutions involved with maintaining public order or protecting national security, is viewed with skepticism. The Hungarian DPA, for example, emphasizes that Big Data is primarily used in the business sphere, such by as banks, supermarkets, media and telecommunication companies. In similar fashion, the Luxembourg DPA states explicitly that it has no knowledge of prominent examples of the use of Big Data in the law enforcement sector or by police or intelligence services in Luxembourg, but points out that other actors do engage with Big Data. The Norwegian DPA argues along the same line: 'There is, as far as we know, no usage of Big Data within the law enforcement sector in Norway. In 2014, the intelligence service addressed in a public speech the need to use Big Data techniques in order to combat terrorism more efficiently. However, politicians across all parties reacted very negatively to this request and no formal request to use such techniques has since been launched by the intelligence service. The companies that are most advanced when it comes to using Big Data may be found within the telecom (e.g. Telenor) and media (e.g. Schibsted and Cxence) sectors. The tax and customs authorities have also initiated projects in which they look at how Big Data can be used to enhance the efficiency of their work.'

35 In similar fashion, the Slovenian DPA stresses that it has not seen prominent examples of the use of Big Data in Slovenia; it suggests that Big Data applications are mainly of interest in insurance, banking and electronic communications sectors, mostly to combat fraud and other illegal practices. Another important field is scientific and statistical research. 'Law enforcement use is to our knowledge currently at development stages (e.g. in the case of processing Passenger Name Records), whereas information about the use of Big Data at intelligence services is either not available or confidential in nature.' The Swedish DPA states that it has not carried out any specific supervision related to the concept of Big Data and does not have any statistics or specific information on how this is used. 'In our opinion, the law enforcement sector does not use Big Data. Their personal data processing is strictly regulated in terms of collection of data, limited purposes, etc.' Finally, the British DPA indicates that it knows 'that companies are actively investigating the potential of Big Data, and there are some examples of Big Data in practice, such as the use of telematics in motor insurance, the use of mobile phone location data for market research, and the availability of data from the Twitter 'firehose' for analytics. We do not have any specific information on the use of Big Data in law enforcement or security.'

36 Noteworthy is that many DPAs suggest that Big Data is used particularly in the private sector and less so in the public sector – in particular, the use

13 <https://secure.edps.europa.eu/EDPSWEB/webdav/site/mySite/shared/Documents/Consultation/Big%20data/14-07-11_EDPS_Report_Workshop_Big_data_EN.pdf>.

of Big Data for security-related activities by the government is rejected. Only a few DPAs, such as the Dutch DPA, refer to the use of Big Data by the government for security purposes. The desk research, however, reveals a different picture, showing that governments do, indeed, use Big Data technologies, including for security purposes. Australia is an example of a country that is already quite well-advanced in using and applying Big Data processes. Among other things, it operates a prototype of the 'Border Risk Identification System' (BRIS). This system can be used at international airports to better estimate which travelers might cause problems. Reference can also be made to the 'Developmental Pathways Project', in which data on children from a variety of sources are linked. Among other things, an assessment will be made of the influence of factors relating to family and the environment on the health of children, the risk of juvenile delinquency, and education. Finally, there is a data tool, Vizie, which has been designed by the Commonwealth Scientific and Industrial Research Organisation (CSIRO), an Australian government corporate entity. This tool follows activity on social media and analyses social media behaviour. A number of government agencies and public sector actors would also like to use this tool, at least according to CSIRO. Some examples can also be found of trials with Big Data in the area of security in the United States. For example, police forces used Big Data analytics to predict the odds that an individual will become involved in criminal activity. An example is Philadelphia, where the police used a tool to predict the chance of repeated offences. In addition, as indicated in the previous paragraph, countries such as Brazil, Israel and the United Kingdom promote the use of Big Data by the police, the intelligence and security services, and the military.

37 All in all, no clear picture has yet emerged as to where the opportunities for the use of Big Data lie. It seems clear that both the public and private sectors agree that Big Data will be used in the private sector and will lead to economic and jobs growth. There is less certainty about both the desirability and effectiveness of the use of Big Data by the government, particularly for security-related purposes. This also relates to the questions that have already been raised regarding the effectiveness of Big Data-type data collections by intelligence services such as the NSA in the United States in the fight against terrorism. Yet, a number of countries have actually implemented such projects involving the intelligence services, the armed forces and the police; for example, in connection with predictive policing. In conclusion, it seems advisable that regulators make an explicit assessment of the desirability and effectiveness of the use of Big Data in the public sector, especially when used for the promotion of national security or public order.

G. What are the dangers of Big Data?

38 This study shows that the dangers of Big Data are mainly assessed along two lines: first, a possible violation of the right to privacy and/or the right to data protection, and second, the danger of discrimination and stigmatization. With regards to the first point, most countries appear to be well aware of the risks that Big Data might pose for the privacy of citizens. For example, the current legal framework is based on the principles of purpose and purpose limitation. Article 7 of the EU Data Protection Directive contains an exhaustive list of the legitimate grounds for processing ordinary personal data; Article 8 does the same with regard to the processing of sensitive personal data (e.g. about race, religion, sexual orientation, etc.). Article 6 states that personal data must be processed fairly and lawfully, and must be collected for specified, explicit and legitimate purposes, and not further processed in a way that is incompatible with those purposes. The prohibition on further processing for different purposes is also known as the 'purpose limitation principle', from which it follows that 'secondary use' is in principle not permitted. The results of both the desk research and the survey show that it is this principle (along with the data minimization principle) that is cited the most when it comes to the tension between Big Data and data protection. Big Data processes often have no fixed purpose – large amounts of data are simply collected and it may only become clear what the value or potential use of that data is after it has been collected. Moreover, in Big Data analysis, different kinds of databases with different types of data are often linked or merged. The original purpose for which the data was collected is then lost. For example, the Swedish DPA argues that the concept of Big Data 'is used for situations where large amounts of data are gathered in order to be made available for different purposes, not always precisely determined in advance.'

39 The second principle that is often mentioned is the principle of data minimization. This principle requires that as little data as possible should be collected, and that the amount of data should, in any event, not be excessive in relation to the purposes for which it is collected. Additionally, personal data must be removed once the goal for which they were gathered has been achieved, and data should be rendered anonymous when possible. This principle, which mainly follows from Article 6 of the Data Protection Directive, obviously clashes with Big Data. The core idea behind Big Data is that as much data as possible is collected and that new purposes can always be found for data already gathered. Data can always be given a second life. This also challenges the requirement that data should be deleted or anonymized when it is no longer needed for achieving the purpose for which it was collected.

Almost all DPAs mention this principle when it comes to the dangers of Big Data. The Luxembourg DPA, among others, refers to a decision in which it stressed the importance of a retention period for data storage. The Dutch DPA summarizes the tension between Big Data and data minimization in very clear terms: 'Big Data is all about collecting as much information as possible.'

40 Articles 16 and 17 of the Data Protection Directive espouse the principle that data should be treated confidentially and should be stored in a secure manner. Many DPAs also mention this principle when discussing the dangers of Big Data; this holds especially for countries and DPAs that establish a link between Big Data and Open Data. The Slovenian DPA, for example, argues that the 'principles of personal data accuracy and personal data being kept up to date may also be under pressure in Big Data processing. Data may be processed by several entities and merged from different sources without proper transparency and legal ground. Processing vast quantities of personal data also brings along higher data security concerns and calls for strict and effective technical and organizational data security measures.'

41 The current framework also requires that the data that is accurate and kept up-to-date. This ensures that profiles created of or applied to an individual person, and any decisions taken on the basis of them, are appropriate and accurate. This study shows that many countries are aware of this tension and that DPAs are concerned about how this principle can be maintained in Big Data processes. Often, Big Data applications do not revolve around individual profiles, but around group profiles; not around retrospective analyses, but around probability and predictive applications with a certain margin of error. Moreover, it is supposedly becoming less and less important for data processors to work with correct and accurate data about specific individuals, as long as a high percentage of the data on which the analysis is based provides a generally correct picture. 'Quantity over quality of data', so the saying goes, as more and more organizations become accustomed to working with 'dirty data'. In the public sector, too, it seems that working with contaminated data or unreliable sources is becoming more common. Examples include the use by government agencies of open sources on the Internet, such as Facebook, websites and discussion forums. The Dutch DPA, for example, refers to the fact that in Holland, there 'has been a lot of media attention for Big Data use by the Tax administration scraping websites such as Marktplaats [an eBay-like website] to detect sales, mass collection of data about parking and driving in leased cars, including use of ANPR data, and profiling people to detect potentially fraudulent tax filings.'

42 An important principle of the Data Protection Directive and the upcoming General Data Protection Regulation is transparency. It includes a right of the data subject to request information about whether data relating to him/her are processed, how and by whom; the controller has a duty to provide the data subject with this information on its own initiative. This principle is also at odds with the rise of Big Data, partly because data subjects often simply do not know that their data is being collected and are therefore not likely to invoke their right to information. This applies equally to the flipside of the coin: the transparency obligation for data controllers. For them, it is often unclear to whom the information relates, where the information came from and how they could contact the data subjects, especially when the processes entail the linking of different databases and the re-use of information. As the Slovenian DPA puts it: 'Big Data has important information privacy implications. Information on personal data processing may not be known to the individual or poorly described for the individual, personal data may be used for purposes previously unknown to the individual. The individual may be profiled and decisions may be adopted in automated and non-transparent fashion having more or less severe consequences for the individual.'

43 The current legal system also puts much emphasis on subjective individual rights and does so to an increasing degree. For example, the forthcoming Regulation gives data subjects additional individual rights, such as the right to be forgotten and the right to data portability. In their response to the survey, DPAs also frequently referred to the principle of informed consent. Individual rights traditionally also come with individual responsibility, namely to protect individual rights and to invoke them if they are undermined. The question is whether this focus can be maintained in the age of Big Data. It is often difficult for individuals to demonstrate personal injury or an individual interest in a case; individuals are often unaware that their rights are being violated, even if they do know that their data has been gathered. In the Big Data era, data collection will presumably be so widespread that it is impossible for individuals to assess each data process to determine whether it includes their personal data; if so, to determine whether or not the processing is lawful; and, if that is not the case, to go to court or file a complaint. This tension appears both from the desk research and from the output of the survey. The British DPA holds, for example, that it 'may be difficult to provide meaningful privacy information to data subjects, because of the complexity of the analytics and people's reluctance to read terms and conditions, and because it may not be possible to identify at the outset all the purposes for which the data will be used. It may be difficult to obtain valid consent, particularly in circumstances where

data is being collected through being observed or gathered from connected devices, rather than being consciously provided by data subjects.'

44 Finally, the current system is primarily based on the legal regulation of rights and obligations. Big Data challenges this basis in several ways. Data processing is becoming increasingly transnational. This implies that more and more agreements must be made between jurisdictions and states. Making this legally binding is often difficult due to the different traditions and legal systems. Rapidly changing technology means that specific legal provisions can easily be circumvented and that unforeseen problems and challenges arise. The legal reality is often overtaken by events and technical developments. The fact that many of the problems resulting from Big Data processes, as also highlighted by a number of DPAs, predominantly revolve about more general social and societal issues makes it difficult to address all the Big Data issues within specific legal doctrines, which are often aimed at protecting the interests of individuals, of legal subjects. That is why more and more national governments are looking for alternatives or additions to traditional black letter law when regulating Big Data – for example, self-regulation, codes of conduct and ethical guidelines. The DPA of the United Kingdom states, for example, that it is notable 'that there is some evidence of a move towards self-regulation, in the sense that some companies are developing what can be described as an 'ethical' approach to Big Data, based on understanding the customer's perspective, being transparent about the processing and building trust.'

45 Besides privacy and data protection principles, DPAs also place a good deal of emphasis on profiling and the risk of discrimination, stigmatization and inequality of power resulting from Big Data. The desk research shows that a number of countries specifically acknowledge this danger. The best overview of these types of dangers is provided in the Working Paper 'Big Data and Privacy: Privacy principles under pressure in the age of Big Data analytics' by the International Working Group on Data Protection in Telecommunications. Four points are made in the working paper in this respect. First, there is a risk of power imbalance between those that gather the data (multinationals and states) and citizens. Second, there is a risk of determinism and discrimination, because algorithms are not neutral, but reflect choices, among others, about data, connections, inferences, interpretations, and thresholds for inclusion that advances a specific purpose. Big Data may, the Working Group makes clear, consolidate existing prejudices and stereotyping, as well as reinforce social exclusion and stratification. Third, there is the risk of chilling effects, which is the effect that people will restrict and limit their behavior if they know or think that they might be surveilled. Fourth and finally, the Working groups signal the chance of echo chambers, which may result from personalized advertising, search results and news items. 'The danger associated with so-called 'echo chambers' or 'filter bubbles' is that the population will only be exposed to content which confirms their own attitudes and values. The exchange of ideas and viewpoints may be curbed when individuals are more rarely exposed to viewpoints different from their own.'[14]

46 It, therefore, appears that in addition to opportunities, there are significant risks associated with Big Data processes. It should be emphasized that these threats again vary with respect to their impact on citizens according to their application. Instances of discrimination are always problematic, but if the police discriminates, this may obviously be more serious than in the case of personalized advertisements. Consequently, when regulating Big Data, account should be taken of the likelihood and the magnitude of potential problems relating to privacy and/or discrimination, and this must be weighed against the potential benefits.

H. Are the current laws and regulations applicable to Big Data?

47 Both the desk research and the results of the survey show that in most countries, the current rules in the area of privacy and data protection, as developed in their respective jurisdictions, are applied to Big Data processes. There is Germany with its distinctive personality right, the United States without an umbrella law for the regulation of privacy, but with sectoral legislation, and most other countries with relatively similar rules concerning privacy and data protection. In addition, a number of countries have specific laws on telecommunications and special rules for organizations such as the intelligence services and archives. In Australia, for example, there is specific regulation covering data matching in terms of tax records by governmental agencies, in which protocols are established for linking this data. Government departments working with files from the tax department must fulfill the requirements of the 'Data-matching Program (Assistance and Tax) Act 1990'. There are also mandatory guidelines for the implementation of the data-matching program.

48 It appears that current legislation is generally applied to Big Data projects, including in several court cases. In July 2015, for example, the French Constitutional Court, the Conseil Constitutionnel, gave its opinion on the French law governing the

14 <www.datenschutz-berlin.de/attachments/1052/WP_Big_Data_final_clean_675.48.12.pdf>.

intelligence and security services. In this ruling, the court specifically stated which provisions of this law are in line with the French Constitution and which parts or provisions of the law are not. Some provisions were declared unconstitutional, including a provision regarding the permission given by the Minister to monitor communications sent from abroad or received from abroad. In the United States, the case of the United States v Jones from 2011 may be of importance because this lawsuit had a limiting effect on the large-scale data gathering of location data by the police. In ACLU v Clapper, the Second Circuit Court of Appeals ruled that the mass collection of metadata about phone records by the NSA is illegal – this activity is not covered by section 215 of the Patriot Act. Meanwhile, however, the Foreign Intelligence Surveillance Court has ruled that the collection of metadata may continue. In the United Kingdom, in the case of Google Inc. v Vidal-Hall & Others, the Court of Appeal was asked to rule on the interpretation of the Data Protection Act 1998. The case revolved around the complaint by users of Apple's Safari browser, who believed that Google was gathering data through that browser in violation of the Data Protection Act 1998. The Court ruled that browsing information may be personal information and abuse of personal information should be considered as a tort.

49 From the survey among the DPAs, it also appears that current legislation is considered to be generally applicable to Big Data. They mostly refer to the national implementation of the Data Protection Directive. Yet, there are a number of countries with specific laws. Because the Estonian DPA sees Big Data as part of the Open Data movement, it refers to the Open Data legislation, namely the Public Information Act, which is currently pending in Parliament. In Hungary, the Information Self-Determination and Freedom of Information ('Privacy Act') applies. The Swedish DPA refers to special legislation for public services, such as the tax authorities, and to telecommunications law which partially constitutes an implementation of the European e-Privacy Directive. The survey also shows that the current legislation is applied in legal cases by national courts and in the opinions of the DPAs. The Belgian DPA refers to its advice on profiling, the DPA of Luxembourg to a report on smart metering and the Dutch DPA to lawsuits regarding the Tax Authorities and the use of data collected by the police through traffic cameras operated by the Tax Authorities.

50 In conclusion, it seems that the current legislation is generally declared to be applicable to Big Data; both courts and DPAs have successfully applied current principles when assessing Big Data-related projects. This should be taken into account when regulating Big Data. Replacing the current regulation with new 'Big Data' regulation would be to throw the baby out with the bathwater. If additional regulation is required, it seems more logical to develop new rules that could be applied in addition to the current regulatory framework. Whether, and to what extent, there is a need for such additional legislation will be discussed next.

I. Is there a need for new legislation for Big Data?

51 It is evident from the foregoing sections that in most countries, Big Data initiatives are treated under existing legislation with regard to issues such as privacy and data protection. Furthermore, the DPAs are agreed that the current data protection principles must be maintained. The Slovenian DPA, for example, explicitly points out that Big Data brings substantial challenges 'for personal data protection and these challenges must firstly be well understood and adequately addressed. In our view, new concepts and paradigms, such as cloud computing or Big Data should not lower or undermine the current levels of data protection as a fundamental human right. Existing central data protection principles, such lawfulness, fairness, proportionality, rights of the data subjects and finality should not be undermined with the advent of Big Data. The rights of the individuals to informational self-determination should be cornerstone in modern information society, protected by modern data protection framework delivering efficient data protection for the individual, while allowing lawful and legitimate interests, often also in the interest of the individual, to be attained.' Yet, most DPAs are also aware of the fundamental clash between Big Data and data protection principles, as discussed previously.

52 It is remarkable from the survey it appears that despite this fact, as of yet, little new legislation seems to be being developed that specifically addresses the new dangers posed by Big Data. Some DPAs refer to the forthcoming General Data Protection Regulation and indicate that they hope that those rules will help them to adequately curb the dangers of Big Data. For example, the British DPA suggests 'that the proposals for the new EU General Data Protection regulation incorporate some of the measures we have identified as being important in ensuring compliance in Big Data e.g. clearer privacy notices, privacy impact assessments and privacy by design. We welcome the fact that these measures are being foregrounded, although we are concerned that that they should not be seen as simply a bureaucratic exercise.' Moreover, the Estonian parliament is discussing new legislation on Open Data (including Big Data). Also, a number of DPAs refer to co-regulation and self-regulation as a possible solution.

53 Yet, the desk research supports the idea that governments are, in fact, actively thinking about new legislation, partly because current laws are seen as hindering technological innovation. Japan may be a case in point here. In 2013, the Strategic Headquarters for IT produced an amendment to various statutory provisions on privacy and data protection: 'Directions on Institutional Revision for Protection and Utilization of Personal Data'. A summary containing the main points of its policy, issued in 2014, discusses technological developments, including Big Data, that have occurred since the introduction of the Data Protection Act of 2003. According to the Strategic Headquarters for IT, there are now several barriers to the use of personal data. Furthermore, even organizations that respect the law and do not infringe rights are worried about criticism over potential privacy violations and the use of personal data; as a consequence, data are not used optimally. The growth envisaged by the Japanese government can only be achieved if personal data is used optimally and if Big Data flourishes. That is why the government wants to remove these barriers. An environment must be created in which violations of rights are prevented and in which personal information and privacy are protected, but in which, at the same time, personal information can be used for innovation. Furthermore, the UK Parliament has commissioned a study on the legislative framework for sharing data between public authorities. In July 2014, a commission published a report with three recommendations, suggesting among other things that the legal reform should go beyond simply stipulating rules for the sharing of data between public authorities; it should also regard the sharing of information between government agencies and organizations with public tasks. Finally, reference can be made to Germany. The Minister of the Interior has proposed a new principle for forthcoming legislation: the minimization of risk. He has also announced that Germany will propose the inclusion of provisions about pseudonymisation and profiling.

54 Consequently, when answering the question of whether it is desirable to formulate new rules for Big Data processes, three specific issues seem important. First, almost all countries and DPAs acknowledge that Big Data poses new and fairly fundamental risks to the current regulatory framework, and in particular the underlying principles. Second, the current regulatory framework is perceived as being (too) restrictive in relation to the deployment of new technologies and technological innovation, particularly in the private sector. Thirdly, many stakeholders are unsure how the current regulatory framework should actually be applied and interpreted in relation to Big Data. Two dangers might follow from this: on the one hand stakeholders, for fear of breaking the law, might forgo many technological innovations and data uses that would in fact be legitimate. On the other hand, parties might use – or rather, abuse – the existing grey area to deploy certain technologies that would not be in accordance with the current regulatory framework. Whether and how a new regulatory framework might provide a solution for these challenges needs to be assessed carefully by regulators.

J. What concept should be central to Big Data regulation?

55 In short, a diffuse picture emerges, with respect to the extent to which developing a special regulatory Big Data regime is necessary or even desirable. What is evident is that regulating Big Data will be especially difficult for two reasons. First, it is difficult to choose a good starting point for the regulation of Big Data; this will be discussed in this section. Second, it will be difficult to pinpoint a specific person or institution to serve as data controller or, more generally, a natural or legal person that is responsible for compliance with the regulatory principles in Big Data processes. This will be discussed in the next section. Regarding the starting point, it should be noted that the current regulation is primarily based on the individual and their interests – this holds for human rights such as privacy and for data protection, which is based on the concept of 'personal data', i.e. data that enables someone to identify or individualize a natural person. However, Big Data processes do not so much revolve around the storage and processing of data at an individual level – rather, the trend is to work increasingly with aggregated data, general patterns and group profiles. Consequently, it is questionable whether the focus on the individual, on personal data, can still be maintained in the Big Data era. The statistical correlations and group profiles do not qualify personal data, but can be used inter alia to alter, shape or influence the living environment of people to a great extent. Furthermore, the trend towards the use of metadata also ties into this problem, because it is unclear to what extent metadata will always qualify as personal data.

56 In addition, many DPAs point out that in Big Data processes, personal data or profiles may be created through the use, combination or analysis of data that do not qualify as personal data. The EPDS states explicitly that a lot of data is gathered in Big Data processes, but also suggests: 'Not all of these data are personal, but many players in the digital economy increasingly rely on the large scale collection of and trade in personal information.' The Working Party 29 states that: 'In addition, Big Data processing operations do not always involve personal data. Nevertheless, the retention and analysis of huge amounts of personal data in Big Data environments require particular attention and care. Patterns

relating to specific individuals may be identified, also by means of the increased availability of computer processing power and data mining capabilities.' The DPA from Luxembourg suggests that Big Data 'allows for the correlation of information which previously could not be linked. From a data protection point of view it can raise many concerns, when it contains personal data, such as the respect of data subjects' rights – for example in the context of data mining – and their ability to exercise control over the personal data or the respect fundamental principles of data protection such as that of data minimization or purpose limitation. Moreover practices such as linking separate databases or computer analytics can turn anonymous data or any kind of non-identifiable information into personal data which would need to be protected under data protection law.' As a final example, reference can be made to the DPA from Slovakia, which argues: 'As a research topic, we would like to suggest examining boundaries between personal and non-personal information. In the Big Data environment, you are able to connect non-personal information and, based on this information, identify the data subject which represents potential risk to rights of the data subjects.'

57 Consequently, it is questionable whether the individual, individual interests and concepts such as personal data, which are explicitly linked to individual natural persons, still serve as a good starting point for building a regulatory framework in the Big Data era. Irrespective of whether the regulator chooses to leave the current legislation largely intact, whether it opts to amend current legislation or chooses to develop a new Big Data framework, it seems that at a certain point in time it will be necessary to address the fact that it is increasingly difficult to take 'personal data', or a related concept, as the basis for rules and obligations. It should finally be noted that the nature of the data is also becoming less and less static; rather, data increasingly goes through a lifecycle in which its nature might change constantly. While the current legal system is focused on relatively static stages of data, and linked to them specific forms of protection (e.g. for personal data, sensitive data, private data, statistical data, anonymous data, non-identifying information, metadata, etc.), in reality, data go through a circular process: data is linked, aggregated and anonymized and then again de-anonymized, enriched with other data and profiles, so that it becomes personally identifying information again, and potentially even sensitive data, and is then once again pseudonymised, used for statistical analysis and group profiles, etc.

K. How should the responsibilities be distributed?

58 A final question that needs to be answered when regulating Big Data is who should bear responsibility for enforcing the rights and obligations; or, in data protection terms, who should be the data controller. This issue exists irrespective of whether the regulator chooses to leave the existing legislation untouched, seeks to amend current legislation or opts to develop new Big Data legislation. The problem of allocating responsibility was prominent both in the desk research and the survey and, in general, manifests itself on three different levels. Firstly, there was already a fair degree of awareness of the increasingly transnational nature of data processing activities. The problem is that different countries have different levels of data protection. The danger is that private parties will settle in those countries where the regulatory pressure is low. But public sector organisations might act in similar ways as well. For example, in the Netherlands, there is a court case pending on the cooperation between the Dutch intelligence services and their counterparts abroad. Although the Netherlands limits the capacities of its intelligence services to collecting information about Dutch citizens, the US intelligence services, which are less constrained regarding the collection of data on Dutch nationals, might collect such data and then pass it on to the Dutch intelligence services. This might work the other way around, too. Consequently, intelligence services might effectively circumvent the rules that apply to them, by cooperating with other international actors that are not bound by those rules.

59 Secondly, it is also apparent from the desk research that there is increasing cooperation between the public and the private sectors, voluntary or otherwise. For example, in Australia, there is collaboration between industry and academia; the Brazilian police use a system that was originally developed by Microsoft and the New York police; China stresses the need for cooperation between the public and the private sector; and the Estonian DPA refers to the cooperation between public and private parties with respect to the development of regional policies. Again, the question is which responsibilities should be borne by which party. Often, it is not clear at first sight what role an organization has played in the value chain of the data processing activity. Also, very different regulatory frameworks often apply to public sector and private sector institutions, as also noted by a number of DPAs in their response to the survey.

60 Thirdly and finally, there is also a trend towards sharing data and linking databases between governmental organisations. This implies that

governmental agencies that have a limited legal capacity to gather and store data may still obtain a wealth of information from other governmental organisations that have a greater legal capacity to gather and store such data. For example, the Dutch DPA refers to a lawsuit that revolves around the use by the Tax Authorities of information gathered by the police. Again, the question is which party should bear responsibility for enforcing the legal regime and the restrictions it imposes. More generally, it should be noted that data flows are becoming more fluid and elusive, meaning that more and more organizations are involved and more and more parties share partial responsibility. This complicates the attribution of responsibilities.

61 Just as the lifecycle of data is becoming increasingly circular, so the division of responsibilities is a clearly shifting from a rather static reality, in which one party collects and processes data, is the main controller of the data and should therefore enforce the different rules and obligations encapsulated in the legislative framework, to a world in which different parties collect, share and link data; in which parties from the private and the public sectors cooperate; in which different governmental institutions share data and databases; and in which international data flows are becoming increasingly common. Consequently, when regulating Big Data, it seems logical to make a choice regarding the distribution and attribution of responsibility. The regulator may, despite these developments, opt for a relatively static model in which one party is the main controller and is responsible for enforcing the legal obligations; or it could opt for a more dynamic model, in which the distribution and attribution of responsibilities is shared and might change as the nature of the data processing activities change. The Data Protection Directive could provide a basis for the latter option, as it defines the controller as 'the natural or legal person, public authority, agency or any other body which alone or jointly with others determines the purposes and means of the processing of personal data.'

L. Summary of main findings

1. *What is the definition of Big Data?* It is impossible to give an exact definition of Big Data. From the research conducted for this report, it follows that a number of different phases must be taken into account when defining Big Data, namely the collection, analysis and use of data. Big Data revolves around collecting large amounts of data (volume), from varied (variety) and often unstructured data sources. Big Data refers to the speed (velocity) of the analyses, often with the use algorithms, machine learning and statistical correlations. The results are often predictive in nature (predictive analytics) and are formulated on a general or group level. The use of the results is usually carried out through profiling. Many of the definitions used in the field contain some of these concepts; none of them mentions all of them. It therefore seems premature to give an exact and precise definition. Two things must be taken into account when regulating Big Data. First, the fact that Big Data cannot be easily defined; this will complicate the making of specific Big Data regulations or laws. Second, the fact that the Big Data process occurs at three levels: collection, analysis and use. These are communicating vessels and must be treated and possibly regulated in connection to each other.

2. *Is Big Data an independent phenomenon?* Big Data should be viewed in its interrelationship and in conjunction with other phenomena. Big Data is part of and in some sense the umbrella term for many of the technological developments that are taking place right now. Terms that are often mentioned as part of the definition of Big Data or as related to Big Data are: Open Data, Re-Use, Internet of Things, smart applications, Profiling, Algorithms and Cloud Computing. Also, machine learning, commodification, datafication, securitization and risk society are sometimes brought up. If the government chooses to regulate Big Data, it should take into account that Big Data is not an isolated phenomenon, but is a development which by its nature very strongly correlates with a number of technical, social and legal developments that are already taking place. The government will have to take a holistic approach when regulating Big Data and related phenomena.

3. *Big Data: fact or fiction?* Right now, Big Data plays a small role, but it will, nevertheless, become increasingly important as time progresses. Consequently, Big Data should not be seen as either an actual practice or fiction, a hype that will blow over, but mainly as a trend that will play a major role of significance in 5 or 10 years from now and will have a significant impact on the operations of governments and businesses and will significantly affect the everyday life of citizens. Only then will many of the effects of Big Data become clear. The government should develop future-oriented policies that follow and preferably anticipate this trend. If it starts to regulate Big Data only in about 5 or 10 years, many of the projects will already have started. The potential negative consequences will have materialized, and it will be difficult to adjust or cancel the projects that have already started. It should also be remembered that good and clear regulation can contribute to innovation

and the use of Big Data. Because the frameworks for Big Data projects are not always clear at the moment, some government agencies and companies are reluctant to use new technologies for fear of breaking the law. New regulation may give more clarity on this point.

4. *What is the scope of Big Data?* Generally speaking, the use of Big Data can be divided into three types. First, the use of Big Data for specific government tasks - examples include the use of Big Data by intelligence services, the police, tax authorities and other public bodies; for example, in the context of formulating economic policies. Second, the use of Big Data by the private or semi-public sector for achieving their tasks and/or goals. Examples include the use of Big Data by companies to create risk profiles, the use of Big Data in the healthcare sector and the use of Big Data in scientific projects. Thirdly, Big Data is used by both governments and companies to improve their service to citizens and customers - for example, this could involve increasing the transparency of activities, strengthening the control citizens have on data processing, etc. The regulation of Big Data will have to take into account the impact the use of Big Data has on the individual, the type of data and data analysis that is used, and the possible danger of a mismatch between a general profile and a specific individual. A distinction must be made between the type of body that executes Big Data projects and the specific purpose for which it is used - the general interest served by the use of Big Data should also have an impact on what is legally permissible.

5. *What are the opportunities for Big Data?* The first opportunity that Big Data offers is to improve the service to the citizen or customer, to improve transparency in the public or private sector, and to give more control to individuals. This practice is generally unproblematic as it serves the interests of the citizen. The second possibility is the use of Big Data in the private and semi-public sector. Big Data is expected to provide a substantial growth in the number of companies, especially start-ups, the number of jobs and the profits generated by these companies. Both the public and the private sector see the biggest opportunities for Big Data in this field of application. However, the use of Big Data in the private sector is not unproblematic. When advertisements or services are personalized through the use of Big Data, the impact on the individual will be relatively small, but this may be different when risk profiles are created by banks or health insurers when deciding who may get a loan or insurance, and on what condition. There exists controversy about the question whether governments should make use of Big Data, especially with respect to security-related purposes. On the one hand, some countries already use Big Data, also for security-related purposes. On the other hand, there are considerable doubts about both the efficacy and the desirability of these projects. The regulator should particularly assess the efficacy and the desirability of the use of Big Data by the public sector institutions when used for security-related purposes. With regard to the use of Big Data by the private sector, a distinction should be made between the type of application.

6. *What are the dangers of Big Data?* This study shows that the dangers of Big Data are assessed mainly along two lines. First, a possible violation of the right to privacy or the right to data protection. Second, the danger of discrimination and stigmatization. Regarding the first point, it appears from underlying research that most countries are well aware of the risks to the privacy of citizens. With regard to the risk of discrimination and stigmatization, this appears to be true to a lesser extent. Consequently, the government will have to weigh the dangers of a breach of privacy and of discrimination against the potential benefits. It should be stressed that both the right to privacy, the right to data protection and the right to freedom from discrimination are fundamental human rights that may be limited only in exceptional circumstances, if necessary in a democratic society.

7. *Are the current laws and regulations applicable to Big Data?* From both the desk research and the results of the survey, it appears that, in most countries, the current regulations in the area of privacy and data protection are applied to Big Data processes. Germany with the distinctive personality right, the United States without an umbrella law for the regulation of privacy, but with sectoral legislation, and most other countries with relatively similar rules concerning privacy and data protection. In addition, a number of countries has specific legislation in the field of telecommunications; also, there are often special rules for organizations such as the intelligence services and archives. Current legislation is generally applicable to Big Data; both courts of law and DPAs are not empty-handed when confronted with Big Data-like processes. This should be taken into account by the government when regulating Big Data. Replacing the current regulation by new 'Big Data' regulation would be to throw the baby out with the bathwater. Rather, it should consider formulating new rules in addition to the current regulatory framework.

8. *Is there a need for new legislation for Big Data?* In most countries, the existing laws are applied to Big Data initiatives. Also, the DPAs are in agreement that the current privacy and data protection principles must be safeguarded. Yet, most DPAs are also aware of the fundamental tension between Big Data and data protection principles. It is remarkable that despite this fact, little new legislation seems to be developed that specifically addresses the new dangers posed by Big Data. Some DPAs refer to the upcoming General Data Protection Regulation and hope it will contain new rules that could help to tackle the dangers posed by Big Data. A number of DPAs refer to co- and self-regulation as a possible solution. Still, some countries seem to be thinking about new regulations for data processing techniques, such as Estonia, France, Japan and Great-Britain. This is partly motivated by concerns over the protection of privacy, but also by the thought that the current laws hinder technological innovation. When answering the question whether it is desirable to formulate new rules for Big Data processes, the government will need to take into account three issues. First, almost all countries and DPAs see new and fundamental risks for the current regulatory framework and, in particular, its underlying principles in the Big Data era. Second, it appears that the current regulatory framework is regarded by some to be too restrictive, muffling the use of new technologies and technological innovation, particularly in the private sector. Third, many parties are unsure how the current rules and laws should be applied to and interpreted in the light of Big Data processes. There are roughly two dangers: on the one hand, for fear of breaking the law, parties may forgo many technological innovations that would be legitimate to use; on the other hand, parties may abuse the existing gray area and take steps that circumvent basic constitutional principles. Whether and how a new regulatory framework can solve these problems needs to be considered by the government.

9. *What concept should be central to Big Data regulation?* Current regulations are often based on the individual and his interests - this applies to individual human rights and to data protection, which regulates the processing of personal data, that is, data that can identify or individualize a natural person. Since increasingly, data are not collected and processed at an individual level, and rather, use is made of aggregated data, which lead to general patterns or group profiles, the question is whether the focus on the individual can still be maintained. This ties up to the use of metadata – it is often unclear to what extent metadata can qualify as personal data. Finally, it should be noted that the nature of the data is less and less static and that data increasingly go through a circular life. While the current legal system is focused on relatively static stages of data and attaches to these stages a specific protection regime (such as for personal data, sensitive data, statistical data, private data, anonymous data, metadata, etc.), in practice, data go through a circular process: data are linked, aggregated and anonymized and then again de-anonymized, enriched with other data for the making of personal or even sensitive profiles, and then again pseudonymised, used for statistical analysis and group profiles, etc. It seems to go too far to simply regulate 'data', but the direct connection to a specific individual, such as is the case with 'personal data', also seems difficult to sustain in the Big Data era. The government will have to determine whether 'personal data' as a concept is still adequate to serve as a basis for data regulation in the Big Data era.

10. *How should the responsibilities be distributed?* Like the life cycle of data that is increasingly circular, with regard to the attribution of responsibilities, a clear shift may be seen from a world in which one controller collects, processes and uses the data and is, therefore, the party solely or primarily responsible for respecting the legal principles, to a world in which data are increasingly shared between governmental organizations, between the private and the public sector and between international public and private sector parties. With regard to the attribution and distribution of responsibilities in the Big Data era, the government has to make a principled choice. Will it, despite the observed trend, maintain the model in which one party has the sole or primary responsibility, and if so, who will bear the burden, or will it choose for a more dynamic model, and if so, how will the responsibility of the parties be divided and established?

Appendix I

	Is a specific definition of Big Data used?	Is Big Data used within the government?	Is there a public-private partnership?
Australia	For the purpose of the Big Data Strategy, the following definition is used: *"1. The data analysis being undertaken uses a high volume of data from a variety of sources including structured, semi-structured, unstructured or even incomplete data; and* *2. The size (volume) of the data sets within the data analysis and velocity with which they need to be analysed has outpaced the current abilities of standard business intelligence tools and methods of analysis".*	The Australian Public Service Big Data Strategy is one of the most prominent examples. This strategy, and accompanying documents, were drafted by the Australian Department of Finance. Parallel to this, a center for the entire the government was set up, headed by the Department of Finance, for improving the data analytics capacity of the government. In the Strategy, several current Big Data projects or pilots of Big Data projects are listed, such as: Border Risk Identification System (BRIS) and the Development Pathways Project.	There is a law that facilitates the use of data from the private sector for the tax authorities, called the Data-matching Program. This law can facilitate a public-private partnership.
Brazil	-	One of the most prominent examples from Brazil is the Big Data tool, 'DataViva', used by the government of the province of Minas Gerais. DataViva combines data from databases belonging to three Ministries and an U.N. database on trade, concerning exports and imports, labour and education, from all over the country. Another prominent example is the system that is used by the Sao Paulo police, 'Detecta'. Detecta is an intelligent system for monitoring crime. Large datasets held by the Sao Paulo police are combined in this tool and subsequently, Detecta makes connections between the data. The system gives of warning signals to relevant authorities and reveals patterns in the crimes committed in the region.	-
China	-	According to the State Council, Big Data is used to make the government more efficient. This entails more personalized service delivery by the government, greater efficiency in the administrative approvals process, with preference being given to companies with a good credit score and those with a poor credit rating being restricted. The premier of the State Council also announced that the government is working on Big Data. An example can be found in the new credit system that will be introduced in China. Another example is the judicial Big Data center, linking all China's judicial bodies.	-

France	-	This is unclear. The French government considered the future challenges for 2025 for the national mail delivery system. The research suggests that the government has five options for resolving these problems. One possible strategy is to focus the service delivery more on e-commerce and to use Big Data analytics to improve the chain of production. It is not yet clear which of these five directions is preferred	-
Germany	The Federal Ministry for Education and Research is the Ministry that is most concerned with Big Data in Germany. According to this Ministry, Big Data is synonymous with: "*den intelligenten Umgang mit solchen großen oder auch heterogenen Datenmengen*" (intelligent use of large or heterogeneous datasets).	This is unclear. There are investments and research projects concerning Big Data. In 2014, the Ministry announced that it would be providing financial support for the construction of two Big Data centers: the Berlin Big Data Centre and the Competence Center for Scalable Data Services in Dresden. In addition to building the two centers, the Ministry will promote further research in support of Big Data, as illustrated by the funding initiative launched in 2013. Specifically, the Ministry will focus attention on 'Industry 4.0' projects and on the bio- and geosciences. A research project focusing on Big Data is ABIDA ('Interdisziplinäre Analyse der gesamtgesellschaftlichen und wirtschaftlichen Folgen beim Umgang mit großen Datenmengen'), funded by the Ministry of Education and Research.	-
India	-	The Indian Ministry of Science and Technology has started a Big Data initiative. The Ministry lists four focus areas for the development of a sustainable data analysis system. Aadhaar is a government-wide project being implemented by the Unique Identification Authority of India. It involves the collection of biometric and demographic data of the Indian population. The Indian Government has not specifically labelled this as a Big Data project.	Not in the sense of a partnership, but the Indian government does make datasets publicly available online to make large amounts of non-sensitive data available to society.
Israel	-	C4i is the department of the IDF that is specifically engaged in information and computer technology. An interview with the commander of this unit makes clear that it is no longer just about passing on information to divisions of the armed forces. Rather, C4i should be seen as a tool which can be deployed in the area of Big Data analytics. The IDF makes use of several Big Data systems such a 'Crystal	The Israeli Ministry of Health sent out a tender in August 2015 for a partner in Big Data analytics. The Ministry has an enormous dataset containing all the medical data on the Israeli population as well as data on the health care system. The Ministry wants to put this dataset to good use and to be able to translate it into specific recommendations.

		Ball' and a GPS system to direct the troops.	
Japan	-	The Japan Science and Technology Agency (JST) is the body responsible for implementing the technology policy of the Japanese government. One of JST's research programmes, 'CREST', involves team-based research to achieve the strategic goals of the government. The programme involves research on Big Data, under the auspices of two main projects: 'Advanced Application Technologies to Boost Big Data Utilization for Multiple-Field Scientific Discovery and Social Problem Solving' and 'Advanced Core Technologies for Big Data Integration'.	There is not a specific partnership, but the sharing of data between the two sectors is encouraged by the government, especially data relating to earthquakes.
South-Africa	-	With the Square Kilometre Array (SKA), a large multi-radio telescope project, South Africa is seeking to put itself on the map as a Big Data hub. The data science capacity that comes with the SKA project must be provided by a network of universities, grouped together in the 'Inter-University Institute for Data-Intensive Astronomy (IDIA)'.	-
United Kingdom	-	The British government published a strategy for Big Data: 'Seizing the data opportunity. A strategy for uk data capability', and made several large investments in Big Data Research Councils. One of the projects funded by a Council is the 'Big Data for Law' initiative, allowing Big Data research on legislation. There are several Big Data projects scattered over various sectors, these projects are described in 'POSTnotes' by the Parliamentary Office of Science and Technology.	The government has founded several Big Data centers which are used by the private sector, in which data from the government sector and private sector is used, or in which researchers and the business sector work together. The British government also makes use of said data.
Unites States	The Podesta report refers to the definition given by Gartner and adds that: "More precisely, Big Datasets are 'large, diverse, complex, longitudinal, and/or distributed datasets generated from instruments, sensors, Internet transactions, email, video, click streams, and/or all other digital sources available today and in the future".	In March 2012, the Obama Administration launched the 'Big Data Research and Development Initiative'. Under this initiative, six federal government departments and agencies announced the investment of 200 million dollars in additional improvements to the processing of enormous volumes of data. In the fact sheet dated 29 March 2012, 'Big Data Across the Federal Government', dozens of ongoing government projects and partnerships related to Big Data are mapped, in all sectors. Some examples can also be found of trials with Big Data in the area of security in the United States.	The US government appeals to the private sector to "join with the Administration to make the most of the opportunities created by Big Data. Clearly, the government can't do this on its own". Whether this should take the form of a partnership between both sectors remains unclear.

	To what goal is Big Data used by the government?	Which laws are especially relevant for Big Data?	Are there judicial decisions relating to Big Data?
Australia	According to the Australian Public Service Big Data Strategy, the strategy is intended to advance the possibilities of Big Data while safeguarding the privacy of the individual. Improving the possibilities for Big Data analytics for the government should lead to improved services and better policy advice. In this Strategy, the mission of the Australian government in relation to Big Data is described as: "The Australian Government will be a world leader in the use of Big Data analytics to drive efficiency, collaboration and innovation in the public sector".	The Freedom of Information Act 1982, the Archives Act 1983, the Telecommunications Act 1997, the Electronic Transactions Act 1999, the Data-matching Program (Assistance and Tax) Act 1990, the Privacy Act 1988, the Privacy Amendment (Enhancing Privacy Protection) Act 2012, the Privacy Regulation 2013.	-
Brazil	At first, the aim of the DataViva tool was to help in drafting economic policy, but it became clear that it offered opportunities as a Big Data tool as such; the relationships and dynamics that the tool exposes provide an insight into the economy for public and private actors and support them in their decision-making. The Detecta system is used to combat and prevent crime.	An amendment to the legislation on data protection is currently being developed. The government has released a draft bill for this law, entitled: "On the processing of personal data to protect the personality and dignity of natural persons".	-
China	There is an emphasis on the use of Big Data to make government services more efficient and to stimulate economic growth.	China does not have overarching privacy legislation such as is present in many European countries. At the end of 2012, the Chinese parliament drafted a resolution consisting of 12 articles and regulating privacy and data protection: the 'Decision of the Standing Committee of the National People's Congress to Strengthen the Protection of Internet Data'.	-
France	Big Data is highlighted by the French government as one of the key developments for modern reforms in French industry.	The 'Loi Informatique et libertés'1978, which has been amended several times since its introduction.	The highest French constitutional court, the Conseil Constitutionnel, issued a ruling in July 2015 regarding the French law governing the intelligence and security agencies. In this ruling, the court declared specifically which provisions of this law are in accordance with the French Constitution and which provisions or parts of provisions are not. What is for example permitted, subject to certain conditions, is the collection of data in real time in order to prevent terrorism, and obliging service providers to identify connections (the parameters of which are set out in the order) which suggest a terrorist threat.

Germany	Several research initiatives for Big Data are aimed at researching how Big Data can be used sensibly and how to handle Big Data. Big Data is seen as a great opportunity for the ICT sector and can improve the competitive position of the German business and science sector, but is also seen as "one of the major challenges" of our time.	The central data protection legislation in Germany is the Bundesdatenschutzgesetz, originally dating from 1990.	-
India	The Indian government uses its Big Data strategy to focus on a sustainable system of data analysis.	The "Information Technology (Reasonable security practices and procedures and sensitive personal data or information) Rules, 2011. A new bill is in the making, the 'Privacy Bill 2014'. In 2012 the Ministry of Science and Technology developed a national policy for data sharing and accessibility.	-
Israel	The focus of Big Data initiatives in Israel lies on making the best use possible of the government's data and using Big Data to protect the country and make the military system more efficient.	The right to privacy is enshrined in Section 7 of the Basic Law on Human Dignity and Liberty. In 1981 a law was also introduced which is tailored specifically to this right, the Protection of Privacy Law 5741 – 1981. To implement this this law, special legislation was drafted governing data flows from Israel to other countries. In 2010 an amendment to the privacy legislation was introduced, adding provisions relating to the security of databases.	-
Japan	The Japanese Prime Minister stated that in order to achieve its economic goals the Japanese government was among other things making changes to optimize the it sector. The law on the protection of personal data would be changed to make it easier to use personal information as part of Big Data. The 'it Strategic Headquarters', established within the Japanese Cabinet, published an open data strategy for the government, in which it argued that government data is a public asset and that the sharing and use of that asset should be encouraged.	The Act on the Protection of Personal Information from 2003, in 2013 amendments were made to this law, inter alia because of Big Data.	-
South-Africa	South Africa is seeking to put itself on the map as a Big Data hub, further goals of the Big Data project are to reduce poverty and improve the country's economic competitiveness.	The right to privacy is explicitly enshrined in Article 14 of the South African Constitution. The Protection of Personal Information Act 2013 is relevant.	-
United Kingdom	Big Data is used for various purposes, such as: creating efficient motorways and traffic flows, predicting crime, researching diseases and facilitating Big Data research on legislation. There is no	The Data Protection Act 1998, the Human Rights Act 1998 (section 8), the 2000 Regulation of Investigatory Powers Act and the Intelligence Services Act 1994.	In 2015, the case of Google Inc v Vidal-Hall & Others was heard by the Court of Appeal. The case related to data protection and the Data Protection Act 1998. The Court ruled that browser

	focus on one specific goal. The Minister for Universities and Science and the Minister for Skills and Enterprise state the following about data: Governments around the world must change the way they engage with citizens, the way they develop policy and deliver services, and the way they are held to account (...) The UK government is determined to position the UK to make the most of the data revolution."		information can be regarded personal data and that abuse of personal data should be regarded as a tort. With regard to data protection, the High Court pronounced a verdict in July 2015 in the case of Davis & Others v SSHD in relation to the Data Retention and Investigatory Powers Act 2014. In this case the Court declared this law partially invalid due to conflicts with European law, and specifically the section in which the competence is established to request telecommunications service providers to retain communications data.
Unites States	The Big Data review produced five overarching conclusions which can be seen as goals the government can aim for in following the report: First, more research must be carried out on the protection of privacy, and action should be taken in the area of legislation on the protection of privacy. Second, there should be more attention for the responsible handling of data collected in the context of education, especially data regarding children. Third, the federal government is advised to be on its guard for discrimination of citizens, which can be caused by Big Data analytics. Fourth, the authorities responsible for enforcement and safety are advised to make maximum use of the legal possibilities for Big Data analytics. The Big Data initiatives that are already in place focus on several goals, varying with the sector of the government that they are used within.	The United States does not have an overarching law for the regulation of privacy, and certainly not for the specific regulation of Big Data. Besides the constitutional protection, the United States has a system of sector-specific regulation of privacy risks. The Consumer Bill of Privacy Rights was introduced in 2012. This is not legislation in the sense of being enforceable, but more of a guideline for the business sector.	A court case on limiting the effects on large-scale location data collection by the police was The United States v. Jones from 2011. Another interesting case is Sorrell v. IMS Health Inc., which was also heard by the Supreme Court in 2011. In this case, involving the commercial use of medical data, the Court ruled that there is a limited scope for datamining when in breach of the freedom of expression. On 7 May 2015, the Second Circuit Court of Appeals ruled in ACLU v. Clapper that the large-scale collection of metadata concerning telephone records by the NSA is unlawful. However, the Foreign Intelligence Surveillance Court ruled that the collection of metadata could continue.

Appendix II Responses from the DPAs to the survey

	1. Are you familiar with the debate on Big Data? If so, how would you define Big Data?	**2. Are there prominent examples of the use of Big Data in your country, especially in the law enforcement sector, by the police or by intelligence services?**	**3. Have you issued any decisions/ reports/opinions on the use of Big Data? If so, could you provide us with a reference and your main argument?**
Belgium	We have no official national definition. However we follow closely the definitions; The EDPS states on its website "Big Data means *large amounts of different types of data produced at high speed from multiple sources, whose handling and analysis require new and more powerful processors and algorithms. Not all of these data are personal, but many players in the digital economy increasingly rely on the large scale collection of and trade in personal information. As well as benefits, these growing markets pose specific risks to individual's rights to privacy and to data protection"* (<https://secure.edps.europa.eu/EDPSWEB/edps/Consultation/big_data>) Also, the Working Party 29 has issued a general statement on Big Data. (<http://ec.europa.eu/justice/data-protection/article-29/documentation/opinion-recommendation/files/2014/wp221_en.pdf>) The Consultative Committee of the Convention 108 has appointed an expert that has to write a report on Big Data, expected to become public in 2016 (<www.coe.int/t/dghl/standardsetting/dataprotection/TPD_documents/OJ_TPD32(2015)_11%2006%2015_Fr.asp>)	Not to our knowledge for the indicated sectors in the strict meaning (there is no obligation to notify our DPA of such projects in these sectors). However, in the approach of the fiscal and social fraud, the projects and discussion on the use of Big Data or the steps in this process (profiling, data mining,...) exist since 2012. We have addressed several opinions since 2012 that address a part of the Big Data issue (mainly data mining and profiling)	On profiling by facebook : Aanbeveling 04/2015 van 13 mei 2015 uit eigen beweging met betrekking tot 1) Facebook, 2) de gebruikers van internet en/of Facebook alsook 3) de gebruikers en aanbieders van Facebook diensten, inzonderheid social plug-ins, gepubliceerd op <www.privacycommission.be/sites/privacycommission/files/documents/aanbeveling_04_2015.pdf> At the request of our Commission the inter-university research center EMSOC/SPION (see <www.law.kuleuven.be/icri/en/news/item/icri-cir-advises-belgianprivacy-commission-in-facebook-investigation>) conducted a detailed study into the way in which Facebook deals with its members' personal data. And that of citizens who do not use Facebook or who explicitly opted out of its service. On profiling of energy and water clients: Advies nr. /2015 van 17 juni 2015 betreffende Hoofdstuk II van het Ontwerp van wet houdende diverse bepalingen, betreffende de verbruiksgegevens van nutsbedrijven en distributiebeheerders
Croatia	The Republic of Croatia is familiar with the concept of Big Data, and a definition /explanation with which we most agree is from the text "What is really Big Data and where is it used?" By Luka Stepinac from 12. May 2014. published at the www.ictbusiness.info in which stands „Definition that we can find the most often refers to "3V": Volume - a large amount of data collected, processed and made available for analysis; Velocity - continuous collection of large amounts of data in real time; Variety - the data are available in various forms and sources, and in fact are usually unstructured, or, in one sentence, Big Data is a technology that enables the collection and processing of large amounts of structured and unstructured data in real time."It is	At this moment we do not have an appropriate/adequate information.	No.

	necessary to point out that the Republic of Croatia regularly monitors technological innovations which in most cases allows the use of information from the field of Big Data, and most often in commercial purposes.		
Estonia	Estonian Data Protection Inspectorate is familiar with the debate on Big Data. In our opinion Big Data could be defined as collected and processed open datasets, which are defined by quantity, plurality of data formats and data origination and processing speed.	Yes, some public sector authorities in cooperation with the private sector (e.g. mobile operators) and universities have applied Big Data to their analysis. For example, *Bank of Estonia* (Eesti Pank) and *Statistics Estonia* on tourism statistics, *Ministry of the Interior* with municipalities have used Big Data in the development of regional policy. Based on open datasets, private company *Big Data Scoring* provides background information to loan companies.	No.
France	The CNIL is familiar with the debate on Big Data and is actively working on the subject. In August 2014, a definition of the term 'Big Data' was adopted by the French General Commission on terminology and neology (*Commission générale de terminologie et de néologie*). The official translation of this term in French is '*mégadonnées*' and the definition is 'structured data or not whose very large volume require appropriate analytical tools'. The Gartner definition is also a reference: 'Big Data is high-volume, high-velocity and/or high-variety information assets that demand cost-effective, innovative forms of information processing that enable enhanced insight, decision making and process automation'. With reference to this definition, three 'Vs' are generally associated with Big Data: volume, variety and velocity. Our Data protection authority (DPA), as other actors, considers that other 'Vs' are also relevant, in particular value and veracity. Many examples of Big Data operations involve processing of personal data, in various business sectors. The projects have different goals and use different categories of data. But, beyond this diversity of projects and objectives, the notion of 'Big Data' reveals a new approach of the data, appeared with the development of new storage and analytical capacities. And privacy challenges are associated to Big Data because, thanks to sophisticated algorithms, Big Data can ultimately be used to identify profiles, predict the behavior of individuals or groups of individuals, and take decision affecting them.	There are various examples of the use of Big Data in France, for instance in the fields of marketing, insurance, credit scoring, anti-fraud mechanisms, tourism or research. Data controllers can use specific compliance tools *i.e.* simplified standards or single authorizations that allow interconnecting databases (See AU39 fraud detection in insurance sector for a recent example <www.cnil.fr/documentation/deliberations/deliberation/delib/318/>). Regarding the law enforcement sector, different data processing operations can be considered as Big Data analysis. For example, opinions of the CNIL on such processing operations are available on our website (<www.cnil.fr/nc/linstitution/actualite/article/article/publication-de-lavis-sur-le-projet-de-loi-relatif-aurenseignement/;www.cnil.fr/documentation/deliberations/deliberation/delib/302/>).	At this stage, there is no report on the use of Big Data drafted by our DPA. However, different presentations were made during conferences on this topic as well as analytical articles (see, for example, the article '*Big Data et protection des données personnelles : quels enjeux ?*', Sophie Vulliet-Tavernier, Revue Statistique et société <www.statistique-et-societe.fr>). The CNIL also participated in the elaboration of International opinions (Statement of the WP29 on the impact of the development of Big Data on the protection of individuals with regard to the processing of the personal data in the EU; Working paper on Big Data and Privacy of the International Working Group on Data Protection in Telecommunications, Berlin Group). Besides, in 2011, the CNIL issued a warning against the company *Pages Jaunes* (*deliberation n° 2011-203, September 21, 2011*), for having obtained personal data contained in profiles available on different social media websites, without data subjects' knowing. This online directory proposed a 'webcrawl' function on its website enabling to add information from the accounts of web users to the search results provided by the directory. About 25 million people were concerned and the captured data included the names and first names, pseudonyms, photographs, the names of their school, the names of their employer, their geographical location... In particular, the CNIL considered that the fact that the data were public on the internet did not authorize a third party to massively, repetitively and indiscriminately collect

			such data without informing the data subjects before posting these information on its website. Consequently, the collection of the personal data was unfair. Moreover, it was difficult for the data subjects to exercise their rights. *Pages Jaunes* (Solocal Group) introduced an appeal before the *Conseil d' État* against the warning of the CNIL but the Supreme Court for administrative justice confirmed the analysis of the CNIL (*Conseil d'État, 10ème et 9ème sous-sections réunies, 12/03/2014, 353193*).
Hungary	The Hungarian National Authority for Data Protection and Freedom of Information accepts the Big Data definition of the International Working Group on Data Protection and Telecommunications. According to the Working Group's Working Paper on Big Data and Privacy: "Big Data is a term which refers to the enormous increase in access to and automated use of information. It refers to the gigantic amounts of digital data controlled by companies, authorities and other large organizations which are subjected to extensive analysis based on the use of algorithms." Big Data is, to a certain extent, used to analyze data in order to identify and predict trends and correlations.	As far as we know, there are no prominent examples in Hungary for the use of Big Data in law enforcement sector, by the police or intelligence services.	The Hungarian National Authority for Data Protection and Freedom of Information has not issued any decision, report or opinion on the use of Big Data so far. Besides that our Authority participated in the drafting of the working paper on Big Data by the International Working Group on Data Protection and Telecommunications. It is available online on the following address: <https://datenschutz-berlin.de/content/europa-international/international-working-group-on-data-protection-in-telecommunications-iwgdpt/working-papers-and-common-positions-adopted-by-the-working-group/>
Latvia	We do not have a specifically determined definition for Big Data, even though we are familiar with the debate on it.	No, there aren't.	No, we have not.
Lithuania	The State Data Protection Inspectorate is involved in discussions on Big Data, insofar as regards the performance of supervisory functions.	In Lithuania there is a Home Affairs Information System, which is a system performing data processing in which on the basis of the joint infrastructure of information technology and telecommunications operates the state and institutional registers and information systems (Criminal Offences register, Police information systems and etc.) managed by the MI and institutions under the MI.	Not yet.
Luxembourg	Big Data stems from the collection of large structured or unstructured datasets, the possible merger of such datasets as well as the analysis of these data through computer algorithms. It usually refers to datasets which cannot be stored, managed and analysed with average technical means, due to their size. Personal data can also be a part of Big Data but Big Data usually extends beyond that, containing aggregated and anonymous data. It allows	To our knowledge, there are no prominent examples of the use of Big Data in the law enforcement sector or by police or intelligence services in Luxembourg. There are however other actors which deal with Big Data. At a national level, a system of smart metering for electricity and gas has been launched. The project is, however, still in a testing phase. At the level of the University of Luxembourg, the Luxembourg Centre for Systems	The CNPD has not issued any decisions, reports or opinions that are directly dealing with Big Data. The Commission has however issued an opinion in a related matter, namely with regard to the problematic raised by smart metering. In 2013, the CNPD issued an opinion on smart metering (Avis de la Commission nationale pour la protection des données relatif au projet de règlement grand-ducal relatif aux modalités du comptage de

	for the correlation of information which previously could not be linked. From a data protection point of view it can raise many concerns, when it contains personal data, such as the respect of data subjects' rights – for example, in the context of data mining – and their ability to exercise control over the personal data or the respect fundamental principles of data protection such as that of data minimization or purpose limitation. Moreover, practices such as linking separate databases or computer analytics can turn anonymous data or any kind of non-identifiable information into personal data which would need to be protected under data protection law.	Biomedicine uses Big Data in the health sector. The Interdisciplinary Center for Security, Reliability and Trust (SnT) is also involved in Big Data projects. A partnership with Choice Technologies allows the SnT to conduct research into the new analytical methods in the domain of "Big Data". Moreover there are private companies that use Big Data. NeXus for example is, a company "which surfs the wave of Big Data and security by developing services that fall in the pure concept of "Industry 4.0". "With objects, people and data in constant move, nexus creates a dynamic identity for each end point and keeps track, connects and provides security to the information shared."200	l'énergie électrique et du gaz naturel, Délibération n° 566/2013 du 13 décembre 2013 (<www.cnpd.public.lu/fr/decisions-avis/2013/12/comptage-energie-gaz/566_2013_Deliberation_MinistereEconomie_avis-prj-rgd-comptage-energie-electrique-et-gaz-naturel.pdf>). The main argument of the opinion highlights the necessity to clearly define the purposes of the data processing as well as the retention periods of the data related to smart metering.
Netherlands	Yes, we are familiar with the broad concept of Big Data. Big Data is all about collecting as much information as possible; storing it in ever larger databases; combining data that is collected for different purposes; and applying algorithms to find correlations and unexpected new information. We refer to the speech of our chairman on Big Data, at URL: <https://cbpweb.nl/sites/default/files/atoms/files/2._sp ech_jko_panel_ii_privacy_with_no_territorial_bounds.pdf>	Yes, there are examples of the use of Big Data in the Netherlands. There has been a lot of media attention for Big Data use by the Tax administration (scraping websites such as Marktplaats to detect sales, mass collection of data about parking and driving in leased cars, including use of ANPR-data, and profiling people to detect potentially fraudulent tax filings, see for example the interview with the general manager of the IRS, at <https://decorrespondent.nl/2720/Baas-Belastingdienstover-Big-Data-Mijn-missie-is-gedragsverandering/83656320f6e78aaf>). Next to that, there are many pilots currently being conducted by different municipalities to combine different statistical, social care and medical care data, related to a shift in financial responsibility for social care duties. Recently, an interview was given by high ranking police officers describing the introduction of datamining tools for preventive policing. See URL: <www.politieacademie.nl/kennisenonderzoek/kennis/mediatheek/pdf/89539.pdf>	Next to the speech of our chairman, we refer to international opinions and resolutions from The International Working Group on Data Protection and Telecommunications (<www.datenschutz-berlin.de/attachments/1052/WP_Big_Data_final_clean_675.48.12.pdf?1407931243 The Article 29 Working Party (http://ec.europa.eu/justice/data-protection/article-29/documentation/opinion-recommendation/files/2014/wp221_en.pdf>) and The resolution from the International Commissioners conference (<https://cbpweb.nl/sites/default/files/atoms/files/resolution_big_data.pdf>). Our key concern is that data protection should be about surprise minimisation, while Big Data entails the risk of surprise maximization. There is a real risk that those who are involved in the development and use of Big Data are ignoring the basic principles of purpose limitation, data minimisation and transparency. And an additional frightening fact is that the statistical information, even if the data used is properly anonymised, can lead to such precise results that it essentially constitutes re-identification. When Big Data are used to profile people, it has the potential of leading us on to a - predetermined and maybe sometimes dangerous - path. A path that may in the end undermine the values that underpin our democratic societies, by depriving people of their free choice, of their right to personal development and equal treatment.

Norway	The Norwegian DPA issued a report on Big Data in 2013. The report was very well received and we have been giving talks on this topics for representatives from all sectors, covering finance, health, law enforcement, marketing, telecom etc. In the report we use the definition of Big Data as it was phrased by the the Article 29 Group:201 *Big Data is a term that refers to the enormous increase in access to and automated use of information: It refers to the gigantic amounts of digital data controlled by companies, authorities and other large organisations which are subjected to extensive analysis based on the use of algorithms. Big Data may be used to identify general trends and correlations, but it can also be used such that it affects individuals directly.* We use this definition as a basis, but also add what in our opinion is the key aspect of Big Data, namely that it is about the compilation of data from several different sources. In other words, it is not just the volume in itself that is of interest, but the fact that secondary value is derived from the data through reuse and analysis. This aspect of Big Data, and the consequences it has, is in our opinion the most challenging aspect from a privacy perspective.	There are, as far as we know, no usage of Big Data within the law enforcement sector in Norway. In 2014, the intelligence service addressed in a public speech the need to use Big Data techniques in order to combat terrorism more efficiently. However, politicians across all parties reacted very negatively to this request and no formal request to use such techniques has since been launched by the intelligence service. The companies that are most advanced when it comes to using Big Data may be found within the telecom (eg. Telenor) and media (eg. Schibsted and Cxence) sector. The tax and customs authorities have also initiated projects in which they look at how Big Data can be used to enhance the efficiency of their work.	The Norwegian DPA published a report on Big Data in 2013. In 2014 we drafted a working paper on Big Data for the International Working Group on Data Protection in Telecommunications (aka the Berlin Group). Following on from this work we were later responsible for drafting a Resolution on Big Data for the 36th International Conference of Data Protection Authorities and Privacy Commissioners. Report on Big Data: <www.datatilsynet.no/Global/04_planer_rapporter/big-dataengelsk-web.pdf> Working Paper on Big Data and Privacy: <www.datenschutz-berlin.de/content/europa-international/international working-group-on-data-protection-in-telecommunications-iwgdpt/working-papers-and-common-positions-adopted-bythe-working-group Resolution> on Big Data: <http://privacyconference2014.org/media/16602/Resolution-Big-Data.pdf> Our main argument in the report can be summarized as follows: "Big Data is challenging key privacy principles, in particular the principles of purpose limitation and data minimisation. The protection provided by these privacy principles is more important than ever at a time when an increasing amount of information is collected about us. The principles provide the foundation for safeguards against extensive profiling in an ever increasing array of new contexts. A watering down of key privacy principles, in combination with more extensive use of Big Data, is likely to have adverse consequences for the protection of privacy and other fundamental rights."
Slovakia	We are following the debate, but we have not adopted any definition yet.	We are not aware of special example of the use of Big Data in Slovakia.	No, we have not issued any documents about the use of Big Data yet.
Slovenia	The Information Commissioner is closely following the debate on Big Data. In terms of definitions of Big Data, we believe that established definitions and descriptions (e.g. Wikipedia) adequately describe the issue. Big Data is a broad term for processing of large amounts of different types of data, including personal data, acquired from multiple sources in various formats. Big Data revolves around predictive analytics – acquiring new knowledge from large data sets which requires new and more powerful processing applications.\ Big Data has important information privacy implications. Information on personal data processing may not be known to the	We have thus far not seen prominent examples of the use of Big Data in our country. To our knowledge, Big Data applications are particularly of interest in insurance, banking and electronic communications sector, mostly to battle fraud and other illegal practices. Another important field is scientific and statistical research. Law enforcement use is to our knowledge currently at development stages (e.g. in the case of processing Passenger Name Records), whereas information about the use of Big Data at intelligence services is either not available or of confidential nature.	So far, given that the use of Big Data in our country has not attained greater acceptance, we have not issued particular papers on Big Data at national level. On the other hand, we co-operate in international fora of privacy advocates and supervisory authorities, such as Article 29 Working Party202, International Working Group on Data Protection in Telecommunications203, European and International Privacy Commissioners conference204, which have already provided their views on the issues surrounding Big Data in resolutions, working papers and opinions.

	individual or poorly described for the individual, personal data may be used for purposes previously unknown to the individual. The individual may be profiled and decisions may be adopted in automated and non-transparent fashion having more or less severe consequences for the individual. Decisions about the individual may be biased, discriminatory and even adopted on grounds of statistics, averages and predictions that could have little or even nothing to do with individual's actual data. Such uses could have severe consequences for the individual particular when used by law enforcement, but also in other sensitive fields, such as health services and health insurance, social transfers, employment and in particularly situations where processing of sensitive personal data may be involved. The principles of personal data accuracy and personal data being kept up-to-date may also be under pressure in Big Data processing. Data may be processed by several entities and merged from different sources without proper transparency and legal ground. Processing vast quantities of personal data also brings along higher data security concerns and calls for strict and effective technical and organisational data security measures.		
Sweden	We are familiar with the debate on Big Data, but we have not produced any definition of this concept ourselves. As we see it, the concept is used for situations where large amounts of data are gathered in order to be made available for different purposes, not always precisely determined in advance.	We have not carried out any specific supervision related to the concept Big Data and do not have any statistics or specific information on how this is used. In our opinion, the law enforcement sector does not use Big Data. Their personal data processing is strictly regulated in terms of collection of data, limited purposes etc.	No
United Kingdom	We are familiar with current debates on Big Data and have contributed to them. We consider that the accepted Gartner definition based on the "three V's" (volume, variety and velocity) provides a useful starting point for defining Big Data. We also consider that other key characteristics of Big Data analytics include: repurposing data; using algorithms to find correlations in datasets rather than constructing traditional queries; and bringing together data from a variety of sources, including structured and unstructured data. Furthermore, we note that Big Data may involve not only data that has been consciously provided by data subjects, but also personal data that has been observed (eg from Internet	We have not carried out a comprehensive market assessment of Big Data but, from our contacts with business and our desk research, our impression is that the take up of Big Data is still at a relatively early stage in the UK. Nevertheless, we know that companies are actively investigating the potential of Big Data, and there are some examples of Big Data in practice, such as the use of telematics in motor insurance, the use of mobile phone location data for market research, and the availability of data from the Twitter 'firehose' for analytics. We do not have any specific information on the use of Big Data in law enforcement or security. The UK Data Protection Act includes a wide-ranging exemption from the data	In July 2014, we published a discussion paper on Big Data and data protection. We invited feedback on this and in April 2015, we published a summary of feedback, together with our response. In our work we have noted that Big Data poses a number of challenges to data protection, in particular: It may be difficult to provide meaningful privacy information to data subjects, because of the complexity of the analytics and people's reluctance to read terms and conditions, and because it may not be possible to identify at the outset all the purposes for which the data will be used. It may be difficult to obtain valid consent, particularly in circumstances where data is being collected through being observed or gathered from

	of Things devices), derived from other data or inferred through analytics and profiling. Given the range of features listed here, we think that it is difficult to produce a comprehensive definition of Big Data which fits all use cases. It is better to see Big Data as a phenomenon, rather than a specific technology. In our discussions with companies about Big Data, they have tended to see the defining characteristics of Big Data as the use of new data sources (eg social media data) and the use of existing data for new purposes, rather than simply the volume of data.	protection principles where it is required for safeguarding national security.	connected devices, rather than being consciously provided by data subjects. Big Data tends to use data for new and unexpected purposes, which may conflict with the purpose limitation principle. Big Data tends to use "all the data", which may conflict with the data minimization principle. Nevertheless, we have stressed that the data protection principles still apply in the world of Big Data; it is not a game that is played by different rules. We have said that organisations need to carry out a realistic assessment of what they are trying to achieve, and balance the benefits of the analytics to the organisation, to the individual and to society against the impact on data privacy. They also need to be innovative in seeking new ways to provide privacy notices. We think that privacy impact assessments (PIAs) have an important role to play in helping to ensure that Big Data analytics meets data protection requirements. We are currently doing further work with organisations to explore how PIAs can be used in the context of Big Data as part of privacy by design approach. We also advocate that, wherever possible and appropriate, the data used for the analytics should be anonymised, so that it can no longer be considered to be personal data. We are planning to publish a new version of our Big Data paper later this year.

	4. Are there any legal cases/judgements by a court with regard to (privacy/data protection) violations following from Big Data practices in your country? If so, could you provide us with a reference and the main consideration of the court?	**5. Which legal regimes are applied to Big Data/ is there a special regime for Big Data in your country? Are there any discussions/plans in parliament to introduce new legislation to regulate Big Data practices?**	**6. Are there any final remarks you want to make/suggestions you have for further research?**
Belgium	We have no judgment, yet, in the Facebook case. We expect that the main discussion will be on the competence of our DPA. See the media of 15 june 2015 (<www.theguardian.com/technology/2015/jun/15/belgium-facebook-court-privacy-breaches-ads>).	No. The general data protection law applies, and we expect that the new data protection regulation will be able to provide a partial answer (profiling) to Big Data issues (legal interpretation of the EU legal framework)	Most Belgian projects seem to be still in a pilot phase and the visibility of Big Data in practice is still low (competition issue). Often, the practice is still labeled differently (data mining, profiling,...) Conclusions seem to be premature at this stage until more experience has been obtained on the practical uses of this new practice. (Gartner's 2013 Hype Cycle for Emerging Technologies, <www.gartner.com/newsroom/id/281991>). Follow-up research seems necessary.

Croatia	At this moment, we do not have an appropriate/adequate information.	At the moment, in Republic of Croatia, there is no separate regulations governing the area of the Big Data, but certainly the part referring to the personal data of natural persons applies the Law on Protection of Personal Data.	No.
Estonia	Inspectorate is not aware of legal cases/ judgements by a court, related to Big Data practices in Estonia.	Estonian Data Protection Inspectorate consider Open Data as a part of Big Data. General requirements of Open Data processing are described in the Public Information Act, which new draft bill is in the parliament.	No additional comments.
France	Please refer to the aforementioned case.	Like the WP29, the CNIL considers that the EU and national legal framework for data protection is applicable to the processing of personal data in Big Data operations, even if the challenges of Big Data might require, in some cases, innovative thinking on how some of the key data protection principles are applied in practice. Regarding the discussions at the national level to introduce new legislation to regulate Big Data operations, we can mention the works relating to a new law for a 'Digital Republic' and a report published by the French Digital Council. At present, the French government is preparing a new law for a 'Digital Republic'. An online consultation was launched on the draft bill on September 2015, and the public was invited to suggest amendments to 30 proposed measures, ranging from net neutrality to open data (until 17 October 2015, < http://www.economie.gouv.fr/projet-loi-numerique>). The draft bill proposes, in particular, an open-data policy for the French state that would make official documents and public-sector research accessible to all online. The bill should be submitted to the parliament at the beginning of 2016. The French Digital Council (*Conseil national du numérique, CNNum*) is an independent advisory commission. The Council issues independent opinions and recommendations on any question relating to the impact of digital technologies on economy and society. The government can consult the Council on new legislation or draft regulations. The Council's thirty members come from across the digital spectrum, and include researchers and activists. In its report handed over on 13 June 2014 to Arnaud MONTEBOURG (Minister of Economy, of Productive Recovery and of the Digital) and to Axelle LEMAIRE, (Secretary	-

		of State charged of the Digital), the French Digital Council held an expanded approach to the neutrality principle: consecrate Internet neutrality and take into account the digital platforms that became the new entrance doors of the digital society. The report recommends to establish guidelines on transparency in the way services operate, particularly algorithms. The relevance criteria and governing principles of algorithms should be explained to users as part of a digital literacy effort. The report is available in English on the website of the French Digital Council (<www.cnnumerique.fr/wp-content/uploads/2014/06/PlatformNeutrality_VA.pdf>).	
Hungary	As far as we know, there hasn't been any legal cases or judgments by Hungarian court with regard to violation following from Big Data practices so far.	In Hungary Act CXII of 2011 on Information Self-Determination and Freedom of Information ("Privacy Act") should be applied to any data protection issues including data protection problems concerning Big Data. Neither the aforementioned act nor other laws includes special regulation on Big Data, so the general legal regulation on data protection and privacy should be applied. There aren't any plans or discussions now in the parliament to introduce special legislation for Big Data practices.	We would like to raise to attention that according to the working paper on Big Data by the International Working Group on Data Protection and Telecommunications the application of Privacy-by-Design principles are crucial for legitimate Big Data practices in most cases. Furthermore, a Privacy Impact Assessment could be also recommended and effective before the installation and use of Big Data services in order to avoid future privacy incidents. Furthermore, we would like to point out that in Hungarian business sphere more and more enterprises, such as banks, supermarkets, media and telecommunication companies use and take advantage of the possibilities in Big Data. Moreover, several international conferences are being organized in Budapest in the topic.
Latvia	We do not have such information.	We do not have information on this issue at this point.	No. But we would like to be informed on the outcome of this survey.
Lithuania	Not yet.	Not yet.	-
Luxembourg	No	There is no legislation directly addressing Big Data. The general data protection legislation applies (Amended Act of 2 August 2002 concerning the protection of individuals with regard to the processing of personal data). To our knowledge, there are no plans in Parliament to introduce new legislation to regulate Big Data practices.	-
Netherlands	Yes, there has been a court procedure in two instances about access to parking data for the IRS (case number HD 200.139.173/01, URL: <http://uitspraken.rechtspraak.nl/inziendocument?id=ECLI:NL:GHSHE:2014:2803>). Furthermore,	The current data protection regime also applies to the use of Big Data, but enforcement of the key values cannot be solely made dependent of the supervisory authority. Our chairman has called for a fierce social dialogue, to make people	-

	complaints about the use of police data from traffic cameras for the investigation of road vehicle usage in compliance with tax law have led to complaints and court cases. In March 2015, the Court of Appeal in Den Bosch ruled that the data that is collected with road surveillance camera's of the police that are installed for safety purposes, may be used by the tax authorities to monitor compliance with the law on road vehicle tax. (The ANPR data case, See: <http://uitspraken.rechtspraak.nl/inziendocument?id=ECLI:NL:GHSHE:2015:1087>)	aware of the risks to our intrinsic values that is posed by Big Data and to think together about how we can effectively address these risks and unwanted consequences. With regard to the security and intelligence services, a Bill has been consulted publicly and will be introduced to parliament soon to extend powers to allow for mass interception of communications data. With regard to scientific and academic research, sector- specific rules apply. For example, the law on higher education and scientific research.	
Norway	There are no legal cases	There are no special regimes for Big Data in Norway or plans to introduce new legislation. We rely on the national "Personal Data Act" which builds on the European Data Protection Directive.	Knowledge and awareness of the privacy challenges associated with Big Data are important among the enterprises that implement the technology. We urge the trade organisations to place these challenges on their agendas, and provide training in how they can be handled, for example through the use of privacy by design. Knowledge of data protection and the privacy challenges associated with the use of Big Data should be part of the curriculum for universities and colleges where data analysis or data science are taught. It is also crucial that supervisory authorities possess the necessary knowledge and awareness of the potential that lies in Big Data. This is important so that they can function as efficient and effective enforcers of the regulations that have been established to protect key societal assets. Research on the social and privacy consequences of Big Data is also of great importance. Big Data is still a relatively new phenomenon. It will be important to research how access to ever-increasing volumes and additional types of data will affect how we make decisions and organise our society in the future. At the Norwegian DPA we are currently looking into how it affects our privacy when personal data is more and more turning into a valuable commodity in all sectors of the economy. We are writing a report on how Big Data is used within the advertising industry, and how the use of automated, personalised marketing triggers an enormous appetite for and exchange of personal data.
Slovakia	We have no knowledge about the case or judgements about the Big Data in our country to this date.	We have no special regime for Big Data so far. General data protection law will apply when the personal data will be processed within the Big Data. We are not planning to issue a new legislation connected with	We think that the issue of Big Data is a very challenging topic. Finding the right balance between protection of personal data and the business models based on Big Data will need to be examined and

		Big Data practices yet.	legislated. As a research topic we would like to suggest examining boundaries between personal and non-personal information. In the Big Data environment, you are able to connect non-personal information and based on this information identify the data subject which represents potential risk to rights of the data subjects.
Slovenia	Not to our knowledge.	There is no special regime for Big Data. If processing of personal data is involved, then Personal Data Protection Act applies with its existing provisions. To our knowledge, there are no plans to introduce new legislation to regulate Big Data practices. The Information Commissioner has the competence to issue non-binding decisions regarding proposals for new legislation and will and would be able to comment on such proposals.	Big Data brings substantial challenges for personal data protection and these challenges must, firstly, be well understood and adequately addressed. In our view, new concepts and paradigms, such as cloud computing or Big Data should not lower or undermine the current levels of data protection as a fundamental human right. Existing central data protection principles, such lawfulness, fairness, proportionality, rights of the data subjects and finality should not be undermined with the advent of Big Data. The rights of the individuals to informational self – determination should be cornerstone in modern information society, protected by modern data protection framework delivering efficient data protection for the individual while allowing lawful and legitimate interests, often also in the interest of the individual, to be attained. Further research issues could cover the following topics: Understanding and managing privacy risks arising from the concept of Big Data. Adequacy and effectiveness of the notion of consent in the age of Big Data. Benefits and pitfalls of the notion of "legitimate interests" as legal ground for processing personal data in Big Data environments. The principle of finality vis a vis exploiting the benefits offered by Big Data. Privacy by design and privacy enhancing technologies in connection with Big Data. Accountability and other notions of demonstrative and effective data protection vis a vis Big Data. Automated decision making and profiling – which privacy safeguards are needed?
Sweden	No	Personal data processing in general is regulated in the Personal Data Act, which in principle applies to all sectors of society. However, many public agencies have their own personal data legislation which is specifically adapted to each agency's particular activity and needs. To the extent that public agencies collect large amounts of data, this is therefore usually specifically regulated (e.g. the Tax	-

		authority which processes data for taxation purposes but also for population register purposes). Telecom and Internet service providers' collection of data may involve collection of large amounts of data and this is specifically regulated in an act that implements the e-Privacy directive. This personal data processing does not fall under our supervision but, instead, under supervision of the National Post and Telecom Agency. It might also be worth noting that further to the aim to strengthen the right to privacy, the Swedish Constitution was amended in 2010 and now explicitly mentions the right to protection against privacy infringements by surveillance or mapping of the individual's personal circumstances without his/her consent. This means that the creation of large databases which contain information that provides a comprehensive image of an individual person, must be specifically permitted in an Act by the Parliament. We are not aware of any specific plans for Big Data regulation.	
United Kingdom	We are not aware of any cases specifically to do with Big Data. This may be due to the fact that Big Data analytics can be opaque to the data subject, and so people do not necessarily realise how their data is being used.	There is no specific legal regime for Big Data, other than the Data Protection Act. It is notable, however, that there is some evidence of a move towards self-regulation, in the sense that some companies are developing what can be described as an 'ethical' approach to Big Data, based on understanding the customer's perspective, being transparent about the processing and building trust.	We note that the proposals for the new EU General Data Protection regulation incorporate some of the measures we have identified as being important in ensuring compliance in Big Data, eg. clearer privacy notices, privacy impact assessments and privacy by design. We welcome the fact that these measures are being foregrounded, although we are concerned that that they should not be seen as simply a bureaucratic exercise.

* *LL.M. MPhil Bart van der Sloot* is a senior researcher at the Tilburg Institute for Law, Technology, and Society (TILT), Tilburg University, the Netherlands. During 2014 and 2015, he participated in the 'Big Data, Privacy and Security' project of the Netherlands Scientific Council for Government Policy.

Sascha van Schendel (LL.M.) is a Phd student at the Tilburg Institute for Law, Technology, and Society (TILT), Tilburg University, the Netherlands, she worked as a trainee in this project. The basis for this article was published as a Working Paper on the internet: <http://www.wrr.nl/fileadmin/en/publicaties/PDF-Working_Papers/WP_20_International_and_Comparative_Legal_Study_on_Big_Data.pdf>.

CAD Files and European Design Law

by **Viola Elam***

Abstract: Three-dimensional printing ("3DP") is an additive manufacturing technology that starts with a virtual 3D model of the object to be printed, the so-called Computer-Aided-Design ("CAD") file. This file, when sent to the printer, gives instructions to the device on how to build the object layer-by-layer. This paper explores whether design protection is available under the current European regulatory framework for designs that are computer-created by means of CAD software, and, if so, under what circumstances. The key point is whether the appearance of a product, embedded in a CAD file, could be regarded as a protectable element under existing legislation. To this end, it begins with an inquiry into the concepts of "design" and "product", set forth in Article 3 of the Community Design Regulation No. 6/2002 ("CDR"). Then, it considers the EUIPO's practice of accepting 3D digital representations of designs. The enquiry goes on to illustrate the implications that the making of a CAD file available online might have. It suggests that the act of uploading a CAD file onto a 3D printing platform may be tantamount to a disclosure for the purposes of triggering unregistered design protection, and for appraising the state of the prior art. It also argues that, when measuring the individual character requirement, the notion of "informed user" and "the designer's degree of freedom" may need to be reconsidered in the future. The following part touches on the exceptions to design protection, with a special focus on the repairs clause set forth in Article 110 CDR. The concluding part explores different measures that may be implemented to prohibit the unauthorised creation and sharing of CAD files embedding design-protected products.

Keywords: Community design; CAD file; 3D printing; EUIPO; disclosure; informed user; spare parts; scope and criteria of protection; infringement

Recommended citation: Viola Elam, CAD Files and European Design Law, 7 (2016) JIPITEC 146 para 1.

A. What is three-dimensional printing and what is a CAD file?

1 The term "three-dimensional printing" ("3DP") can be considered as an umbrella term that stands for a set of related technologies building physical objects by the consecutive addition of liquids, sheet or powdered materials in ultra-thin layers. Hence, in contrast with traditional "subtractive manufacturing" technologies, which mostly rely on the removal of material (e.g. cutting, drilling and milling), 3DP is an "additive manufacturing" technology. The peculiarity about 3DP is that every physical object is created directly from a digital file, the so-called Computer-Aided Design (CAD) file. The latter is a virtual 3D model that serves to send information to the printer on how to build the object.

2 A CAD file can be obtained in different ways. First, it can be created from scratch, by using modelling software ("CAD software"). A number of open-source software tools are freely available online. They enable individuals with no prior experience in 3D modelling to create their own designs, with some programs providing pre-rendered shapes[1]. Furthermore, many websites offer tutorials on modelling best practices to assist users who are not design professionals[2].

3 Second, an existing object could be turned quickly into a virtual 3D model by using a 3D scanner. The latter is a device that collects a huge amount of data from a real-world object, by means of lasers

1 E.g. *FreeCAD, Sketchup or ThinkerCad*.

2 For example, *Sculpteo* provides a tutorial for users of the "*Sketchup*" 3D modelling software, available at: <http://www.sculpteo.com/en/tutorial/prepare-your-model-3d-printing-ketchup/.>

or x-rays. Hence, it reproduces a high-resolution and accurate digital model of the scanned object ("3D visualization"). Third, photogrammetry is a valid alternative to 3D scanning. It is a photography technique that uses software tools for stitching a series of 2D photographs – taken from different angles – together into a 3D model.

4 A CAD file can be saved in different formats, such as the .stl format ("STereoLithography") or the .amf format ("Additive Manufacturing Format"). The .stl format merely describes the surface geometry of a three-dimensional object as a set of triangular faces, whereas the .amf format is an XML-based format inclusive of information about the volumetric structure of the interior, composition, colour, geometry and material.

5 At a second stage, CAD files need to be processed, in order to become printable. Hence, a (CAD or scanned) 3D model has to be segmented into a number of layers by specialized software, so-called "Computer-Aided Manufacturing" software or "slicer". The latter generates a G-code for each layer, which contains commands to tell the printer how to manufacture the object[3]. The slicing programs are usually included with the printer or available online for download[4].

6 It emerges from the above considerations that three consecutive steps have to be followed in an ordinary 3DP process: the creation of a virtual 3D model; the deconstruction of the 3D model into a series of slices ("slicing"), which are sent to the 3D printer through a computer code; the final print, consisting in a layer-by-layer deposition of suitable materials.

7 3DP has gained a wider distribution among the general public in recent years. The launch of Open Source Hardware initiatives, such as the "Replicating Rapid Prototyping" ("*RepRap*") project[5], together with the expiration of a number of key patents on 3D printing technologies, have contributed to a steady improvement in the quality of personal 3D printers and to a considerable reduction in hardware costs[6]. The technology has, therefore, crossed over into the consumer sphere, with over 100,000 desktop 3D printers having been sold so far[7].

8 Furthermore, online platforms dedicated to the dissemination of CAD files ("digital-design-file-sharing") have grown in popularity. These platforms have contributed to the creation of a communication infrastructure that is a powerful tool for co-creation. They enable individuals to connect to a vast and distributed network, where they can upload, download, edit, remix, share or indeed sell a CAD file, from which a 3D printed product will emerge.

9 Some recent studies, conducted by Rayna et al.[8], Moilanen et al.[9], and Mendis et al.[10], provide examples of the diversity of existing 3DP platforms. The latter include platforms, such as *Thingiverse*, where users license their CAD files – rather than selling them – under Creative Commons licences (CC) or General Public Licence (GPL). By using CC licences, the CAD file's proprietor can withhold certain rights (e.g. the right of attribution and the right to make derivative works), and impose that derivatives should be licensed under the same terms as the licence of the original CAD file (the "Share Alike" clause). Furthermore, the "Non Commercial Use" clause restricts the possibility for the licensee to use the CAD file for commercial purposes.

10 Other platforms, such as *Cuboyo*, offer paid downloads to users' CAD files (i.e. the 30% of the sale price goes to the website, whereas the remaining 70% goes to the seller)[11]. Moreover, online platforms, such as *Shapeways* and *Sculpteo*, offer printing and delivery services on demand. Taking as an example the architecture of *Sculpteo*, the 3DP process takes place in the following way: individual users upload their CAD files onto *Sculpteo* website; *Sculpteo* automatically repairs any defect and optimizes the digital blueprint, with its own 3D tools; then, it prints the object and delivers it to costumers in finished form, charging a price for its activities.

11 Whether personal 3DP will reach its full potential in

3 The CAD and CAM functions could also be integrated into a single CAD/CAM program.

4 E.g. *Slic3r*, *Cura* and *Skeinforge*.

5 This project was launched by a research team at the University of Bath. The idea was to create an open source 3D printer capable of reproducing its own spare parts. The specifications of the hardware (e.g. CAD files, mechanical drawings, diagrams, etc.) were made freely available online for anyone to use, modify and update. The *RepRap* project could be realised because key patents, covering the fused deposition modelling technique, had expired.

6 Before 2009 the cheapest personal 3D printer on the market was offered for around €15.000. Today, the price for a personal 3D printer ranges from €500 to € 2000.

7 Mendis, Secchi, report commissioned by the UK Intellectual Property Office, *A Legal and Empirical Study of 3D Printing Online Platforms and an Analysis of User Behaviour* (March 2015), p. 2.

8 Ranya, Striukova, Darlington, *Open Innovation, Co-Creation and Mass Customisation: What Role for 3D Printing Platfroms?*, T. D. Brunoe et al. (eds.), Proceedings of the 7th World Conference on Mass Customization, Personalization, and Co-Creation (MCPC 2014), Aalborg, Springer (2014).

9 Moilanen, Daly, Lobato, Allen, *Cultures of Sharing in 3D Printing: What Can we Learn From the Licence Choices of Thingiverse Users?*, Journal of Peer Production (6), Disruption and the Law (2015), available at: <http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2440027>.

10 *Supra* note 7.

11 Moilanen et al. (2015), *supra* note 9, p. 4.

the near future is not altogether clear yet. For the time being, 3D printing by individual makers reveals major technical limitations to seriously hinder the market for quality products. The 3DP community is still a small niche, and the dissemination of CAD files on the Internet is not a mass phenomenon yet. So far, digital-design-file-sharing shows a high level of participation by 3DP enthusiasts[12]. Furthermore, to date, there is only evidence of IP infringement occurring on a small scale, on online platforms such as *Thingiverse*[13].

12 Having noted that, the speed of technical developments in consumer 3D printing is undeniable. The level of precision and accuracy attainable by a desktop 3D printer is steadily rising. The price of printing material has dropped drastically, and the technology has also become faster, more reliable and cheaper.

13 Furthermore, there are continuous attempts to make 3D printing more easily accessible and affordable for the average consumer. To give a recent example, an Italian company specialized in rapid prototyping and digital fabrication, called *Solido* 3D, came up with one of the latest innovation in 3DP. It developed a device, named "OLO", which enables printing 3D objects directly from a smartphone[14]. Because of its size, weight, and battery power source, this device is considered as the first portable 3D printer, available for sale at the price of $ 99[15].

14 It is therefore maintained that, although personal 3D printers are still far from being ubiquitous, it is just a matter of time until ordinary people will manufacture – directly from a digital file – an increasing number of items at the comfort of their home. Furthermore, in the event that products are not capable of being printed by means of personal 3D printer, it is possible to outsource the actual manufacture to bureau services, such as *Sculpteo* and *Shapeways*. Users can create online shopfronts, where they display their own CAD files for products. The Internet operator will then 3D print the product on demand and deliver it worldwide.

12 Based on data extracted from 17 online platforms, the total number of downloads, for the time-period 2008-2014, is around 40.000. *See* Mendis, Secchi, UK Intellectual Property Office (2015), *supra* note 7, p. 28.

13 *See* Hamdi, *IP Law vs 3D Printing: the 5 Worst Examples*, Trinckle blog (September 18, 2015). Retrieved February 7, 2016 from <https://www.trinckle.com/blog/ip-vs-3dp/>.

14 *See* <www.olo3d.net>.

15 OLO is a crowd-funded project, launched on *kickstarter* on March 21, 2016. The aim of the developers of this portable 3D printer was to raise $ 80.000 in one month, whereas they reached this goal in just thirty-three minutes. At the end of the fundraising campaign, they collected $ 2,321,811 with 16,180 backers from all over the world.

I. Is a CAD file a computer program?

15 An emerging body of literature has addressed the question of whether CAD files warrant copyright protection[16]. Some scholars have suggested that the definition of computer programs is perfectly compatible with CAD files[17]. The present writer, however, rejects former existing assumptions on the application of this analogy.

16 As noted above, a CAD file simultaneously encompasses both a "design drawing component" and a "code component". The latter serves to give a series of instructions to the printer (i.e. where to move the print head and how fast to deposit the material). Even if a CAD file embeds a code, it is not the equivalent of a computer program. This in light of the fact that the designer of a CAD model does not write the code herself, at least not directly.

17 As noted by Dolinsky, "the CAD designer ... "creates" the code necessary to print the object only by creating the design", whereas the CAD software programmer has already predetermined the code associated with a pre-made shape or a free-hand drawing[18]. It is therefore only the CAD software that finds protection under the Software Directive 2009/24/EC, not the CAD file itself.

18 Although an enquiry on the copyright status of CAD files goes beyond the purposes of the present analysis, it is here suggested that a CAD file merely serves as the medium in which a copyright protected work (i.e. artistic work) is recorded. To the extent that the design drawing component of the CAD file is the expression of the author's creativity, and is not dictated by purely functional considerations, it may qualify as a copyright protected work. By contrast, the file itself is just the medium in which the work is recorded. The fact that the work exists in digital file format does not change its nature. In this respect, a CAD file bears a certain similarity to other files,

16 Among others: Rideout, *Printing the Impossible Triangle: The Copyright Implications of Three-Dimensional Printing*, The Journal of Business, Entrepreneurship & the Law (2011), 5(1); Simon, *When Copyright Can Kill: How 3D Printers Are Breaking the Barriers Between "Intellectual" Property and the Physical World*, PIPSELF (2013), 3; Weinberg, *What's the Deal with Copyright and 3D Printing*, available at <https://www.publicknowledge.org/news-blog/blogs/whats-the-deal-with-copyright-and-3d-printing> (2013); Dolinsky, *CAD's Cradle: Untangling Copyrightability, Derivative Works, and Fair Use in 3D Printing*, Washington and Lee Law Review (2014), 71(1); Mendis, *Clone Wars Episode II – The Next Generation: The Copyright Implications Relating to 3D Printing and Computer-Aided Design (CAD) Files,* Law, Innovation and Technology (2014), 6(2).

17 *See*, among others, Bradshaw, Bowyer, Haufe, *The Intellectual Property Implications of Low-Cost 3D Printing*, SCRIPTed (2010), 7(1), p. 24.

18 Dolinsky, *supra* note 16, p. 641.

such as JPG or PDF files, which respectively embed a photograph or a literary work.

19 Therefore, the distinction between what is the work of authorship, as opposed to the medium of its expression, takes on a particular significance when claiming copyright protection of CAD files.

II. Aim of the present analysis

20 This paper explores whether design protection is available under the current European regulatory framework for designs that are computer-created by means of CAD software, and, if so, under what circumstances. The key point is whether the appearance of a product, embedded in a CAD file, could be regarded as a protectable element under existing legislation. To this end, it begins with an inquiry into the concepts of "design" and "product", set forth in Article 3 of the Community Design Regulation No. 6/2002 ("CDR"). Then, it considers the EUIPO's practice of accepting 3D digital representations of designs. The enquiry goes on to illustrate the implications that the making of a CAD file available online might have. It suggests that the act of uploading a CAD file onto a 3D printing platform may be tantamount to a disclosure for the purposes of triggering unregistered design protection, and for appraising the state of the prior art. It also argues that, when measuring the individual character requirement, the notion of "informed user" and "the designer's degree of freedom" may need to be reconsidered in the future. The following part touches on the exceptions to design protection, with a special focus on the repairs clause set forth in Article 110 CDR. The concluding part explores different measures that may be implemented to prohibit the unauthorised creation and sharing of CAD files embedding design-protected products.

B. Designs in the European Union

21 For the purposes of the Community Design Regulation ("CDR"), "design" means "the appearance of the whole or a part of a product"[19]. While there is no definition of appearance, Article 3(a) CDR provides a non-exhaustive list of elements that one may have to consider, for appraising the external aspect of a product. These elements include the lines, contours, colours, shape, texture and/or materials of the product itself and/or its ornamentation. These features are all perceivable by the human eye or by the sense of touch, whereas sounds and smells are not contemplated[20]. Yet the CDR does not make the eye appeal a necessary prerequisite for registration.

22 Furthermore, as noted by the European Commission in the 1991 "Green Paper on the Legal Protection of Industrial Design"[21], the external aspect of a product is of considerable economic importance. The notion of appearance, therefore, should be broad enough to encompass any economic value attached to the aspect of a product.

23 Article 3(b) CDR goes on to define a product as "any industrial or handicraft item other than computer programs". It then offers some guidance as to the type of designs that are eligible for protection. The latter include both three-dimensional designs, such as packaging and get-up, and two-dimensional designs, such as graphic symbols and typefaces[22]. Designs of parts of products, for which no assembly is required, and designs of component parts, which are intended to be assembled in a larger complex product, can also be protected[23].

24 The concept of product, therefore, is central to the whole structure of the CDR. The essence of a design is the appearance of a product. Furthermore, as explained below in more detail, there should be a product, to which the design is applied, in order to commit an infringement.

25 The following enquiry aims at analysing whether the appearance of a product that is represented digitally as a CAD file may attract design protection under existing EU legislation.

I. The visual element of a CAD file

26 The ultimate generation of designs created by means of CAD software embed all information that is needed to define the outer appearance of a product. Embedded data can describe the geometry, as well as the colours and materials of the product. In this respect, the design-drawing component of a CAD file differs from traditional blueprints or technical drawings. In most cases, blueprints only define the geometrical aspect of an object. They may be seen as graphical abstractions of the intended product that need to be interpreted by a human being.

27 CAD files, instead, may define all the properties

19 Article 3(a) CDR.

20 *See* Suthersanen, *Design Law: European Union and United States of America*, 2nd edn., Sweet and Maxwell (2010), p. 95.

21 European Commission, Green Paper on the Legal Protection of Industrial Design, June 1991, at 2.1.2.

22 The list of products enumerated in this provision is not intended to be exhaustive.

23 Article 3(c) CDR.

and attributes of the product to be printed. They may contain the entire product design that, when printed, will be a finished 3D product. In such a case, in parallel with a photograph, the visual element of a CAD file – i.e. the image of the product stored therein – may be regarded as a view of the appearance of the finished product, for which protection is sought.

28 The CAD file can be seen as the medium in which the design is first recorded. Hence, as noted below, the appearance of the product embodied therein may enjoy Community design protection, irrespective of whether the product comes into existence or not.

29 By contrast, in case where a CAD model does not clearly reveal the outer appearance of the product, it may be allegedly considered as a blueprint protected in class 19-08 of the Locarno classification (i.e. "other printed matters")[24].

II. Digital item embedded in a CAD file: design protection?

30 The issue at stake is whether a digital item, which is computer created by means of CAD software and recorded in a CAD file, may attract design protection in its own right, as a graphic symbol.

31 At first reading, the notion of a "design" seems to be confined to the appearance of products having some physical form, insofar as the CDR makes express reference to "industrial or handicraft items", and expressly excludes computer programs.

32 Hence, the definition of "product" set forth in the CDR may give rise to a certain degree of uncertainty as to whether a design, that is not applied to a product in the sense of a physical, tangible object, should likewise be considered as a protectable element under existing regulation. The inclusion of graphic symbols in the meaning of products indicates that protectable designs need not be tied to a physical dimension. The European Union Intellectual Property Office ("EUIPO") guidelines provide some assistance in this respect. The Office's practice allows registration of screen displays, icons, and other visible elements of a computer program, such as graphical user interfaces ("GUIs"), in Locarno class 14, Subclass 04 ("screen displays and icons")[25].

33 Therefore, design protection appears to cover all digital items – with the sole exclusion of sounds and animated images – that appear on electronic devices, such as computer screens or mobile phones.

34 By contrast, computer programs as such are excluded from design protection. The reason for inserting this exclusion is explained in the Explanatory Memorandum of the first proposal of the CDR. The rationale was mainly to avoid any potential interference with the Software Directive which might arise whenever copyright protection provided under the aforementioned Directive is supplemented or reinforced by a protection of the "look and feel" of a computer program by way of design protection[26].

35 In theory, if a computer item is eligible for protection, the digital item represented in a CAD file could likewise enjoy Community design protection. One may argue that, in parallel with other computer icons, the item embodied in a CAD file is a graphic symbol that appears on a computer screen when the file is loaded. Since the CDR enumerates graphic symbols as a category of product of their own, their appearance can be protected under the title of Community designs.

36 This analogy appears questionable, however. There is a substantial difference between a computer icon and a product embedded in a CAD file. The former fulfils its function exclusively once it is displayed on a computer screen. As noted by Margoni, it does not even possess the characteristics to be manufactured or printed into an industrial or handicraft item[27]. It is "intangible" by its very nature. On the contrary, not only can an item embedded in a CAD file become tangible once it is shown on a computer screen, but it could also be turned into a physical product at the click of a button.

37 In other terms, the reason why a person creates a CAD file is to enable the manufacturing of the object embodied therein. Design rights are vested in the appearance of the product to be made from that file.

38 Having noted that, an applicant may also wish to protect the virtual product per se under the scope of European design law in order to avoid the risk that no protection will be available, should a third party make a digital copy of its design, i.e. creates a CAD file depicting the design-protected product and uploads such CAD file on the Internet. In fact, as discussed below in more detail, Article 19 CDR seems to confine infringing use of a design to use in relation to physical goods or corporeal movables[28].

24 *See* EUIPO Guidelines for Examination of Registered Community Designs, version of 01/08/2016, at 4.1.1.

25 *Id.*, EUIPO Guidelines, at 4.1.3. Also, the Explanatory Memorandum clarifies that the exclusion provided in the CDR for computer programs does not extend to "specific graphic designs as applied, for example, to icons or menus". *See* EU Commission, Explanatory Memorandum on the Proposal for a European Parliament and Council Regulation on the Community Design, 3 December 1993, p. 11.

26 *Id.*, Explanatory Memorandum, p. 11.

27 Margoni, *Not for Designers: On the Inadequacies of EU Design Law and How to Fix It*, JIPITEC, (2013), 4(3), p. 232.

28 *See* below, part G.

39 For the sake of clarity, it seems rather unrealistic to assume, in the absence of a specific provision, that digital items represented in the form of CAD files could be seen as "products", whose appearance deserves protection in its own right. The definition of "design" would need to be broadened in future legislation in order to cover a wider range of "immaterial" protectable elements.

III. CAD drawings as graphical representations

40 Under current practice, the drawing component of a CAD file may serve as a "graphic representation" of the design for which protection is sought. The following analysis attempts to clarify this point.

41 A person seeking protection for their design at the EU level has the option to either apply for a Registered Community Design ("RCD") through the EUIPO, before disclosing it, or, alternatively, opt for an anti-copying right, relying on Unregistered Community Design ("UCD") protection.

42 In the former case, in order to have a valid application for a RCD, Article 36 CDR requires to include "a representation of the design that is suitable for reproduction". There are several ways in which a design can be represented. The EUIPO accepts drawings, photographs and computer-made representations (i.e. CAD representations), either in black or in colour, provided that they are of quality, permitting all details of the design to be clearly distinguished[29].

43 Hence, it may well be that, in parallel with a photograph, a CAD representation is used to disclose the features of the design for which protection is sought. By way of explanation, if the applicant wishes to register the design of a table knife, rather than affixing a photograph of such a knife, she can affix the 3D representation, created by CAD software, of the same household good. The applicant will then have to indicate "knives" as the relevant product category (class 7-03 of the Locarno classification)[30].

44 However, in order to be of quality, a CAD representation should enable to determine, with clarity, the subject matter of the protection afforded by the RCD to its holder. Hence, it should contain clear and intelligible information about the sizes, dimensions, and colours of the item in which the design is incorporated or to which the design is applied.

45 Interestingly, a recent decision from the UK Supreme Court, *PMS International Group Plc*[31], is notable for stressing the importance of the images affixed in the application form, for determining the scope of Community design protection. As noted by Lord Neurberger, when it comes to deciding the extent of protection afforded by a RCD, the question "must ultimately depend on the proper interpretation of the registration in issue, and in particular of the images included in the registration"[32]. Therefore, it will almost always be the images that "exclusively identify the nature and extent of the monopoly" which the applicant is claiming[33].

46 The case concerned an alleged infringement of a RCD, which consisted of six images prepared by CAD software of an item (a ride-on animal suitcase) whose main body appeared as a uniform grey, but which had black strips in the front, a black strap on the top and black wheels. After analysing these images, it was not clear whether the two-tone colouring on the CAD images – i.e. the contrast in colour between grey and black – was simply an artefact of the computer-generated process or a visual cue to indicate that the wheels and the strap should be considered as separate components. The problem, therefore, was whether the RCD was to be considered as protection for the shape only, or for the shape in two contrasting colours. Only in the latter case, the overall impression created by this contrast in colour could be considered.

47 It might be, therefore, that CAD representations depict some unnecessary tonal contrast. This, in turn, could generate confusion and be understood as limiting the scope of design protection to certain colours only.

48 The application for a RCD should also indicate the products in which the design in intended to be incorporated or to which it is intended to be applied. In this respect, it is worth noting that product classification mainly serves administrative purposes and does not affect the scope of design protection[34]. Once the design is registered, it is protected against any use in relation to any product that does not produce a different overall impression on the informed user.

49 Once the request for registration is filed, the EUIPO carries out an *ex officio* examination of the two absolute grounds for non-registrability, set forth in

29 EUIPO Guidelines, *supra* note 25, at 3.3.1.

30 The EUIPO has recently released an e-filing tool, called "3D image uploader", that allows the applicant to upload and store its CAD files. The applicant can move the 3D image, zoom in and out, take some pictures from different views, and select between a maximum of 7 static views.

31 *PMS International Group Plc v Magmatic Ltd* [2016] UKSC 12.

32 *Id.*, at 30.

33 *Id.*, at 31.

34 EUIPO Guidelines, *supra* note 25, at 6.1.4.1.

Article 47 CDR. Namely, the Office verifies whether the subject matter of the application corresponds to the definition of a design foreseen in Article 3(a) CDR, and whether or not it is contrary to public policy and accepted principles of morality. Therefore, the registration procedure is kept to a minimum. Compliance with the novelty and individual character requirements will only be examined at a second stage if a third party submits an application for a declaration of invalidity.

50 It should also be noted that the EUIPO examines whether the appearance of the "product" is disclosed in the light of the design itself. Whether the product is actually made or used, or can be made or used, in an industrial or handicraft fashion, is not taken into consideration[35]. In fact, there is no requirement to submit a specimen of the claimed RCD.

51 This, in turn, implies that a person can: create a CAD file for a product by means of CAD software; include in the application for a RCD an image taken from such a CAD file; obtain a design registration covering the product design represented therein, irrespective of whether the product is actually manufactured or not.

52 This leads to the outcome that, although the entire regulatory framework in EU design law is structured on the concept of "product", a design is protectable regardless of whether a product comes into existence or not. Accordingly, legal protection does not depend on whether designs represented as CAD models exist as tangible articles or not.

C. Unregistered design protection of CAD files

53 An Unregistered Community Design ("UCD") is based on the same substantive provisions postulating the validity requirements for a RCD. The meaning of "design", "appearance" and "product" are the same for both RCD and UCD. As a general matter, any design capable of being registered at the EU level could also benefit from the protection granted to UCD.

54 There are, however, substantial differences between RCD and UCD. A RCD confers a true monopoly, whereas an UCD grants the right to prevent any commercial use of a design that is an intentional copy of the protected one. Yet, it should not be demonstrated that the alleged infringer acted in bad faith. Furthermore, a RCD confers protection for up to 25 years, subject to renewal each five years, whereas an UCD affords protection for only three years.

55 Protection of the UCD commences from the date on which the design has been "made available to the public" within the EU. As Recital 16 CDR puts forth, there is no need to register products having a short market life. A designer can introduce a new design testing the market and file an application for registration at a second stage. In fact, the designer is entitled to register her design within a 12-month period ("grace period") from the date of the first disclosure. In other words, in the event that the designer files an application for a RCD, disclosure during the year preceding the date of filing shall not be taken into consideration when appraising novelty and individual character of the design in question, pursuant to Article 7(2) CDR.

56 Let us now assume that a CAD file for a product to which a design is applied is uploaded onto a website which is a 3DP marketplace or repository. This, in turn, raises a number of questions. Should the act of uploading a CAD file onto an online 3DP platform be tantamount to a disclosure of the design to the public, which triggers UCD protection? Has the design been "made available to the public", and become known in the normal course of trade? Has the 12-month grace period commenced?

57 The following part of this paper detects the circumstances under which a design shall be deemed to have been made available to the public. The phrase "made available to the public", for the purposes of identifying the date on which UCD protection commences, is defined under Article 11(2) CDR. This provision mirrors Article 7(1) CDR, which clarifies when a design has been disclosed for considering questions of novelty and individual character, for both registered and unregistered designs. In fact, all designs made available to the public, prior to the relevant date (indicated at Article 5(1)(a)&(b) and 6(1)(a)&(b)), are to be taken into account to determine whether a design is new and if it has individual character. This, in turn, raises an additional question: should we consider all the CAD models that have been previously uploaded onto 3DP online platforms as antecedent designs in the prior art?

58 It should also be noted that a disclosure should take place within the territory of the European Union in order to create an UCD. Hence, UCD protection is not afforded to designs that have first been made available outside the EU. On the contrary, this requirement is not imposed under Article 7 CDR, which defines the notion of disclosure that is relevant for determining the state of the prior art.

35 *Id.*, at 4.1.

D. The concept of "made available to the public"

59 Articles 7(1) and 11(2) CDR provide some guidance to assess whether a design has been 'made available to the public'. This is the case if "it has been published following registration or otherwise, or exhibited, used in trade or otherwise disclosed". The following part of these provisions set forth the so-called "safe-guard" clause, stipulating that a disclosure shall not be taken into consideration if these events (publication, exhibition, and use in trade) could not have become known "in the normal course of business to the circles specialised in the sector concerned", operating within the Community.

60 The EUIPO[36]'s case law from 2004 onwards allows enough clearance on which acts constitute a disclosure of a design to the public, which could also become known in the normal course of business to specialised circles.

61 A remarkable ruling that helps us to understand better whether the publication of a CAD file on an online platform would amount to a disclosure to the public is the Board's decision in *Crocs, Inc. v Holey Soles Holdings Ltd*[37]. The holder of a RCD for Crocs clogs, which was published in the Bulletin of 8 February, 2005, conceded that the design had already been published on www.crocs.com before 28 May, 2003. Nonetheless, the right owner argued that such disclosure on the website did not destroy novelty of the design in question, since it could not have reasonably become known in the Community.

62 At that time, the website was unsophisticated and virtually impossible to access. The website merely functioned as an information tool for persons "who might have learnt about the clogs from people who had already bought them" and was not used as a large mail order service. Websites that will be regarded as a source of inspiration for developing new designs are those of the established footwear companies, such as *Nike or Adidas*, whereas *Crocs Inc.* was not an established manufacturer at the relevant date[38].

63 The Third Board of Appeal dismissed the appellant's findings. In the first place, the Board found that the Internet is a formidable information tool and is used by designers in footwear as well as in other fields as a resource in the development of their designs. Moreover, *Crocs* website was an active website already at that date and was configured to function as a sales channel. Henceforth, the audience targeted by the website was not only composed by those who knew *Crocs* from before[39].

64 Accordingly, when a design is published on a website, it will *per se* be publicly disclosed and reasonably become known in the normal course of business, even if the circles specialised in the sector were not aware of the website owner at that date[40]. This is further confirmed in recent case law from the EUIPO. As a matter of principle, information disclosed on the Internet or in online databases forms part of the prior art and is considered to be publicly available as of the date the information was posted[41].

65 Moreover, neither restricting access to a limited circle of people (for example, by using password protection) nor requiring payment for access (in the same way as requiring a payment for subscribing to a journal or purchasing a book) prevent a webpage from being part of the prior art. The European circles specialised in the sector concerned could reasonably meet the accessibility requirement[42].

66 A disclosure shall be deemed to be obscure and irretrievable only in situations in which a design disappears from mankind's memory over time and is available only in a local museum or traded on a remote local market. This is not the case for prior designs made available online. Users – either the broad public or experts in a particular field of industry – use the service of web browsers, such as *Google* or *Yahoo*, to search on the Internet. By using keywords, they can easily find websites dealing with a particular subject matter. Therefore, once a design is published on the Internet it becomes automatically accessible and retrievable[43].

67 For the purposes of applying Articles 5 and 6 CDR, a disclosure could also take place outside the EU, insofar as the design has become known in the trade circles in the European Union. The question of whether events taking place outside the EU could reasonably have become known to persons forming part of specialized circles in the EU is a question of fact, dependent on the particular circumstances of each individual case[44]. In theory, even where the

36 Formerly called Office for Harmonization in the Internal Market ("OHIM").

37 OHIM Third Board of Appeal, decision of 26 March 2010 – R 9/2008-3.

38 *Id.*, at 10(d).

39 *Id.*, at 85-92.

40 Suthersanen, *supra* note 20, p. 126.

41 OHIM Invalidity Division, *Mariusz Adamski Adams Group v Abakus Direct Ltd*, decision of 10 July 2014, at 13. In the present case the holder had disclosed its design on *eBay* prior to the RCD's filing.

42 OHIM Invalidity Division, *Napco Beds B.V. v Leopold Meijnen Oosterbaan*, decision of 24 February 2015, at 13.

43 OHIM, Invalidity Division, *Samsung Electronics CO. Limited et al. v Apple Inc.*, decision of 05 July 2013, at 70-71.

44 See the CJEU's ruling in *H. Gautzsch GroBhanden GmbH & Co. KG v Munchener Boulevard Mobel Joseph Duna Gmbh*, C- 479/12,

design has been disclosed to a single undertaking within the EU, a disclosure of that kind may, indeed, be sufficient for that purpose[45].

68 Making a design available overseas, therefore, may destroy novelty on the basis that Article 7(1) CDR is not geographically restricted to the EU. On the contrary, the same disclosure taking place outside the EU may not be sufficient to commence UCD protection, given the territorial qualification contained in Article 11(1) CDR[46].

69 It is therefore maintained that, in principle, the act of uploading a CAD file onto an online platform should be a sufficient ground for "disclosing" the design represented therein, for the purposes of applying Articles 5 and 6 CDR. A CAD file is retrievable and easily accessible by Internet users, including experts in the field. This might be the case for both CAD files that have been made available to the public, subject to a Creative Commons licence, and those offered for sale in 3DP marketplaces.

70 It follows that whether the design is new and has individual character would need to be considered, taking into account the already-available body of designs, including all antecedent CAD files that have been previously disclosed. In other words, product designs embedded in CAD files that have already been distributed online will form part of the state of the prior art[47].

71 The publication of the CAD file on a EU website can also trigger UCD protection from the date of the first online publication, if the criteria for protection (i.e. novelty and individual character) are met. The designer would then have the option to register the design within one year.

72 An unsettled issue is whether UCD protection is activated if the CAD file is first uploaded onto a website that is hosted outside the EU (such as *Thingiverse*). If the website is easily accessible by European users, a positive answer may appear as more appropriate in light of the above-mentioned case law, which focuses on the retrievability of Internet publications, whereas a literal interpretation of Article 110a (5) CDR may suggest the opposite.

73 It should also be noted that the CAD file made available online should clearly reveal the outer appearance of the product for which protection is sought. Lacking a clear representation of the product design, the act of publishing the CAD file on a website will not constitute a relevant disclosure for the purposes of Articles 7 and 11 CDR.

74 The option of making CAD files available online, therefore, constitutes an interesting possibility for those designers that want to prevent third-parties from using their 3D models to obtain design protection[48]. When the CAD file is disclosed, all later designs will have to produce a different overall impression on the informed user.

E. Requirements that a design has to meet towards design protection

75 Articles 5 and 6 CDR state that a design has to be new, has to have individual character, and must not fall foul of any of the stipulated exceptions, in order to enjoy design protection. These requirements will be analysed in turn, focusing on the implications that 3DP carries.

I. Novelty and individual character

76 A design is new only when it differs materially from everything that has been produced before. In fact, Article 5(2) CDR states that differences between two designs are irrelevant whenever they relate to mere "immaterial details". In this regard, the novelty requirement is much closer to that for utility patents, rather than the originality requirement for copyright protection. It follows that users who download already-existing CAD models from a 3DP platform will have to modify them substantially in order for their designs to be new.

77 In this respect, a critical issue that 3DP poses is whether customized designs differ materially from other designs that have been made available before. Today, many companies, such as *eMachineShop.com* or *Shapeways*, manufacture customized products based on consumers' CAD files. From an IP perspective, a key issue is whether customized products provide "added value" because they imprint true novelty, or because they just enhance the value inherent in the design of the core product. It may well be that customized designs lack in novelty, since they differ from the core product design in details that are immaterial, banal or commonplace.

at 34. See also OHIM Board of Appeal, *Kirschenhofer GmbH v WS Teleshop International Handles-GmbH*, decision of 11 July 2007.

45 *Id.*, *H. Gautzsch GroBhanden GmbH*, at 15.

46 See the decision of the German Federal Supreme Court of October 9, 2008, *Gebackpresse* I ZR 126/06, [2009] GRUR 79.

47 More precisely, in order to pass the novelty and individual character test, the design embedded in the CAD file shall differ from all the designs made available before: the date on which the file itself was published on the 3DP website, with respect to UCD; the date of filing or validly claimed priority, with respect to RCD.

48 Margoni (2013), *supra* note 27, p. 241, at 113.

78 Novelty and individual character overlap to a certain extent. The main difference between these criteria lies in the kind of examination carried out by the EUIPO. When assessing novelty, the EUIPO makes a comparison between the overall appearances of the two designs. In contrast, when measuring individual character, the EUIPO considers the overall impression that the design produces on the "informed user". Therefore, any reference to the informed user is not justified when assessing novelty. It is the Board's task to measure the differences between the designs under examination on the basis of their overall appearance[49].

79 The test for individual character is less straightforward and is likely to give rise to slightly more subjective appraisals[50]. In *Karen Millen Fashions*[51], the CJEU held that, in order for a design to be considered to have individual character, the overall impression which that design produces on the informed user must be different from that produced on such a user "not by a combination of features taken in isolation and drawn from a number of earlier designs, but by one or more earlier designs, taken individually".

80 Therefore, the assessment as to whether the product design embedded in a CAD file has individual character must be conduced in relation to individualised, defined and identified designs that have been made available to the public previously.

81 Furthermore, in its recent decision in *H&M Hennes & Mauritz BV*[52], the CJEU held that the assessment of the individual character of a Community design is the result of a four-stage examination, which consists in deciding upon: first, the sector to which the products belong; second, the identity of the informed user of those products; third, the designer's degree of freedom in developing his design; fourth, the outcome of the comparison of the designs at issue. The designer's degree of freedom cannot, on its own, give rise to an outcome as regards the assessment of individual character, but can only "reinforce" this evaluation. The starting point should always be the perception of the informed user.

82 The problem is how to carry out the four-stage examination of the individual character requirement with respect to CAD files. In order to be protectable, a product design in the form of a CAD file should produce an overall impression on the informed user that differs from the impression produced by all previous designs. Therefore, such a design will only pass the individual character test if it differs from: a) any CAD file for a product that has been previously uploaded onto a 3DP platform; b) any product that has already been marketed.

83 The situation is further complicated by the contention that the informed user of an item represented as a CAD file might need to be distinguished from the informed user of the corresponding physical product. Arguably, the former should be the user of a 3DP platform, who wants to 3D print the item, rather than the person who purchases the product in a retail store.

84 Let us assume that a CAD file represents a bottle opener, and that a later CAD file depicts a similar bottle opener. In potential litigation, the informed user for assessing the individual character requirement of the disputed design could be: a private individual who drinks wine; a professional (e.g. waiter or sommelier); the user of a 3DP platform, who wants to manufacture the bottle opener at home.

85 Therefore, a number of issues need to be addressed. Who is the informed user of CAD files? How should we evaluate the degree of freedom of the CAD file's designer? Will the individual character threshold become less strict in the future if the market sectors become overcrowded? The next paragraph suggests some possible answers to these questions.

II. The "informed user" in the 3D printing landscape

86 For the purposes of this analysis, it is worth asking, in the first place, who would be the notional informed user, if an increasing number of individuals engage in the creation of CAD models and in digital-design-file-sharing. Everyone can now design a product from scratch by using CAD software. Users can also download third parties' CAD files and use online tools to transform, adapt or recast the pre-existing designs. Individual makers are both users and designers. Hence, the following analysis suggests that, if it becomes common practice that people not only print but also design their own product at home, the notion of informed user might need to be revisited in the future. It argues that informed users would tend to belong to the circles specialised in the sector concerned, and resemble the "person skilled in the art" in patent law.

87 The legal concept of "informed user" differs from that of "average consumer" in EU trademark law. The possibility of imperfect recollection on the

49 OHIM Third Board of Appeal, *Imperial International Limited v Handl Cookware Limited*, decision of 2 September 2008, at 11-12.

50 OHIM Third Board of Appeal, *Daka Research Inc. v Ampel 24 Vertiebs-GmbH & Co. KG*, decision of 22 November 2006, at 20.

51 *Karen Millen Fashions Ltd v Dunnes stores et al.* C-345/13 ECDR 17, at 35.

52 *H&M Hennes & Mauritz BV & Co. KG v OHIM – Yves Saint Laurent* (handbags) T-526/13, at 32-34.

part of the average consumer plays a vital role in trademark law, which is aimed at preventing consumer confusion or deception. To the contrary, design law protects the appearance of a product. This implies that the informed user should not merely half-remembering the articles but also have a certain degree of familiarity with the item goods in which the design is incorporated[53].

88 Hence, according to established case law, the informed user shall be particularly observant, aware of the state of the art in the sector concerned, and use the product related to the RCD in accordance with the purpose for which the product is intended[54]. The background knowledge of the items is certainly higher than average, but not even too specific. She is more than a mere consumer, but is less than a design expert. Moreover, Lord Justice Jacob, in *Procter & Gamble Company v Reckitt Benckiser* (UK) *Limited*[55], highlighted that the informed user is not the same sort of person as the 'person skilled in the art' of patent law. The equivalent to that person in the field of design would be some sort of average "designer", not a "user".

89 Originally, the EUIPO's Invalidity Division adopted a rather different approach. The informed user was found to be a person aware of the prior art known in the normal course of business to "the circles specialised in the sector concerned". She does not ignore the specific methods and techniques of production[56]. For example, in a case concerning an application for a declaration of invalidity of a RCD for "wheels for bicycles", the Invalidity Division found that the informed user is aware of the requirements that bicycle wheels must fulfil in order to perform their function. Therefore, the informed user also "takes into account whether the degree of freedom of the designer is limited by the requirement that a wheel has to be laced with the spokes between the hub and the rim and to transfer the weight of the rider to the rim"[57].

90 It thus seems that the notion of informed user was once much closer to that of a design expert. The Invalidity Division used to consider the informed user as belonging the "circles specialised in the sector concerned". Nonetheless, as noted above, this criterion should only apply when establishing what is a relevant disclosure to the public, and potential conflicts with an already-existing design corpus, under Article 7 CDR. The person of the informed user, who is the reference for evaluating individual character, shall not be part of any specialised circle, lacking this sort or requirement in Article 6 CDR.

91 A correct interpretation of these two provisions should be that a design is considered to have individual character if the overall impression it produces on the informed user differs from that of an earlier design, which has already been disclosed to the public. However, a design shall *not* be deemed part of the prior art if not even the circles specialised in sector concerned, operating in the territory of the EU, are aware of its existence[58].

92 Therefore, in a recent ruling, the Board of Appeal found that the informed user of clogs is "neither the manufacturer nor a seller of clogs, but the person who wears clogs. Without being a designer or a technical expert, the informed user knows the various designs for clogs as a result of the relevant product range offered in retail shops or over the Internet"[59]. In the present context, footwear designers and footwear industry, operating in the EU, represent the circles specialised in the sector concerned.

93 This paper argues that 3DP may blur the distinction between the notions of informed user and that of design expert. Users may become more and more aware of the specific methods and techniques of production. If this is the case, one will have look at early case law from the EUIPO in order to detect who should be considered the informed user, in a new ecosystem where the person of the designer and that of the user conflate to a greater extent.

III. How to evaluate the designer's degree of freedom

94 Following established case law from the EUIPO, the designer's degree of freedom is likely to be lower if she has to comply with technical constraints. Similarly, if a field of application is already very crowded, minor advances from the prior art might produce a different overall impression on the informed user[60].

53 *Procter & Gamble Co v. Reckitt Benckiser* (UK), Ltd [2007] EWCA Civ 936, per LJ Jacob at 27.

54 Judgement of the General Court (First Chamber), 9 September 2011, in Case T-10/08, *Kwang Yang Motor Co. Ltd. v OHIM,* at 23.

55 *Supra* note 53, at 16.

56 OHIM Invalidity Division, *Eredu S. Coop v Arrmet S.r.l.*, decision of 27 April 2004, at 18: "in particolare, l'utilizzatore informato non ignora lo stato della tecnica quale è conosciuta nel corso della normale attività commerciale negli ambienti specializzati del settore considerate".

57 OHIM Invalidity Division, *Rodi Commercial S.A. v ISCA S.p.A.*, decision of 30 August 2005, at 27.

58 See, *inter alia*, opinion of Advocate General Wathelet, 5 September 2013, in case *H. Gautzsch GroBhanden*, *supra* note, at 44.

59 OHIM third Board of Appeal, *Hessy s.r.o. v Crocs, Inc.*, decision of 14 September 2015, at 16.

60 By way of example, the OHIM third Board of Appeal, in *Mafin S.p.A. v Leng-D'Or S.A.*, decision of 4 November 2010, at 20-21, found that the presence of so many shapes for

95 On the one hand, when applying this reasoning to 3DP, one could maintain that the designer's degree of freedom will be gradually reduced. Assuming that an increasing number of users and companies will start producing and distributing their own versions of CAD files, and that such files form prior art, many market sectors will be thoroughly soaked. If a specific sector is saturated, it inevitably entails compromises, since minor differences in the appearance of products might be enough to lead to a different overall impression on the part of the informed user. The appearance of a contested design, therefore, might be very similar to that of an earlier design and, nonetheless, lead to a different overall impression.

96 Besides, it is worth considering that the designer has to work within certain constraints in order to make a 3D model suitable for printing. In the first place, there are some dimensional constraints. The designer has to comply with height and size requirements in order for the 3D printer to be used. In other words, when designing the 3D model using modelling software, the designer should take into account that printed objects are limited by the printers' size[61]. Furthermore, a 3D model should have a minimum thickness, at any given point ("minimum wall thickness"), which depends on the material used. Arguably, all of these technical constraints limit the designer's freedom.

97 On the other hand, one may argue that 3DP enhances the designer's freedom, since it enables the creation of much more complex geometries, as opposed to traditional manufacturing processes. Furthermore, individuals have gained the capacity to design all sorts of products with a relatively low experience. It is also possible to find tutorials on the Internet on how to use modelling software, such as CAD software. 3D scanners enable the designer to digitize without difficulty any physical object. The newly-created 3D model can then be modified, adapted and optimized.

98 Thus, it is questionable whether the designer's degree of freedom should be considered lower in 3DP than in other design processes. This issue, however, is dependant on whether the technology will or will not become widespread. As noted above, for the time being, individual users engaging in the creation and sharing of CAD files mainly include 3DP enthusiasts.

F. Exceptions to Community Design protection: the non-harmonisation of the repairs clause

99 The scope of design protection for the appearance of items represented as CAD files – and the corresponding 3D-printed products – is narrowed by a series of exceptions, set forth in the CDR. The first functionality exclusion, provided in Article 8(1) CDR, states that a Community design shall not subsist in features of appearance of a product, which are solely dictated by its technical function. Such features shall not only be necessary, but essential to obtain a technical result. Thus, the level of functionality required is higher than that provided under trademark law.

100 In a way, such exclusion emulates the idea and expression dichotomy in copyright law. In fact, in the 1991 Green Paper on the Legal Protection of Industrial Design[62], the European Commission made clear that if the designer has a choice among various forms, in order to arrive at the technical effect, the features in question could be protected. This, in turn, means that features of appearance of a product, represented as a CAD file, will not be granted protection if they are only indispensable for achieving a specific technical result. It does not follow, however, that the whole design will automatically be denied protection.

101 Over and above the general exclusion of "technical function", Article 8(2) CDR provides the so-called "must-fit" exception or "interface" exclusion. This exclusion is aimed at enabling technical replacement products and ensuring mechanical interoperability. Hence, no protection is given to those features that must necessarily be reproduced in their exact form and dimension in order to permit the product, in which the design is incorporated, to be mechanically connected to another product (for example, exhaust pipes or coupling sleeves are examples of "must fit" designs in the automotive industry). This permits the possibility that either product may perform its function.

102 This provision turned out to be rather redundant, insofar as spare parts, which are not visible in normal use[63], and those that are solely dictated by their technical function, are anyway excluded from design protection[64].

"snacks items" is evidence of the broad possibilities open to the designer and, at the same time, the limits thereof. The designer freedom is not limitless, since the overcrowding of the market sector and industrial feasibility of the goods item determine much more constraint on a competing company operating in the same market sector. Accordingly, the designer's degree of freedom was found to be average, rather than broad or limitless, and implying a gradual decline in the shapes that are still available.

61 It is however likely that in the future it will be possible to produce 3D printed products in larger sizes.

62 At 5.4.6.2.

63 See Article 4(2)(a) CDR.

64 Both the functionality and the must-fit exclusions do not

103 One of the most problematic issues the EU legislators had to face concerns the so-called "must-match" exclusion[65]. This exclusion deals with the visual synchronisation and aesthetic appearance of a complex product, rather than with functionality. In other terms, the must-match provision concerns the design of a component part, which should be used for the purpose of the repair of a complex product so as to restore its original appearance (e.g. the design of a car body panel that is used to restore the original appearance of the vehicle).

104 The protection of must-match spare parts has occasioned the greatest controversy among a wide range of stakeholders, especially in the automotive industry. The following analysis provides a brief overview of the legislative history on this issue. This will help explain why the dispute is not resolved yet.

105 The original idea in the 1993 proposals for a Regulation on the Community design[66], and for a Directive on the legal protection of designs[67], was to introduce a must-match exception in Europe, specifying that only after a period of three years, from the first placing on the market of a complex product, the rights conferred by a RCD could not be exercised to prevent third parties from using the design of a component part, in order to restore the original appearance, or to permit the repair of, the complex product. The Council of Ministers rejected this option.

106 The European Parliament advanced a different solution in the Amended Proposal for the Design Directive, opting for a compulsory licensing regime that allowed the use of component parts, for repair purposes, immediately after the placing on the market of the complex product, in exchange for a fair and reasonable remuneration of the right holder[68]. Manufacturers of component parts were required to inform the public as to the origin of their products used for the repair by means of a trademark or trade name. They also had to notify the right holder of the intended use of the design, and regularly inform her as to the scale of such use. Nonetheless, no agreement on the compulsory licensing clause was reached by the European Council.

107 Ultimately, the disagreement between EU institutions was the subject of a Conciliation Committee meeting, where the Council insisted on its position against a remuneration scheme. It recommended, instead, an extension of the period of exclusivity over component parts for a period ranging from three to seven years.

108 In such a tense context, the European Union opted for the so-called "*freeze plus*" solution, stating that until amendments to the Directive are adopted on a proposal from the Commission on this subject, Member States shall maintain in force their existing legal provisions. Member States should only change their laws if they wished to liberalise their market for spare parts, pursuant to Article 14 Directive 98/71/EC. Therefore, Member States had alternative options: they could introduce a clause allowing any use of the design for repair purposes; adopt a remuneration system; provide a term-limited design protection; or craft their own exception, which is a combination of the second and third options.

109 Article 110 CDR codified another "freeze plus" or transitional provision, mirroring the one set forth in the Directive. Thus, in 2004 the Commission made its third attempt to achieve harmonisation in this convoluted area, issuing a proposal designed to liberalise the aftermarket for spare parts[69]. This proposal, known as the "repairs clause", purported to increase legal certainty and allow market operators and consumers to take full advantage of a uniform Internal Market for spare parts[70].

110 In fact, the situation at that time was characterised by opposed regimes, where nine Member States, including Italy and the UK, have liberalised, whereas sixteen Member States had *de jure* design protection to spare parts (among them, Austria, Denmark, Finland, Germany, Portugal, Sweden). The European Commission found that the *status quo* – with mixed protection regimes of design protection for spare parts – was altogether unsatisfactory and created trade distortion in the Internal Market[71]. The non-

apply to design features which allow the multiple assembly or connection of mutually interchangeable products within a modular system (Recital 11, Article 8(3) CDR). Hence, design subsists in interconnection features of construction toys or modular furniture. Cornish et al., in *Intellectual Property: Patents, Copyright, Trade marks and Allied rights*, 7th edn Sweet & Maxwell (2007), p. 613, maintain that the special treatment offered to toy manufacturers has no reasonable explanation, except that it shows how determined lobbying can squeeze special concessions into legislation.

65 The "must-match" terminology comes from the UK legislation on UK Unregistered Design Rights. Such exception was first introduced within the UK Community Designs and Patents Act 1988.

66 Proposal for a European Parliament and Council Regulation on the Community Design, COM (93) 342 final-COD 463, 3 December 1993, Article 23 of the Draft Regulation.

67 Proposal for a European Parliament and Council Directive on the Legal Protection of Designs, COM (93) 344 final-COD 464, 3 December 1993, Article 14.

68 Amended Proposal Design Directive [1996] OJ C1 42/7, Article 14.

69 European Commission (2004), Proposal for a Directive of the European Parliament and of the Council amending Directive 98/71/EC on the Legal Protection of Designs: Extended Impact Assessment.

70 *Id.*, at 2.

71 *Id.*, at 1.1.1.

harmonisation of the must-match exclusion means that independent manufacturers are only able to sell their products and offer their services in some Member States but not in others.

111 Following a lack of progress at Council level, in May 2014 the proposal was withdrawn. Successively, the Commission launched a comprehensive legal and economic evaluation of the overall functioning of EU design systems[72]. In the framework of this evaluation, an external contractor, Europe Economics, presented "The Economic Review of Industrial Design in Europe"[73]. The latter suggests that, among various policy options, full liberalisation, meaning a complete elimination of design protection for spare parts within the EU, would be the best outcome. In an age of widespread availability of 3D printers, consumers and independent manufacturers think that they are entitled to produce their own 3D-printed spare parts for the purpose of repair. Hence, a *de facto* repairs clause might become inevitable anyway. Insofar as it is impossible to enforce design law against all infringers in the 3DP landscape, a full liberalisation has to take place.

112 In response to this argument, one could maintain that 3DP makes the introduction of a repairs clause a more delicate issue than it was in the relatively recent past, because 3D printed products might not meet quality and safety standards. Any proposal for full liberalisation should foresee a method to ensure that component parts are safe and useable, when it becomes possible for different industries to manufacture spare parts using 3DP.

113 In a study commissioned by the UK Intellectual Property Office ("IPO"), Reeves and Mendis stressed, in this regard, that it is rather unrealistic to assume that 3DP will be heavily used, in the near future, to make component parts in certain industrial sectors, such as the automotive aftermarket[74]. The component parts that, according to the UK IPO's study, are not yet suited to additive manufacture include: tyres, batteries, oil filters, air conditioning, etc. There are also aftermarket parts whose manufacture is technically possible by means of 3DP, but not economically viable yet, since the production costs would be higher than the current aftermarket value. The latter include: exhaust pipes, distributor caps, water pumps, and radiators[75].

114 As noted by the authors, one of the biggest limitations to the production of 3D printed spare parts lies in the lack of credible design data from which to print. In the Office's opinion, it is erroneously "assumed that parts can be simply scanned and reverse engineered, with the resulting data then being stored on the cloud" for downstream 3DP. It is of fundamental importance to have access to the original CAD files, to understand "issues such as tolerances, loading conditions and material requirements".

115 Hence, whether the impact of 3DP on the liberalisation of the aftermarket sector will be significant in the next future is not altogether clear yet[76].

116 For the sake of completeness, it is also worth recalling a recent Order from the CJEU in *Ford Motor Company v Wheeltrims s.r.l.*[77], dealing with trademark law.

117 At first instance, in the Italian proceedings, the claimant *Ford Motor Company* claimed that the defendant, a company operating in the automotive aftermarket, had infringed its registered trademark "Ford". *Wheeltrims* was marketing wheel caps bearing the registered trademarks of the original manufacturers – including Ford's trademark – without the owners' authorisation. The defendant raised the repairs clause defence, arguing that Article 241 of the Italian Industrial Property Code, implementing Article 14 of the Design Directive, should apply as a defence to trade mark infringement. The use of the trademark "Ford" was justified for the purpose of restoring the original appearance of the complex product, in derogation of the Trade Mark Regulation (EC) 207/2009 and Trade Mark Directive 84/104/EC. The Tribunale Ordinario di Torino made a reference for a preliminary ruling to the CJEU on the interpretation of the repairs clause set forth in the DD and CDR.

118 The CJEU answered the referred questions by Order, stating that Article 14 of DD and Article 110 CDR must be interpreted as not allowing – by way of derogation from the provisions of the Trade Mark Directive 2008/95/EC and Trade Mark Regulation 2009/207/EC– a manufacturer of replacement parts and accessories for motor vehicles to affix to its products a sign, which is identical to a trademark registered for such products by the original manufacturer, without the latter's authorisation, on the ground that the use thus made of the trade mark is the only way to restore the original appearance of the complex product.

119 Hence, the CJEU has made clear that, in its current

72 *See <http://ec.europa.eu/growth/industry/intellectual-property/industrial-design/protection/index_en.htm>.*

73 Europe Economics, *The Economic Review of Industrial Designs in Europe*, a study commissioned by DG Internal Market and Services (January 2015).

74 Reeves, Mendis, report commissioned by the UK IPO, *The Current Status and Impact of 3D Printing Within the Industrial Sector: An Analysis of Six Case Studies* (March 2015), p. 19.

75 *Id.*, p. 17.

76 According to Reeves and Mendis it will not be significant for the next 10 years. *Id.*, p. 20.

77 Order of the CJEU (Third Chamber) of 6 October 2015, Case C-500/14.

form, the repairs clause that is anchored in European design law does provide a defence to an alleged trademark infringement. As a result, a third party who replicates by means of 3DP a component part, to which the manufacturer's own trademark is affixed, may be found liable for trademark infringement, provided that the private use exception does not apply[78].

G. Exclusive rights conferred by a design

120 In the event that a design is registered, the holder of a RCD is granted an exclusive right to use it and to prevent any third party not having her consent from using it. Pursuant to Article 19 CDR, the right to use the design covers different sorts of activities, such as the making, offering, putting on the market or using of a "product" in which the design is incorporated. In contrast, an UCD confers the right to prevent the same aforementioned activities, but only insofar as the contested use results from copying the protected design, and is not the result of an independent work of creation.

121 The owner's exclusive rights extend towards any third party, without any differentiation between primary and secondary infringers. This, in turn, implies that the holder of a RCD can pursue claims for direct infringement against intermediaries (e.g. online 3D platforms).

122 Furthermore, as already mentioned, infringement is not confined to the use of the design on the same product, in which the design was incorporated in the first place. Protection extends toward any use of the design, in relation to any products. It is also worth remembering that infringement cannot occur with respect to acts done privately and for non-commercial purposes[79], and acts done for experimental purposes (Article 20 CDR).

123 In light of the above considerations, the question of whether 3D printing a design-protected product from a CAD file constitutes or not an infringing activity is straightforward. There is no doubt that the acts prohibited under Article 19 CDR will encompass the manufacture of objects via 3DP, that is done in the context of a commercial activity and outside the scope of the private use exception (i.e. "making" the design)[80]. Infringement will not be actionable, instead, against an individual, who 3D prints a design product at home, for private and personal use.

124 Moreover, the fabrication of products, by means of 3DP, done for scientific research will be exempted, irrespective of whether it is for a private or commercial purpose[81]. As noted by Suthersanen, this exception should be interpreted narrowly and be only allowed if the experimental usage of the design is in the general interest. A demarcation should always be made between acts of experimental nature, and those that seek to exploit the design[82].

125 Whether the scope of design protection should also include the act of making a scanned representation and/or a CAD file from a design already existing as a tangible article is less clear-cut. Also, does the unauthorized act of copying and marketing a third party's CAD file, in which a design is incorporated, amount to infringement of the design right?

126 The unsolved issue, therefore, is whether activities carried out in relation to CAD files fall foul of Article 19 CDR, and constitute an illegitimate "use" of the design. Moreover, who is the party responsible for the infringement: the one who uploads, downloads or markets the CAD file? Should the host of the file-sharing site be held liable too?

127 A strict interpretation of the law would suggest that the answer to these questions should be no. Just as a design requires there to be a product, infringement should only occur where a person uses a physical product[83]. The latter should not necessarily be the same product to which the design was incorporated in the first place, but it should however be an industrial or handicraft item.

128 Furthermore, the CDR does not provide protection against indirect use of a design, differently from patent law. There is no specific provision that confers on the holder of a Community design the right to prevent third parties, not having her consent, from supplying the "means" for using the design (e.g. marketing a complete kit that, when made up, constitutes the design)[84]. A CAD file could be seen as a "means" enabling the fabrication of the product in which the design is incorporated. As a consequence,

78 According to Article 10 of Directive 2015/2436/EC, in order to commit an infringement the use of a third party's trademark should be "in the course of trade", i.e. in the context of a commercial activity with a view to economic advantage and not as a private matter. *See* the CJEU's ruling in *Arsenal Football Club plc. v Reed*, C-206/01 [2002] ECR I-10273, [40].

79 These criteria are cumulative. Use should be both private and for purposes that are not commercial.

80 *See* Malaquias, *The 3D Printing Revolution: an Intellectual Property Analysis* (8 August 2014), available at: <http://papers.ssrn.com/sol3/Papers.cfm?abstract_id=2495416>.

81 Article 20(1)(b) CDR.

82 Suthersanen (2010), *supra* note 20, p. 140.

83 Bently, Sherman, *Intellectual Property Law*, 3rd ed., Oxford: OUP (2008), p. 666.

84 *See* Article 30 of the Convention for the European Patent for the Common Market (Community Patent Convention) 76/76/EEC.

making and distributing a CAD file would constitute an authorized (indirect) use of the design[85].

129 Therefore, a crucial issue to address is the extent to which a design right can be used against a new form of exploitation that does not imply the making of physical objects, but the creation and sharing of digital files.

130 It is here suggested that confining the scope of design protection to use on material products only is overly restrictive, in the light of the current technological change brought about by 3DP. This technology is blurring the line between the physical and the immaterial worlds. An increasing number of undertakings might decide to make their CAD files available online in the course of their business. Digital networks might emerge as an ordinary means of distributing 3DP templates of protected designs. In this way, undertakings would not need to mass-produce or distribute their products any longer.

131 Once a design is made available in the form of a CAD file, it then becomes extremely easy for anyone to replicate it, either by entrusting a third party with the task of printing the product, or by using personal hardware. Future advancements in personal 3D printers will further expand this capability. Hence, design-based industries have to be equipped for the digitalization of things. In order for alternative business practices to come to light, it is of utmost importance to ensure that material protected by an IP is respected.

132 This, in turn, calls for a reinterpretation of the legal basis on which right holders shall receive protection. Arguably, they should be exclusively entitled to use – and prevent third parties from using and dealing with – the CAD files of their protected designs.

133 There are several ways to address this issue. A first option would be to consider the digital representation of a design as a "product" within the meaning of Article 19 CDR. Accordingly, this provision would cover different activities, such as the unauthorized making of CAD files (i.e. the making of a product), sharing of CAD files with other Internet users (i.e. use of the design), and the sale of CAD files on 3DP marketplaces (i.e. offering the product and putting it on the market). Furthermore, a possible interpretation of Article 20 CDR would be that the private use exception exempts from liability a third party who simply downloads a CAD file and saves it on her computer.

134 One may support this conclusion arguing that requiring products to have some physical form would be unduly limiting[86]. From a systemic perspective, it seems rather contradictory to allow registration of graphic symbols – including computer icons – and, at the same time, postulate that the notion of "product" is tied to a physical dimension for infringing purposes.

135 Furthermore, the scope of design protection is not limited to a certain category of products; rather, it covers any use of the design, in relation to "any" product that does not produce on the informed user a different overall impression. As noted by Malaquias, it seems very difficult to ascertain that a CAD file "will produce on the informed user a different overall impression from the protected design, considering that its purpose is to replicate the existing design in three-dimensions"[87].

136 In the opinion of the present writer, the preferable solution is to follow the recommendation, made by the European Commission in the "Legal Review on Industrial Design Protection in Europe"[88], to introduce an infringement provision stating that the creation of a design document amounts to an infringing use[89].

137 As suggested by the European Commission, a template for such provision may be Section 226(1) of the UK CDPA 1988, which states that "the owner of a design has the exclusive right to reproduce the design for commercial purposes [...] by making a design document recording the design for the purposes of enabling such articles to be made".

138 In the UK jurisdiction, "design document" is defined in Section 51(3) CDPA as: "any record of a design, whether in the form of a drawing, a written description, a photograph, data stored in a computer or otherwise". The scope of this provision is wide enough to include CAD files as design documents[90].

139 Furthermore, EU design law could fashion an additional provision similar to Section 226(3) CDPA 1988, specifying that it is a primary infringement of a design right to do or "authorise" another to do, without the design right owner's permission,

85 Bently, Sherman (2008), *supra* note 83, p. 666.

86 This expression is used by Bently and Sherman, *id*., p. 667, footnote 66.

87 Malaquias (2014), *supra* note 80, p. 27.

88 MARKTD2014/083/D.

89 *Id*. 133.

90 It should however be borne in mind that in the UK jurisdiction, pursuant to Section 51(1) CDPA 1988, copyright in a design document (i.e. in the CAD file) will not infringed by making a 3D article from it, where the design is for anything other than an artistic work or a typeface. Hence, if a CAD file embodies a utilitarian design (for example, the design of automotive spare parts), printing the object will not result in copyright liability. In this respect, the UK model would not be a good model to replicate for the EU.

anything which is the exclusive right of the design right owner.

140 In the first place, this provision would clarify that making a CAD file from an existing design-protected product, for the purposes of 3D printing such product, amounts to an infringement of the design right. This provision would also specify that intermediary parties (such as 3DP online platforms) might also be directly liable for "authorising" design infringement. As stressed in the Commission's review, the benefit of such a provision is that neither actual nor constructing knowledge would be required for a positive finding of infringement[91].

H. Conclusions

141 A clear message emerges from the arguments developed in this paper. European design law should adapt to the reality of digitized goods and accommodate greater protection for right owners.

142 To date, the EUIPO accepts 3D digital representation of designs as "representations of the design that are suitable for reproduction", within the meaning of Article 36 CDR. Such a representation is enclosed in the application form for a RCD to show, in the same way as a photograph, the design for which protection is sought.

143 It has also been noted that, although the CDR is structured on the concept of "product", the EUIPO does not take into consideration whether a product is actually made or used, or can be made or used, in an industrial or handicraft fashion. This, means that, in theory, the CAD representations included in the application for a RCD will determine the scope of design protection, regardless of whether the product is actually manufactured or not.

144 It has also been contented that in case a CAD file clearly unveils the outer appearance of a product, its publication online will be tantamount to a "disclosure" for the purposes of Article 7 CDR. As a consequence, all later products – and CAD files for products – will have to produce a different overall impression on the informed user. By contrast, it is not entirely clear whether publishing a CAD file on a website that is hosted outside the EU will trigger UCD protection from the date of the first online publication, given the geographical limitation contained in Articles 11 and 110(a) 5 CDR.

145 Hence, there are many issues that have to be clarified. First, who is the informed user of a product represented digitally as a CAD file, as opposed to the informed user of the finished product? Second, is the designer's degree of freedom enhanced or limited by the fact that she creates a product design using CAD software? Third, if many individuals begin to create their own CAD files for products and upload them online, thereby disclosing the design for which protection is sought, will many market sectors suddenly become overcrowded? Will all subsequent designs have to depart from the considerable amount of CAD models already made available online?

146 Besides, the ease of converting a CAD file into a physical item leads us to suggest that design owners should be entitled to claim protection for the CAD representations of their designs. In a hypothetical world of widespread 3D printers, it could be that CAD files become almost interchangeable with end products. The owner of a CAD file might be as satisfied as if she possesses the end product itself. A CAD file would then serve as a substitute for a good, offered to the same or actual potential customers.

147 Many are the fields in which clear-cut rules are needed, since new technologies empower the individual in her creativity and yet should make her responsible for potential infringement of third parties' exclusive rights.

148 In this respect, the present writer supports the following recommendations, made by the European Commission in its recent report "Legal Review on Industrial Design Protection in Europe": first, to introduce a provision that confers upon the design right owner an exclusive right to make a design document, which is a record of the design (i.e. a CAD file); second, to introduce a provision on direct primary infringement by authorisation.

91 As an alternative remedy, the European Commission proposes to introduce "indirect design infringement" as a separate head of liability. As noted above, a CAD file may be seen as a "means" that enables the actual infringement of the design right, i.e. as an "indirect" use of a design. In addition, the European Commission focuses on the possibility to review the private and non-commercial use exception. One way to restrict the scope of this exception is to employ the 3-step language adopted in Article 26 of the TRIPs Agreement ("*provided that such exceptions do not unreasonably conflict with the normal exploitation of protected industrial designs, and do not unreasonably prejudice the legitimate interests of the owner of the protected design, taking account of the legitimate interests of third parties*"), in order to provide greater flexibility and achieve a balance between the legitimate interests involved. The latter recommendation does not seem advisable. The language employed in the three-step test may lead to ambiguity and to a non-uniform interpretation. Rather than representing a useful tool, it may create additional confusion. Cf. 6.1 of the report (MARKTD2014/083/D).

* *Viola Elam* is a Ph.D. Researcher at the European University Institute, Fiesole, Italy. She is very grateful to an anonymous referee for insightful comments.

Personal Data and Encryption in the European General Data Protection Regulation

by **Gerald Spindler and Philipp Schmechel***

Abstract: Encryption of personal data is widely regarded as a privacy preserving technology which could potentially play a key role for the compliance of innovative IT technology within the European data protection law framework. Therefore, in this paper, we examine the new EU General Data Protection Regulation's relevant provisions regarding encryption – such as those for anonymisation and pseudonymisation – and assess whether encryption can serve as an anonymisation technique, which can lead to the non-applicability of the GDPR. However, the provisions of the GDPR regarding the material scope of the Regulation still leave space for legal uncertainty when determining whether a data subject is identifiable or not. Therefore, we inter alia assess the Opinion of the Advocate General of the European Court of Justice (ECJ) regarding a preliminary ruling on the interpretation of the dispute concerning whether a dynamic IP address can be considered as personal data, which may put an end to the dispute whether an absolute or a relative approach has to be used for the assessment of the identifiability of data subjects. Furthermore, we outline the issue of whether the anonymisation process itself constitutes a further processing of personal data which needs to have a legal basis in the GDPR. Finally, we give an overview of relevant encryption techniques and examine their impact upon the GDPR's material scope.

Keywords: GDPR; Encryption; Anonymisation; Pseudonymisation; Personal Data; Material Scope; ECJ; Advocate General; Data Protection; Secure Multiparty Computation

Recommended citation: Gerald Spindler and Philipp Schmechel, Personal Data and Encryption in the European General Data Protection Regulation, 7 (2016) JIPITEC 163 para 1.

A. Introduction

1 Seventeen years ago, *Lawrence Lessig* wrote that "encryption technologies are the most important technological breakthrough in the last one thousand years".[1] This might be a slight exaggeration, but it emphasises the importance of encryption technologies in today's digital world. Encrypted data plays a significant role in the protection of data subjects' privacy. Its legal problems are closely related to the scope of the data protection laws and the legal effects of anonymisation and pseudonymisation.

2 Encrypting personal data is becoming increasingly important for many business models in a data-driven economy and for preserving data subjects' privacy with regard to today's monitoring and profiling possibilities – both of government institutions and of high-tech companies. Be it for the processing of sensitive health data, for the *Internet of Things* or for connected cars, for the privacy preserving use of Big Data or cloud computing technologies[2], encryption can be a key to protect an individual's privacy and

1 *Lessig*, Code and Other Laws of Cyberspace, 1999, p. 35.

2 See e.g. the PRACTICE project, funded by the EU-FP7-programme, which aims to build a secure cloud framework that allows for the realization of advanced and practical cryptographic technologies, <https://practice-project.eu/>.

can make several IT innovations possible, which would otherwise conflict with the data protection framework. For many years, the discussion about the material scope of the European Data Protection Directive[3] (DPD) and about the exact definition of personal data and the interpretation of the term "identifiable" has been one of the "key issues"[4] of European data protection law.[5] Additionally, the legal effects of encrypted data for the applicability of data protection law and for the personal references of data have still not sufficiently been examined. These questions regarding personal data and encryption once again occur in the new EU General Data Protection Regulation[6] (GDPR).

3 Encrypting personal data and deleting any personal reference from the data could also be a way to work with this information when it is transferred to third countries outside of the EU.[7] EU standard contractual clauses or compliance to the new Privacy Shield when transferring data to the U.S. would therefore not be necessary if the data lost all of its personal reference. However, legal uncertainty concerning whether the encryption of personal data has the effect that such data loses its personal reference or not may discourage controllers to use these privacy preserving measures. Thus, in this article we will examine the legal effects of encryption in regards to the applicability of the GDPR.[8]

4 The GDPR only applies if "personal data" is processed. Thus, the notion of *personal* data is crucial for the application of the GDPR. Depending on how "personal data" is defined and interpreted, the effect a valid encryption of this data takes may be different. Furthermore, we will examine how encrypted data is treated in the GDPR – as anonymised or pseudonymised data – and where and how in the GDPR encryption can be used as a technical and organisational measure.[9] With regard to some important encryption tools for the transport, storage and processing of personal data we will demonstrate the effect of encryption on the material scope of the GDPR.[10]

B. The General Data Protection Regulation and Encryption

5 After years of intensive negotiations, the GDPR has now been passed and will finally come into force from 25 May 2018 (see Article 99 Par. 2 GDPR) and will, according to Article 94 Par. 1 GDPR, repeal the old Directive 95/46/EC.[11] Due to its legal form of as a Regulation, the GDPR will be binding in its entirety and will be directly applicable in all Member States of the European Union.[12] We will examine the material scope of the GDPR and the effect of encryption on personal data.

6 The importance for controllers of knowing the exact scope of the Regulation and whether the data they process will be considered as personal data or not, e.g. due to the use of encryption, increases with the GDPR's very broad territorial scope, especially the rules for controllers not established in the EU will be changed dramatically.[13] According to Article 3 Par. 1 the GDPR "applies to the processing of personal data in the context of the activities of an establishment of a controller or a processor in the Union, regardless of whether the processing takes place in the Union or not". Moreover, Par. 2 states that the Regulation applies even to the processing of personal data where the processing activities are related to the offering of goods or services or the sheer monitoring of the data subject's behaviour as long as their behaviour takes place within the Union. "Monitoring" means *inter alia* the online tracking of natural persons to create profiles in order to take decisions, for analysing or predicting personal preferences, behaviours and attitudes (see Recital 24 S. 2 GDPR).[14]

3 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, Official Journal of the European Communities, L 281, pp. 31-50.

4 *Boehme-Neßler*, Datenschutz und Datensicherheit 2016, p. 419.

5 See e.g. *Article 29 Data Protection Working Party*, Opinion 4/2007 on the concept of personal data, WP 136, pp. 6 et seq.; *Hon/Millard/Walden*, Queen Mary University of London – Legal Studies Research Paper No. 75/2011, pp. 8 et seq., available at: <http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1783577##>, accessed 26 August 2016; *Article 29 Data Protection Working Party*, Opinion 05/2014 on Anonymisation Techniques, WP 216, pp. 5 et seq.

6 Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation), Official Journal of the European Union, L 119, pp. 1-88.

7 Cf. *Esayas*, European Journal of Law and Technology, Vol 6, No 2 (2015), p. 13; Commission Decision 2000/520/EC of 26 July 2000, Official Journal of the European Communities L 215/7 (24).

8 See *infra* B.II.2.d.).

9 See *infra* B.II.1.

10 See *infra* B.II.3.

11 See for an overview of the legislative process of the GDPR *Albrecht*, Computer Law Review International 2016, pp. 33 et seq.

12 *Reding*, International Data Privacy Law 2012, p. 119 (121).

13 See *Kindt*, CiTiP Working Paper 26/2016, pp. 13 et seq., available at: <http://papers.ssrn.com/sol3/JELJOUR_Results.cfm?form_name=journalbrowse&journal_id=1781425>, accessed 8 August 2016.

14 Under the DPD, according to Article 4 Par. 1 (c) controllers targeting EU data subjects only had to comply with the DPD if they made use of "equipment" situated in the EU to process personal data.

7 Thus, the GDPR's broad territorial scope leads towards a new awareness of data controllers (also established outside the Union) regarding their processing of personal data. Therefore, technologies which minimise the use of personal data – especially encryption – and which avoid the application of the GDPR become even more important.

I. Personal Data: The Material Scope of the GDPR

8 As already outlined, the characteristics of personal data are crucial for the application of the GDPR. However, the GDPR does not introduce major changes to the concept of personal data in comparison to the DPD. Just like the DPD, the GDPR follows a "black/white approach", hence the data are either personal or not, which means that if the data has a personal reference, all data protection rules apply and if not, it is outside the GDPR's scope.[15] According to Article 2 Par. 1 GDPR

> *"This Regulation applies to the processing of personal data wholly or partly by automated means and to the processing other than by automated means of personal data which form part of a filing system or are intended to form part of a filing system."*

9 Article 4 No. 1 S. 1 GDPR defines that "'personal data' means any information relating to an identified or identifiable natural person[16] ('data subject')" which is the same wording as Article 2 (a) DPD. In this regard, "any information" means virtually *any* information, even publicly available information; when a reference to a natural person can be made the data protection principle of the GDPR always applies regardless of the data's content.[17] However, Article 4 No. 1 S. 2 GDPR introduces a new definition of the concept of an "identifiable natural person", which refers to a person "who can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, an online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person". Thus, the definition of an "identifiable natural person" distinguishes between identifiability on the basis of a reference to an identifier which can clearly identify a natural person, or due to special personal characteristics such as a person's sexual preferences or medical condition.[18]

10 However, it is still highly controversial whether or not a so-called *absolute* or *relative* approach has to be applied for assessing the data controller`s abilities to identify a natural person.

1. The "Identifiable Natural Person"

11 Crucial to understanding the exact scope of the concept of "personal data" is how much a potential data controller has to do in order to establish a link between a natural person and the data, in other words what efforts are required to identify a person.

a.) Absolute Approach

12 The *absolute* approach takes into account all possibilities and chances in which the data controller would be able to identify the data subject individually. Thus, all ways and means for a data controller without any regard to expenses etc. are taken into account. Even theoretical chances of combining data so that the individual is identifiable are included. If identifiability is assessed *absolutely*, then it is sufficient for the application of personal data acts if *anyone* in the world is able to decrypt or decode the encrypted data.[19]

13 In terms of encryption, as long as *anyone* in the world is able to decrypt the data set, the operations of the controller or processor using this encrypted data are subject to data protection legislation, even if they don't possess the key for decryption. Based on this approach data protection legislation is applicable, regardless of the applied encryption technique, as long as one entity holds the key for decoding.[20]

b.) Relative Approach

14 In contrast, the *relative* approach considers the necessary effort required by the data controller

15 *Forgó*, International Data Privacy Law 2015, p. 54 (59).

16 Like in Article 1 Par. 1 of the DPD, the material scope of the GDPR only applies to the processing of personal data of *natural* persons according to Article 1 Par. 1 GDPR.

17 Cf. *Kranenborg*, in: Peers/Hervey/Kenner/Ward (eds.), The EU Charter of Fundamental Rights, 2014, Art 8, Recital 08.85; *Article 29 Data Protection Working Party*, WP 136 (*supra* Note 5), pp. 6 et seq; *Karg*, Datenschutz und Datensicherheit 2015, p. 520 (521).

18 Cf. *Härting*, Datenschutz-Grundverordnung, 2016, Recital 275 et seq.

19 *Kuner*, European Data Protection Law: Corporate Compliance and Regulation, 2nd Ed. 2007, p. 92; *Pahlen-Brandt*, Datenschutz und Datensicherheit 2008, p. 34 (38); *Nink/Pohle*, Multimedia und Recht 2015, p. 563 (565), who criticize that consequently this approach would lead to the result that there would virtually be no more anonymous data.

20 Cf. *Meyerdierks*, Multimedia und Recht 2009, p. 8 (10).

in order to identify the data subject.[21] Therefore, only realistic chances of combining data in order to identify an individual are taken into account – and not highly theoretical identification risks.[22] With regards to encryption issues, data protection legislation is only applicable if the data *controller* is able to decrypt a certain data set[23] – or, at least has reasonable chances of obtaining the decrypting key. In the case law of some courts, the trend is beginning to lean towards favouring a *relative* understanding.[24]

c.) The GDPR's Approach

15 The GDPR utilises a broad approach regarding the interpretation of "identifiable natural person" however, some terms can also be interpreted in a *relative* way. Additionally, both Article 7 and Article 8 of the Charter of Fundamental Rights of the EU (CFR) always have to be taken into account when interpreting the data subject`s rights.[25]

16 Recital 26 S. 3 of the GDPR states that "to determine whether a natural person is identifiable, account should be taken of all the means reasonably likely to be used, such as singling out, either by the controller or by another person to identify the natural person directly or indirectly." On the one hand, the Recital refers to means reasonably likely to be used "by another person" which have to be taken into account, which veers towards an *absolute* approach, because this *third person* could be any person in the world.[26] This is also in tune with the scope of Article 8 CFR, according to which "identifiable" has to be interpreted widely.[27]

17 Moreover, stating in Article 4 No 1 S. 2 GDPR that every "identifier" shall contain personal references is another hint for a rather *absolute* approach of the Regulation regarding the identifiability of a natural person.[28] Additionally, Recital 26 states that using means for "singling out" the natural person directly or indirectly may make this person identifiable. Thus, a data subject may now be singled out for data processing even if it is unlikely that his or her name can be tied to the data, because even this could result in harming his or her privacy.[29]

18 On the other hand, the term "means reasonably likely to be used" suggests limitations through *relative* elements, in particular the notion of "reasonably".[30] Additionally, if a zero risk threshold would be applied for any potential data user, no existing technique could achieve the required level of anonymisation.[31] Moreover, according to the *Article 29 Data Protection Working Party* (interpreting the Data Protection Directive), "a mere hypothetical possibility to single out the individual is not enough to consider the person as 'identifiable'".[32]

19 Recital 26 GDPR continues by stating objective factors which shall be relevant for the interpretation of the means used to identify a natural person:

> *"To ascertain whether means are reasonably likely to be used to identify the natural person, account should be taken of all objective factors, such as the costs of and the amount of time required for identification, taking into consideration the available technology at the time of the processing and technological developments."*

20 These factors illustrate a further attempt to limit the broad *absolute* elements of the GDPR's material

21 *Roßnagel/Scholz*, Multimedia und Recht 2000, p. 721 (723); *Meyerdierks* (*supra* Note 20), pp. 8 et seq.; *Voigt*, Multimedia und Recht 2009, p. 377 (379); *Lundevall-Unger/Tranvik*, International Journal of Law and Information Technology 2010, p. 53 (58); *Hon/Millard/Walden* (*supra* Note 5), p. 14.

22 *Esayas* (supra Note 7), p. 6.

23 Cf. *Spindler*, Verhandlungen des 69. Deutschen Juristentages, Band I, Gutachten, 2012, pp. 115 et seq.

24 England and Wales High Court (Administrative Court), [2011] EWHC 1430 (Admin), Case No. CO/12544/2009, Recital 51 f.; Upper Tribunal (Administrative Appeals Chamber), [2011] UKUT 153 (AAC), Appeal Number: GI/150/2011, GI/151/2011, gi/152/2011, Recital 128; House of Lords, [2008] UKHL 47, recital 27; The Paris Appeal Court, decision of 15 May 2007 – *Henri S. vs. SCPP*; Local Court of Munich, decision of 30 September 2008 – 133 C 5677/08, Recital 26; District Court of Wuppertal, decision of 19 October 2010 – 25 Qs 10 Js 1977/08-177/10; District Court of Berlin, decision of 31 January 2013 – 57 S 87/08; different point of view: The Stockholm Lænsrætt, reference No. 593-2005, publication date 8 June 2005; Local Court of Berlin-Mitte, decision of 27 March 2007 – 5 C 314/06, Recital 20; Administrative Court of Wiesbaden, decision of 27 February 2009 – 6 K 1045/08.WI, Recitals 52 et seq.

25 Cf. *Vedsted-Hansen*, in: Peers/Hervey/Kenner/Ward (eds.) (*supra* Note 17), Art 7, Recital 07.72A.

26 Cf. *Zuiderveen Borgesius*, Computer Law & Security Review 2016, p. 256 (267) who interprets Recital 26 as "an absolute approach to identifiability"; *Polonetsky/Tene/Finch*, Santa Clara Law Review, (Forthcoming) 2016, p. 593 (614).

27 Cf. *Kranenborg*, in: Peers/Hervey/Kenner/Ward (eds.) (*supra* Note 17), Art 8, Recital 08.85.

28 *Brink/Eckhardt*, Zeitschrift für Datenschutz 2015, p. 205 (208); *Buchner*, Datenschutz und Datensicherheit 2016, p. 155 et seq.; *Härting* (*supra* Note 18), Recital 279.

29 *Hon/Kosta/Millard/Stefanatou,* Tilburg Law School Legal Studies Research Paper Series No. 07/2014, p. 9, available at: <http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2405971>, accessed 15 August 2016; *Zuiderveen Borgesius* (*supra* Note 26), p. 256 (267); *Marnau*, Datenschutz und Datensicherheit 2016, p. 428 (430).

30 Cf. *Esayas* (*supra* Note 7), p. 6; *Härting* (*supra* Note 18), Recital 282.

31 *Esayas* (*supra* Note 7), p. 6; regarding "anonymisation of personal data in the GDPR" see *infra* B.II.2.d.).

32 *Article 29 Data Protection Working Party*, WP 136 (*supra* Note 5), p. 15; *Article 29 Data Protection Working Party*, WP 216 (*supra* Note 5), pp. 8 et seq.

scope.[33] Significant objective factors will be *inter alia* the state of science and technology, including future technological developments as well as the time and costs needed to identify somebody.[34]

21 In October 2014, the German Federal Court of Justice (BGH) requested the European Court of Justice (ECJ)[35] for a preliminary ruling on the interpretation of the dispute regarding whether a dynamic IP address can be considered as personal data,[36] in particular if the relevant additional information is held by a third party, such as an internet service provider. The ECJ will most likely resolve the dispute between an *absolute* or *relative* approach regarding dynamic IP-addresses by interpreting Article 2 (a) DPD and especially recital 26 of the DPD.[37] Since Article 2 (a) DPD and Article 4 No. 1 GDPR are very similar, the ECJ's decision will certainly also have a major influence on the general interpretation of defining "identifiability" in the GDPR.[38] On 12 May 2016 the *Advocate General (AG)*, *Campos Sánchez-Bordona* published his opinion regarding this case, however, whilst the ECJ is not bound to follow his opinion, it often does so.[39]

22 In his opinion, the *AG* contradicts an interpretation of "means likely reasonably to be used ... by any other person" in such a way that it would be sufficient that *any* third party might obtain additional data in order to identify a person[40], since this "overly strict interpretation would lead, in practice, to the classification as personal data of all kinds of information, no matter how insufficient it is in itself to facilitate the identification of a user".[41] Moreover, the *AG* emphasises that otherwise "it would never be possible to rule out, with absolute certainty, the possibility that there is no third party in possession of additional data which may be combined with that information".[42] This can be interpreted as a tendency of the *AG* towards a *relative* approach. Furthermore, according to the *AG*, "(j)ust as recital 26 refers not to any means which may be used by the controller (...), but only to those that it is likely 'reasonably' to use, the legislature must also be understood as referring to 'third parties' who, also in *a reasonable manner*, may be approached by a controller seeking to obtain additional data for the purpose of identification".[43] The *AG* concludes that contacting third parties shall not be reasonable when it is "very costly in human and economic terms, or practically impossible or prohibited by law".[44] Otherwise, distinguishing between the different means would be nearly impossible, since it would always be possible to imagine the hypothetical contingency of a third party who could — now or in the future — have additional relevant data to assist in the identification of a data subject.[45]

23 Although the *AG* states that in the future advances in technical means will "significantly facilitate access to increasingly sophisticated instruments for collecting and processing data" and thus, the safeguards put in place in defence of privacy are justified, this shall not result in a failure to take account of "the means likely reasonably to be used" by certain third parties.[46] Consequently, the *AG's* opinion includes several *relative* elements which clearly advocate against an *absolute* approach that would lead to an indefinite scope of the GDPR.

24 Nevertheless, according to the *AG*, it would be sufficient to obtain information "reasonably" if the legal *possibility* of retaining and transferring it to others exists. The possibility that the data *may* be transferred shall itself transform the dynamic IP address into personal data for the provider of services on the Internet.[47] The reasonable means of access shall be *lawful means*, therefore, "the legally relevant means of access are reduced significantly,

33 *Spindler*, Der Betrieb 2016, pp. 937 et seq.

34 *Härting* (*supra* Note 18), Recital 284; *Zuiderveen Borgesius* (*supra* Note 26), p. 256 (262).

35 ECJ, Case C-582/14 – *Patrick Breyer v Bundesrepublik Deutschland*.

36 German Federal Court of Justice (BGH), decision of 28 October 2014 – VI ZR 135/13.

37 German Federal Court of Justice, (*supra* Note 36), Recitals 27, 29 et seq.

38 *Härting*, Der IT-Rechts-Berater 2016, pp. 36 et seq.; *Keppeler*, Computer und Recht 2016, p. 360 (364).

39 Opinion of *Advocate General Campos Sánchez-Bordona*, delivered on 12 May 2016, Case C-582/14 – *Patrick Breyer v Bundesrepublik Deutschland*.

40 Opinion of the *Advocate General* (*supra* Note 39), Recital 64.

41 Opinion of the *Advocate General* (*supra* Note 39), Recital 65.

42 Opinion of the *Advocate General* (*supra* Note 39), Recital 65.

43 Opinion of the *Advocate General* (*supra* Note 39), Recital 68 (emphasis added).

44 Opinion of the *Advocate General* (*supra* Note 39), Recital 68; see also in favour of an "unreasonableness" of using illegal means *Spindler/Nink* in: Spindler/Schuster (eds.), Recht der elektronischen Medien, 3rd Ed. 2015, § 11 TMG Recital 8; *Brisch/Pieper*, Computer und Recht 2015, p. 724 (728), who argue that the wording of „reason" is not compatible with the use of illegal means, but who are, however, against a strict classification of illegal means as unreasonable and thus recommend a consideration of each individual case.

45 Cf. Opinion of the *Advocate General* (*supra* Note 39), Recital 68.

46 Opinion of the *Advocate General* (*supra* Note 39), Recital 66 et seq.

47 Opinion of the *Advocate General* (*supra* Note 39), Recital 72, who additionally names this *possibility* "perfectly reasonable"; cf. regarding the classification of dynamic IP addresses as personal data for *access providers* judged by the EJC, Case C-70/10, judgement of 24 November 2011 – *Scarlet Extended SA v Sabam*, Recital 51, which states that "[IP] addresses are protected personal data because they allow those users to be precisely identified".

since they must be exclusively lawful"[48], however, according to the *AG* it shall not matter how restrictive they may be in their practical application for constituting "reasonable means".[49] Allowing even the *possibility* of obtaining the data is a significant limitation of the above mentioned *relative* elements of the *AG's* interpretation and widens the material scope of the DPD and, consequently, also that of the GDPR significantly.

25 A further broadening of the scope and an orientation towards an *absolute* interpretation of identifiable can be found in the *GA's* statement that alone the sheer *potential* possibility of identification shall be sufficient and not that the dynamic IP address *only* becomes personal data when the Internet service provider receives it.[50] Hence, the *AG's* opinion can be interpreted as a vote for a rather *absolute* approach, which would lead to an even wider scope of the GDPR.

26 However, extending the scope of the Regulation too widely could lead to burdening regulations for data-processing entities which would be incommensurate with the actual risks to the privacy of the data subjects[51] and would thus not be compatible with the purpose of data protection law.[52] Because if the ECJ followed this broad – and nearly absolute – approach of the *AG*, virtually all data would have to be considered as personal data, which would, in the end, weaken the data protection framework and could make it unworkable[53], for instance because of an increase of informed consents and legal permissions to process the data.[54] If all data should be treated as personally identifiable and subjected to the GDPR, this could result in creating "perverse incentives" for controllers to abandon anonymisation and therefore increase, rather than relieve, privacy risks.[55] Thus, the very opposite of the protective intention would occur. Hence, we still hope that the ECJ will not follow the lines of argumentation of the *AG*.

2. Non Personal Data

27 Data which does not have any personal references, for instance sheer machine data or so called *attribute data*, does not fall under the material scope of the GDPR. Sensors that collect data for applications, e.g. made for climate analysis or the monitoring of industrial complexes do not process personal data at any stage.[56]

28 However, this attribute data can still turn into personal data when related to a natural person, for instance in the case of a worker's shift or when being linked with other information in a *Big Data* scenario.[57] Data from the *Internet of Things*[58], e.g. from cars, machines ("Industry 4.0"), smart homes or household applications will in many cases be connected to natural persons and thus be considered as personal data.[59] Moreover, the huge amounts of data can be used in connection with technologies like radio frequency identification tags ("RFID-tags") or monitoring and personal profiling so that identification might be easier than before.[60] How easy a re-identification is was demonstrated by a study carried out by computer science professor *Latanya Sweeney* which showed that the combination of a postal code, date of birth, and gender, is sufficient to identify 87% of individuals in the U.S[61], despite the fact that such data that are usually considered to be non personal data[62].

48 Cf. *supra* Note 44.

49 Opinion of the *Advocate General* (*supra* Note 39), Recital 73.

50 Opinion of the *Advocate General* (*supra* Note 39), Recital 77: "(...) their potential as a means of identifying — by themselves or together with other data — a natural person"; cf. *Keppeler* (*supra* Note 38), p. 360 (362).

51 Cf. *Schwartz/Solove*, California Law Review 2014, p. 877 (887).

52 Cf. Recital 4 S. 2 GDPR: "The right to the protection of personal data is not an absolute right; it must be considered in relation to its function in society and be balanced against other fundamental rights, in accordance with the principle of proportionality".

53 Cf. *Tene/Polonetsky*, Stanford Law Review 2012, p. 63 (66).

54 *Keppeler* (*supra* Note 38), p. 360 (364), who points out the practical problem that an increase of informed consents could mean that the text of the consents will be read even less by the data subjects and that a consent can be withdrawn by the data subject at any time, Article 7 Par. 3 GDPR.

55 Cf. *Tene/Polonetsky* (*supra* Note 53), p. 63 (66).

56 See *Rouvroy*, Council of Europe, T-PD-BUR(2015)09REV, Of Data and Men: Fundamental Rights and Freedoms in a World of Big Data, p. 20, available at: <https://www.coe.int/t/dghl/standardsetting/dataprotection/TPD_documents/T-PD-BUR(2015)09REV_Big%20Data%20report_A%20%20Rouvroy_Final_EN.pdf>, accessed 28 July 2016.

57 Cf. *Karg* (*supra* Note 17), p. 520 (522).

58 See for more use cases of the *Internet of Things*: *Vermesan/Friess* (eds.), Digitising the Industry - Internet of Things Connecting the Physical, Digital and Virtual Worlds, pp. 15 et seq., available at: <http://www.internet-of-things-research.eu/pdf/Digitising_the_Industry_IoT_IERC_2016_Cluster_eBook_978-87-93379-82-4_P_Web.pdf>, accessed 8 August 2016.

59 *Härting* (*supra* Note 18), Recital 268.

60 See regarding RFID and data protection law *TAUCIS*, Technikfolgenabschätzung: Ubiquitäres Computing und Informationelle Selbstbestimmung, 2006, pp. 198 et seq., available at: <https://www.datenschutzzentrum.de/taucis/ita_taucis.pdf>, accessed 29 July 2016; *Schmid*, Radio Frequency Identification Law Beyond 2007, in: Floerkemeier et al., The Internet of Things, 2008, pp. 196 et seq.

61 *Sweeney*, Carnegie Mellon University, School of Computer Science, Data Privacy Lab, Working Paper No. 3, 2000, available at: <http://dataprivacylab.org/projects/identifiability/paper1.pdf>, accessed 15 August 2016.

62 *Schwartz/Solove*, N.Y.U. L.Q. Rev. 2011, p. 1814 (1842),

29 Recital 30 of the GDPR now explicitly states that:

> *"(n)atural persons may be associated with online identifiers provided by their devices, applications, tools and protocols, such as internet protocol addresses, cookie identifiers or other identifiers such as radio frequency identification tags. This may leave traces which, in particular when combined with unique identifiers and other information received by the servers, may be used to create profiles of the natural persons and identify them."*

30 Thus, a lot of the data which originally was *attribute data*, e.g. produced by *Internet of Things* technologies, will become personal data due to the association of online identifiers with natural persons. Data of machines connected to the internet and operated by factory workers, of customers being tracked by RFID-tags, smart grid data, or of devices in smart homes or connected household appliances (e.g. toothbrushes, fridges, watches or TVs) will therefore be considered as personal data.[63] Additionally, natural persons can often be identified or be identifiable by "singling out"[64] their data. Thus, because of the broad material scope of the GDPR and of *Big Data* technologies, there are fewer and fewer possibilities to process data without a personal reference, in particular in the *Internet of Things* era.

3. Conclusion

31 The GDPR's material scope contains several parts which can be interpreted as *relative* approaches regarding the identifiability of natural persons, most prominently with the duty to include means only, if they are "reasonably likely to be used". Moreover, according to the *AG*, illegal means shall not be considered. Nevertheless, several other terms indicate a rather *absolute* approach of the GDPR, be it the wide scope of the online identifiers, the incorporation of "singling out" or that information obtained by a *third* person shall be sufficient to make the data personal for a controller. If the *AG's* opinion that the mere *possibility* of retaining and transferring the data to others is sufficient for a personal reference of data will prevail, the GDPR's material scope will have to be interpreted widely, using a mix of *relative* and *absolute* elements – an approach which could turn out to be a *pyrrhic victory*.

II. Encrypted Data and the GDPR

32 Encrypting personal data is a data security technique which has the effect of rendering data unintelligible to any person who is not authorised to access it due to encoding the information into a mutilated state, so that only parties with access to a decoding mechanism and a secret decryption key can access the information.[65] Encryption of data seems to be one of the promising solutions in order to ensure privacy particularly in cloud computing environments. When a controller encrypts the data before uploading it to a cloud, the data is regarded as personal data for the controller who holds the decryption key and the controller thus remains accountable for the data.[66] As encrypted personal data makes sure that no unauthorized person is able to use the sensitive data, only the original data controller is able to identify the persons related to data stored in the cloud – and not the cloud operator nor third persons. Hence, encryption may serve as a tool to safeguard data protection. Furthermore, when processing is carried out on behalf of the controller, such as in a cloud computing scenario, the GDPR introduces several new obligations to comply with - especially for processors and not only for controllers. Encrypting personal data can thus be a useful way to avoid these obligations for the processor.

33 In the cases where third parties are able to decrypt the data but the controller cannot, the question whether the GDPR shall be applicable for this controller is a point of controversy.[67] Moreover, if decryption has been achieved only by the use of illegal means, the controller who has not used those means shall not be subjected to the GDPR.[68]

34 In this section, we examine the provisions of the GDPR regarding encryption, anonymous and pseudonymous data in order to be able to assess the effect of encrypted personal data on the material scope of the Regulation.

available at: <http://scholarship.law.berkeley.edu/facpubs/1638>, accessed 10 August 2016; *Ohm*, UCLA Law Review 2010, p. 1701 (1705) with further examples.

63 See *International Working Group on Data Protection in Telecommunications*, Working Paper on Big Data and Privacy – Privacy principles under pressure in the age of Big Data analytics, 55th Meeting, 2014, Skopje, p. 4, available at: <http://dzlp.mk/sites/default/files/u972/WP_Big_Data_final_clean_675.48.12%20%281%29.pdf>, accessed 28 July 2016.

64 Regarding singling out people without knowing their names (for behavioural targeting) see *Zuiderveen Borgesius* (*supra* Note 26) pp. 256 et seq. and *supra* B.I.1.c.).

65 Cf. *ENISA*, Privacy by design in big data – An overview of privacy enhancing technologies in the era of big data analytics, 2015, p. 38; available at: <https://www.enisa.europa.eu/publications/big-data-protection>, accessed 9 August 2016; *Gürses/Kundnani/van Hoboken*, Media, Culture & Society, Crypto and empire: the contradictions of counter-surveillance advocacy, 2016, p. 7.

66 *Hon/Kosta/Millard/Stefanatou* (*supra* Note 29), p. 10.

67 See *infra* B.II.2.d.).

68 Cf. the Opinion of the *Advocate General*, *supra* Note 44.

1. Encryption in the GDPR

35 Unlike the proposal of the Parliament[69], the final version of the GDPR does not provide a further definition of encrypted data, but mentions encryption in several provisions as a compliance requirement. According to Article 32 Par. 1 (a) GDPR, encryption is regarded as an appropriate technical and organisational measure to ensure the security of processing. It is apparent that this does not deal with the applicability of the GDPR, but rather with the protection of personal data.[70]

36 Moreover, in case of a data breach, the controller is not required to communicate to the data subject if he or she has implemented encryption as a technical and organisational protection measure (Article 34 Par. 3 (a) GDPR).

37 Additionally, it is one of the "appropriate safeguards" of Article 6 Par. 4 (e) GDPR, which have to be taken into account when assessing the compatibility of a processing for a purpose other than that for which the personal data have been collected. Finally, depending on the classification of encryption as pseudonymisation or not[71], the provisions of the GDPR regarding pseudonymous data[72] may be applicable for encrypted data, too.

2. Is Encrypted Data Anonymised or Pseudonymised?

38 Since the GDPR does not define "encrypted data", we have to examine if encryption is a technique which anonymises or just pseudonymises personal data. In this regard, again the dispute regarding the material scope of the Regulation, as described above, plays an important role. To assess whether encrypted data has to be treated as anonymised or pseudonymised data, we first have to provide an overview of the GDPR's provisions regarding these privacy preserving techniques.

a.) "Anonymous Information" in the GDPR

39 Although technologies to anonymise personal data are considered to be of high value to protect the fundamental privacy rights of the data subjects, the GDPR does not provide a specific article to regulate "anonymous information" in the Regulation, it is only mentioned in one Recital. According to Recital 26 S. 4 and 5 GDPR the:

> *"principles of data protection should (...) not apply to anonymous information, namely information which does not relate to an identified or identifiable natural person or to personal data rendered anonymous in such a manner that the data subject is not or no longer identifiable. This Regulation does not therefore concern the processing of such anonymous information, including for statistical or research purposes."*

40 Thus, the GDPR is not applicable to anonymous data. To examine whether data can be considered as anonymous; once again the problem of the identifiability of data subjects arises.[73] In this regard, the possibility to anonymise personal data in the GDPR can be seen as another hint in favour of a *relative* approach, because given the possibilities to re-identify and combine data (*Big Data*), anonymous information could not be established when following a pure *absolute* approach.[74] However, to determine whether encrypted data may be considered as anonymous data, we will first take a look at the GDPR's provisions regarding pseudonymisation.

b.) "Pseudonymisation" and "Pseudonymous Data" in the GDPR

41 Unlike in the DPD, the GDPR includes a definition of "pseudonymisation". According to Article 4 No. 5 GDPR, pseudonymisation:

> *"means the processing of personal data in such a manner that the personal data can no longer be attributed to a specific data subject without the use of additional information, provided that such additional information is kept separately and is subject to technical and organisational measures to ensure that the personal data are not attributed to an identified or identifiable natural person".*

42 Moreover, "(t)he application of pseudonymisation to personal data can reduce the risks to the data subjects concerned and help controllers and processors to meet their data-protection obligations" (Recital 28 S. 1 GDPR). Furthermore, Recital 28 S. 2 GDPR emphasises that the explicit introduction of "pseudonymisation" does not intend to preclude any other measures of data protection. Thus, the connection between a

69 Article 4 No. 2b of the proposal of the European Parliament for a GDPR (LIBE proposal) defines encrypted data as "personal data, which through technological protection measures is rendered unintelligible to any person who is not authorised to access it", thus, according to LIBE, encrypted data shall just be a subcategory of personal data, which shall not lose its personal reference due to encryption.

70 See Recital 83 GDPR for more details regarding these measures.

71 See *infra* B.II.2.c.).

72 See *infra* B.II.2.b.).

73 See *supra* B.I.1.

74 *Härting* (*supra* Note 18), Recital 291.

natural person and the information on the basis of a corresponding rule remains – pseudonymised data is still qualified as personal data.[75] Hence, pseudonymisation is merely a method which can reduce the likelihood of identifiability of individuals, but does not exclude this data from the material scope of the GDPR. It is handled by the Regulation primarily as a data security measure,[76] and its use is encouraged in several articles of the GDPR; Article 32 Par. 1 (a) names it an appropriate technical and organisational measures to ensure a level of security appropriate to the risk.

43 Moreover, pseudonymisation shall, like encryption, be one of the "appropriate safeguards" of Article 6 Par. 4 (e) GDPR.[77] In addition, in accordance with Article 89 Par. 1 S. 3 GDPR, pseudonymisation is a safeguard to ensure that technical and organisational measures are applied when personal data is being (further) processed for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes. Finally, pseudonymisation is a technical and organisational measure that shall be implemented by the controller as a way to comply with the principle of data minimisation for the newly introduced provisions for "data protection by design and by default".[78] However, the GDPR does not distinguish between the quality of the possible pseudonymisation measures and its consequences for the controller. Nevertheless, to clearly define the unclear provision and the use of pseudonymisation, associations and other bodies representing categories of controllers or processors may prepare "codes of conduct" according to Article 40 Par. 2 (d).[79]

c.) Encrypted Data as Pseudonymised Data or Anonymous Data?

44 When encrypting personal data, in accordance with Article 4 No. 5 GDPR, the encryption key is the "additional information" which is "kept separately" and "subject to technical and organisational measures". Hence safety measures such as a secure key management and the respective encryption method used by the controller have to be used "to ensure that the personal data are not attributed to an identified or identifiable natural person". Therefore, because of its existing assignment rule encryption is an example of pseudonymisation.[80]

45 However, it is controversial whether encrypted personal data, and thus pseudonymised data, can be regarded as anonymised[81] data. Encrypted personal data should nevertheless undisputedly remain personal data to a person who holds the decryption key.[82] The relevant question is whether encrypted data shall also be personal data for a controller or processor who does not have access to the decryption key, for instance a cloud provider. Some academics have argued in this direction[83]; far more important, the *Art. 29 Data Protection Working Party* opines that believing that a pseudonymised dataset is anonymised is a "common mistake".[84] Additionally, the wording of Recital 26 S. 2 GDPR states that "personal data which have undergone pseudonymisation, which could be attributed to a natural person by the use of additional information should be considered to be information on an identifiable natural person".

46 At first sight, this is a clear statement of the EU legislator that pseudonymised data shall always be personal data. Nevertheless, to resolve this dispute, once again the question is crucial whether an *absolute* or a *relative* approach regarding the identifiability of a data subject has to be applied. According to the *absolute* approach, encrypted data will consequently always be personal data, because somebody, at least the key holder or any other party given sufficient time, economic resources and computing power, will always be able to decrypt the data, since no system of encryption can be completely secure[85]. According to this logic, encryption is merely a technical and organisational measure to ensure that data is not accessible to unauthorised persons rather than changing the data's quality. However, with a *relative* approach the data could be regarded as anonymous for the controller.

75 *Karg* (*supra* Note 17), p. 520 (522); see *infra* B.II.2.c.).

76 *Zuiderveen Borgesius* (*supra* Note 26), p. 256 (267).

77 See *supra* B.II.1.

78 According to Recital 78 GDPR, personal data should be pseudonymised "as soon as possible".

79 *Marnau* (*supra* Note 29), p. 428 (431).

80 *Article 29 Data Protection Working Party*, WP 216 (*supra* Note 5), p. 21; *Esayas* (*supra* Note 7), p. 8; *Hennrichs*, Cloud Computing - Herausforderungen an den Rechtsrahmen für Datenschutz, 2016, p. 137.

81 For an overview of existing anonymization techniques such as randomization or generalization see the Opinion of the *Article 29 Data Protection Working Party*, WP 216 (*supra* Note 5), pp. 12 et seq.; *International Working Group on Data Protection in Telecommunications* (*supra* Note 63), pp. 13 et seq., which provides guidelines for procedures for robust anonymisation; *ENISA* 2015 (*supra* Note 65), pp. 27 et seq.; *Lagos*, Indiana Law Review 2014-2015, pp. 187 et seq.

82 *Article 29 Data Protection Working Party*, WP 216 (*supra* Note 5), p. 29; *Borges*, in: Borges/Meents (eds.), Cloud Computing, 2016, § 6 Recital 33; *Polonetsky/Tene/Finch* (*supra* Note 26) p. 593 (613).

83 *Wagner/Blaufuß*, Betriebs-Berater 2012, p. 1751; *Esayas* (*supra* Note 7), p. 8.

84 *Article 29 Data Protection Working Party*, WP 216 (*supra* Note 5), p. 21.

85 Cf. *Kuner*, International Business Lawyer 1996, p. 186.

d.) Requirements for Encrypted Data in order to be considered as Anonymous Data

47 Consequently, we have to examine which level of encryption is sufficient so that with a *relative* approach the encrypted personal data can be considered as anonymous data. As mentioned above, only the knowledge and possibilities of the controller to identify the data subject shall be taken into account, therefore, processing encrypted data without affecting the scope of the data protection law might be possible.[86]

48 In order to concretise whether the means to decrypt the dataset and identify the data subject are reasonably likely to be used, one should take account of objective factors. There are three relevant factors that have to be considered when assessing the level of security of encrypted data against decryption, namely the strength of the encryption algorithm used, the length of the encryption key (the longer the key the safer the encryption will be) and the security of the key management.[87] Obviously, the key always has to be stored separately from the encrypted data in a secure way. If not, attackers may easily be able to decrypt the data[88] and thus, the personal data would no longer be anonymous. The simplest and most common way of decryption is using *exhaustive key search* or *brute-force attacks* which means to try all possible keys and eventually guessing correctly.[89] However, if a secure encryption technology is used, this way of decrypting the dataset cannot be considered as very likely for the controller.[90]

49 Other approaches to get access to the secret key are e.g. legally getting access to a key via a court decision, extracting the key from software or hardware, or by using accidental errors or systematic *backdoors* implemented in the encryption technique for law enforcement.[91] These ways are only considered to be likely for the controller if they do not violate the law or if they can be achieved by the use of computational power which can be reasonably expected. However, if a *backdoor* is implemented by the government into an encryption technology, the GDPR would be applicable for the controller who knows about this (governmental) possibility of accessing the personal data.

50 Additionally, as outlined above[92], the available encryption technology at the time of the processing has to be considered: applying the *AG's* opinion on encryption it would not be reasonably likely if it were practically impossible to decrypt the dataset, thus, if a state of the art encryption technology is enabled, in most of the cases, decrypting will be virtually impossible and therefore not likely and only possible with unreasonable efforts.[93] However, if according to the *AG* even the *potential* possibility of obtaining the decryption key from another person in a lawful way would be sufficient for an identification, the possibilities to avoid the applicability of the GDPR due to anonymisation via encryption would be very restricted.

51 Arguments against this wide interpretation could be sustained by Recital 57 GDPR, which deals with the data subject's right to access personal data held by the controller, where "the personal data processed by a controller do not permit the controller to identify a natural person". Then, "the data controller should not be obliged to acquire additional information in order to identify the data subject for the sole purpose of complying with any provision of this Regulation". This could be a hint against a too wide interpretation of getting access to a key obtained by a third party.

52 Additionally, future technological developments of decryption, e.g. due to more computing power or improved algorithms have to be considered (cf. Recital 26 GDPR), especially regarding the lengths of the secret key. The controller has to assess whether the future development is evidently foreseeable and thus ought to be regarded as a present information.[94] According to the *Article 29 Data Protection Working Party* (regarding the DPD), the controller should

86 Cf. *Hon/Kosta/Millard/Stefanatou* (*supra* Note 29), p. 10; *Borges,* in: Borges/Meents (eds.) (*supra* Note 82), § 6 Recital 33; *Hennrichs* (*supra* Note 80), 2016, p. 137.

87 *Hon/Millard/Walden* (*supra* Note 5), p. 22.

88 Cf. *Article 29 Data Protection Working Party*, WP 216 (*supra* Note 5), p. 22; *Hon/Kosta/Millard/Stefanatou* (*supra* Note 29), p. 10.

89 See with further examples *Gürses/Preenel*, in: van der Sloot/Broeders/Schrijvers (eds.), Exploring the Boundaries of Big Data, 2016, Part I, 3, Cryptology and Privacy in the Context of Big Data, p. 49 (62) available at: <http://www.ivir.nl/publicaties/download/1764.pdf>; accessed 29 August 2016; *Kroschwald*, Zeitschrift für Datenschutzrecht 2014, p. 75 (77).

90 *Cahsor/Sorge*, in: Borges/Meents (eds.) (*supra* Note 82), § 10 Recital 32, who state that using the 128 bits key lengths of AES encryption would make such an attack nearly impossible and thus not likely.

91 *Gürses/Preenel*, in: van der Sloot/Broeders/Schrijvers (eds.) (*supra* Note 89), p. 49 (63); the German and French government are currently deliberating on legal obligations to implement backdoors in encryption techniques for law enforcement reasons, see <http://www.interieur.gouv.fr/Actualites/L-actu-du-Ministere/Initiative-franco-allemande-sur-la-securite-interieure-en-Europe>, accessed 27 August 2016.

92 *Supra* B.I.1.c.).

93 Different opinion: A*rticle 29 Data Protection Working Party*, WP 216 (*supra* Note 5), p. 10, according to which the intentions of the data controller or recipient shall not matter, as long as the data are identifiable, data protection rules shall apply; see regarding the effect of different encryption technologies upon the applicability of the GDPR *infra* B.II.3.

94 *Borges*, in: Borges/Meents (eds.) (*supra* Note 82), § 6 Recital 38.

take into account the technological development for the period of time in which the data is meant to be processed, therefore, if the data shall be processed for ten years, he or she has to take the technological possibilities for these ten years into account; if the data can be decrypted in the ninth year, the data shall become personal data from that date on only.[95]

53 Therefore, due to technical developments, encrypted data will only be anonymous for a certain period of time and thus, the level of encryption has to be checked constantly by the controller and not only when the controller processes the data for the first time.[96] Moreover, if a controller receives an already encrypted dataset, he or she has to obtain further information regarding whether the original dataset has included personal data; if yes, the controller has to regularly check the state-of-the-art of the encryption technique.[97]

54 Thus, if the controller does not have the key to decrypt the data or other means to make it legible, it is in most cases reasonably likely that he or she cannot access the personal information, which consequently has to be regarded as anonymous information. Therefore, according to the GDPR, when using state-of-the-art encryption technique, encrypted personal data can be anonymous data, with the limitation that a potential possibility of obtaining the key, also by a third party and especially due to decryption, always has to be considered, but only if the means used are *reasonably likely*.

e.) Anonymisation as (Further) Processing of Personal Data

55 There is legal uncertainty regarding the lawfulness of the anonymisation process, more precisely whether anonymising personal data means "further processing" of personal data.[98] The *Working Party* states in its WP 216 (still regarding the DPD), that "anonymisation constitutes a further processing of personal data; as such, it must satisfy the requirement of compatibility by having regard to the legal grounds and circumstances of the further processing".[99]

56 Nevertheless, this further processing of personal data is considered to be compatible with the original purposes of the processing but only if the anonymisation process leads to "reliable (...) anonymised information".[100] Furthermore, the data controller's legitimate interest always has to be balanced against the data subject's rights and fundamental freedoms.[101] Consequently, according to the *Working Party*, anonymisation can be compatible with the original purposes of the processing, but it would be a violation of data protection law if personal data was anonymised for purposes that are not compatible with the original purpose and if there were no other legitimate grounds for processing the data, such as the data subject's consent.[102]

57 The *Working Party* clarifies this by providing as an example the anonymisation of the contents of traffic data immediately after its collection by mobile operators which performed deep packet inspection technologies. It was lawful in accordance with Article 7 (f) DPD, because of a legal permission stipulated in Article 6 Par. 1 of the e-Privacy Directive for certain traffic data which has to be erased or made anonymous as soon as possible when it is processed and stored by the provider of a public communications network or publicly available electronic communications service.[103]

58 Another example is given by *Esayas*, according to which the anonymisation of personal data for the purpose of using this data for advertising would constitute a violation of data protection law (unless there are other legitimate grounds for the processing – for instance the data subject's consent), if the data has originally been collected to provide a certain service for the data subject.[104]

59 Applying the *Working Party's* interpretation to the GDPR, the Regulation's requirements regarding further processing need to be fulfilled when anonymising personal data. Thus, it has to be analysed whether anonymisation is a *compatible* use according to the GDPR, then no legal basis separate from that which allowed the collection of

95 *Article 29 Data Protection Working Party*, WP 136 (*supra* Note 5), p. 18.

96 *Spindler*, (*supra* Note 23), p. 115; *Borges*, in: Borges/Meents (eds.) (*supra* Note 82), § 6 Recital 40; different opinion *Lundevall-Unger/Tranvik* (*supra* Note 21), p. 53 (71) who call it "a burden [for the controllers] that they probably cannot be expected to bear" and state that it "will not make controllers in a wired world more inclined to comply with the provisions of the [European data protection law]".

97 *Borges*, in: Borges/Meents (eds.) (*supra* Note 82), § 6 Recital 41.

98 See *El Emam/Álvarez*, International Data Privacy Law 2015, p. 73 (79); *Hon/Kosta/Millard/Stefanatou* (*supra* Note 29), p. 12; *Esayas* (*supra* Note 7), pp. 4 et seq.

99 *Article 29 Data Protection Working Party*, WP 216 (*supra* Note 5), pp. 3, 7.

100 *Article 29 Data Protection Working Party*, WP 216 (*supra* Note 5), p. 7.

101 Cf. *Article 29 Data Protection Working Party*, WP 216 (*supra* Note 5), p. 8.

102 Cf. *Walden*, International Journal of Law and Information Technology 2002, p. 224 (233); *Esayas* (*supra* Note 7), p. 4.

103 *Article 29 Data Protection Working Party*, WP 216 (*supra* Note 5), p. 8.

104 *Esayas* (*supra* Note 7), p. 4.

the personal data would be required (Cf. Recital 50 GDPR). Recital 49 GDPR states that:

> *"(t)he processing of personal data to the extent strictly necessary and proportionate for the purposes of ensuring (...) information security (...) constitutes a legitimate interest of the data controller concerned. This could, for example, include preventing unauthorised access to electronic communications networks and malicious code distribution (...)."*

60 According to this, the anonymisation of personal data could be interpreted as necessary for ensuring information security and be, in accordance with Article 6 Par. 1 (f) GDPR, of legitimate interest to a controller.[105] Apart from this, according to Article 5 Par. 1 (b) GDPR a "further processing for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes shall, in accordance with Article 89 (1), not be considered to be incompatible with the initial purposes".

61 Furthermore, according to the compatibility test of Article 6 Par. 4 GDPR, account should be taken *inter alia* of the possible consequences of the intended further processing for data subjects. Since anonymisation, pseudonymisation and encryption are privacy preserving technologies[106], in most cases applying these tools on the data subject's personal data will be in their interest.

62 However, regarding a possible re-identification of the personal data, the consequences of anonymising personal data for a data subject could also be serious (e.g. when processing special categories of personal data according to Article 9 GDPR) and thus not in its interest if the anonymous data, which is not affected by the scope of the GDPR, is transferred unrestricted from controller to controller.

63 Even though these concerns have to be taken seriously, the *Working Party's* opinion implies a non-existent weakness of the data protection law. Because as long as the data is anonymous, there is no threat to the privacy of the data subjects and as soon as a re-identification of the data is possible the GDPR with all its protective instruments is applicable again. Moreover, the need to justify the process of anonymisation itself could discourage the use of anonymisation and pseudonymisation as privacy-enhancing techniques.[107] However, with the use of Recital 49 GDPR, this dispute could possibly come to an end as soon as the GDPR comes into effect.

3. The Impact of different Encryption Techniques upon the GDPR's Material Scope

64 Finally, we will give a short overview of significant encryption technologies and examine the effect of these technologies on the applicability of the GDPR by determining *inter alia* which technical level of encryption has to be achieved to avoid a decryption or de-anonymisation of personal data and thus the applicability of the Regulation.

65 We have to distinguish between encrypted *transport* of data (e.g. encryption of e-mails or messages of messenger services via *end-to-end encryption*[108]) and encrypted *storing* of data (e.g. online backups in a cloud). However, if personal data is encrypted whilst being stored, applications and programs may not be able to handle and further process that encrypted data unless the data is decrypted and thus once again personal data. Processing stored encrypted data (e.g. in the cloud) in a secure and useful way – hence without the need of spending too much time or computer power – might be possible by using *Fully Homomorphic Encryption*[109] or *Secure Multiparty Computation*[110].

66 However, first of all, a distinction is made between *symmetric cryptography* and *asymmetric cryptography* techniques.

a.) Symmetric Cryptography – Secret Key Encryption

67 In a *symmetric cryptography* scenario, the parties use a publicly known encryption algorithm to transform the personal data into ciphertext or to later decrypt the dataset, the encryption is performed by a secret key which both parties have access to.[111] The level of security of the encrypted data depends significantly upon the secure storing, management, and transportation of the key which often cannot be transmitted safely.[112] Thus, a safe key management is

105 Cf. *Esayas* (*supra* Note 7), p. 5; *Hon/Kosta/Millard/Stefanatou* (*supra* Note 29), p. 12, who criticise that this legitimate interest should also refer to processors.

106 Cf. Recital 29 S. 1 GDPR which gives incentives for controllers to apply pseudonymisation when processing personal data; Article 5 Par. 1 (c) which regulates the principle of data minimisation, which is fulfilled by these technologies that reduce the amount of personal data.

107 *Hon/Kosta/Millard/Stefanatou* (*supra* Note 29), p. 12; *Esayas* (*supra* Note 7), p. 5.

108 See for details regarding WhatsApp's end-to-end encryption <https://www.whatsapp.com/security/?l=en>, accessed 26 August 2016.

109 See *infra* B.II.3.c.).

110 See *infra* B.II.3.d.).

111 Cf. *Gürses/Preenel*, in: van der Sloot/Broeders/Schrijvers (eds.) (*supra* Note 89), p. 49 (53).

112 *Maisch*, Informationelle Selbstbestimmung in Netzwerken, 2015, p. 322.

a necessary condition for avoiding the applicability of the GDPR, however, this can hardly be achieved when only using *symmetric cryptography*, because any holder of the key can easily re-identify the data subjects through decryption of the dataset.[113]

68 However, a safe transportation can be achieved when encrypting the *symmetric* key with an *asymmetric* encryption technique[114] (*hybrid cryptosystem*). Thus, a decryption in this scenario, when not *asymmetrically* encrypting the key, will in many cases be reasonably likely and the data protection law would thus be applicable for the controller or processor of the *symmetrically* encrypted database.

b.) Asymmetric Cryptography – Public Key Encryption

69 In *asymmetric* or public-key cryptography, two different keys are used, the first key (the public key) is used by the sender to encrypt the information, the second key is a private and secret key used by the recipient to decrypt the information.[115] Therefore, the encryption key can be made public, a common secret is not needed to be agreed on by the parties in advance as the second secret key is only known by the recipient.[116]

70 This technique is used mostly for *end-to-end encryption*. Thus, in an *asymmetric* encryption scenario the private key has to be kept secret. The risk that a third party could obtain the key consequently arises e.g. if the secret key is stored at a cloud provider which also holds the public key or by *man-in-the-middle attacks*, if a third party misleads the other parties by pretending to be the respective counterpart. If all necessary security measures are complied with – in the sense of the *relative* approach – it is not reasonably likely that a *man-in-the-middle attack* occurs.

71 However, in light of the *AG's* wide approach it may be sufficient that there is a *potential* possibility of identification by a third party. Thus, if the secret key is held safely by the recipient, a third party, e.g. a cloud provider which stores or transports the encrypted data does not have access to the private key and will, provided that a state-of-the-art key is used, not be able to decrypt the data (with reasonable efforts) and therefore does not fall under the scope of the GDPR. However, the controller always has to monitor the technological development regarding the key used and possible innovative technological ways of decryption.[117] Since *asymmetric* encryption has a significantly lower performance than *symmetric* encryption, in practice *hybrid* encryption is mostly used.

c.) Fully Homomorphic Encryption

72 *Fully Homomorphic Encryption (FHE)* is an encryption technology that allows the performance of an analysis "in the ciphertext in the same way as in the plaintext without sharing the secret key".[118] Therefore, for the processing of the data, no decryption and thus no knowledge of the private key is needed. Moreover, even the result of the processing is encrypted, which can only be decrypted by the user and not by the cloud provider.[119] The cloud provider will never see the data in plaintext. Thus, when processing personal data with the use of *FHE*, the GDPR is not applicable to the cloud provider which consequently does not process personal data. Unfortunately, due to its still very low performance, *FHE* is at present still highly inefficient and currently not a practical alternative to the processing of personal data on plaintext.[120]

d.) Secure Multiparty Computation

73 *Secure Multiparty Computation (SMC)*[121] allows for secure computation of sensitive data sets, such as tax or health data, without having to trust a centralised entity (such as a trusted third party).[122] It refers to a field of cryptography that deals with protocols involving two or more participants who want to mutually compute a useful result without having to

113 *Article 29 Data Protection Working Party*, WP 216 (*supra* Note 5), p. 20.

114 See *infra* B.II.3.b.).

115 *Gürses/Preenel*, in: van der Sloot/Broeders/Schrijvers (eds.) (*supra* Note 89), p. 49 (53).

116 *Gürses/Preenel*, in: van der Sloot/Broeders/Schrijvers (eds.) (*supra* Note 89), p. 49 (53).

117 See *supra* B.II.2.d.).

118 *ENISA* 2015 (supra Note 65), p. 40; *FHE* was first shown to be possible by *Gentry*, A fully homomorphic encryption scheme, 2009; another type of *homomorphic encryption* is *Somewhat Homomorphic Encryption (SHE)* which has a better performance than *FHE* but limits the number of operations.

119 *Gürses/Preenel*, in: van der Sloot/Broeders/Schrijvers (eds.) (*supra* Note 89), p. 49 (58).

120 *ENISA*, Privacy and Data Protection by Design – from policy to engineering, 2014, p. 43, available at: <https://www.enisa.europa.eu/publications/privacy-and-data-protection-by-design>, accessed 10 August 2016; *Gürses/Preenel*, in: van der Sloot/Broeders/Schrijvers (eds.) (*supra* Note 89), p. 49 (58).

121 *MPC* was first introduced by *Yao*, Proceedings of the 23rd Annual IEEE Symposium on Foundations of Computer Science, 1982, pp. 160 et seq.; for further details about *MPC* see *Cramer/Damgård/Nielsen*, Secure Multiparty Computation and Secret Sharing, 2015.

122 *Gürses/Preenel*, in: van der Sloot/Broeders/Schrijvers (eds.) (*supra* Note 89), p. 49 (60).

trust each other with their sensitive data.[123] Every party will provide an input value and learn only the result of their own individual value so that nobody is able to access all the information.[124] A data donor distributes the data into shares using secret-sharing and sends one random share of each value to a single server.

74 When using *Sharemind*[125], each party will receive one share of every secret value. The original secret can only be reconstructed by collecting all the shares of a value and adding them up.[126] After the data has been transmitted and stored, the server can perform computations on the shared data; however, the server does not share the information with other servers so that none of them can reconstruct the input values.[127] An increase of servers reduces the risk of collusions. After finishing the computation, the results of the servers are transmitted and published to the client of the computation. The servers send the share of the results to the client who reconstructs the real result.[128]

75 Thus, *Sharemind* requires three steps: the donors have to be informed whose data shall be provided; the data has to be divided; and then stored on the different servers. If it is necessary for one data donor to specify whose information the other donor has to provide, this has to be considered as processing of personal data. It would then be inevitable to identify the data subjects whose information is needed for the purposes of computation. Alternatively, to reduce the amount of personal data shared, all data can be loaded to *Sharemind* and securely joined using ciphertexts.

76 As outlined above, before the data is stored on the different servers, it has to be divided into the shares. This process must be carried out in plaintext using personal data. Although Article 4 No. 2 GDPR mentions the "alteration" of data as processing, the division of data does not entangle the application of the GDPR as "alteration" refers to the alteration of content, not of its appearance.[129] The secret-sharing of personal data by dividing it thus does not fall under the GDPR's scope. Once the data has been divided, it will be stored on the different servers. Applying an *absolute* approach on the identifiability of data subjects, these data chunks would have to be considered as personal data and this kind of processing would be processing of personal data – however, with the approach opined in this article, the data chunks are not considered to be personal data since it is highly unlikely for a party to receive the other shares.

77 *SMC* is advantageous due to the fact that simply random fragments of personal data are used. The original data can only be restored (and thus turned into personal data) if all fragments are put together. Hence, it is crucial to determine whether the GDPR is applicable to the computation over data fragments. Without the other parts, the file cannot be read in any way. One fragment itself does not contain information regarding a person and thus cannot be regarded as personal data. Only if all fragments of the data were gathered and put together, the Regulation would be applicable. Theoretically, all server providers may collude and reengineer the personal data. However, this is highly unlikely since the providers of the server themselves have a high interest in ensuring safety and confidentiality of the *SMC* and may be legally bound by contract.[130] This unreasonable chance of collusion leads to ruling out the applicability of the GDPR.

78 In contrast to *SMC*, when using *FHE* the parties do not need to be available online and the result is always encrypted.[131] However, *FHE* and *SMC* are special cases that still are not widespread, thus, when processing encrypted data without using these or similar technologies on some point it will always be necessary to decrypt the information with the consequence that in this moment the GDPR will be applicable to the controller again.[132]

C. Conclusion

79 Encrypting personal data can lead to the non-applicability of the GDPR and might thus be an important privacy preserving technology for controllers – however, since the provisions of the GDPR regarding its material scope also include several elements which can be interpreted in an *absolute* point of view and since the *Advocate General* of the ECJ has widened the scope in his opinion a lot, there is still legal uncertainty regarding the applicability of the GDPR for encrypted data. Therefore, controllers

123 Cf. *Bogdanov*, Sharemind: programmable secure computations with practical applications, 2013, available at: <http://dspace.ut.ee/bitstream/handle/10062/29041/bogdanov_dan_2.pdf?sequence=5&isAllowed=y>, accessed 27 August 2016.

124 *Kamm/Willemson*, International Journal of Information Security, 2015, p. 531 (532).

125 See <https://sharemind.cyber.ee//>.

126 *Bogdanov* (*supra* Note 123), p. 34.

127 *Kamm/Willemson* (*supra* Note 124), p. 531 (532).

128 *Kamm/Willemson* (*supra* Note 124), p. 531 (533).

129 Cf. for the German Federal Data Protection Act *Gola/Klug/Körffer*, in: Gola/Schomerus, Bundesdatenschutzgesetz, 12th Ed. 2015, § 3 Recital 30.

130 Regarding the risks for the confidentiality if parties pool their information see *ENISA* 2015 (*supra* Note 65), p. 41.

131 *Gürses/Preenel*, in: van der Sloot/Broeders/Schrijvers (eds.) (*supra* Note 89), p. 49 (60).

132 *Hoppen*, Computer und Recht 2015, p. 802 (804).

have to analyse each encrypted dataset on its own and determine whether a decryption might be reasonably likely, also taking continuously into account the use of future decryption technologies and the security of the key management. We hope that the ECJ does not follow this nearly *absolute* interpretation of the identifiability of natural persons since it would tremendously harm future incentives of controllers to implement privacy preserving technologies.

80 Additionally, encryption serves as a technical and organisational measure to ensure the security of processing in several parts of the Regulation. Controllers have to consider that the process of encryption as well as anonymisation might constitute a further processing of personal data.

81 Using state-of-the-art *asymmetric* encryption technologies especially for transporting personal data is a method which will in most of the scenarios be unlikely to be decrypted and can according to our interpretation prevent the applicability of the GDPR. Storing encrypted data in a cloud can also be done in a secure way without falling within the material scope of the GDPR. Although existing technologies such as *FHE* and *SMC* can exclude the applicability of the GDPR for the processing of encrypted data, processing encrypted data in most cases still has to be undertaken in plaintext by decrypting the data and thus by the use of personal data.

---

* *Prof. Dr. Gerald Spindler* is holder of the chair of Civil Law, Commercial and Economic Law, Comparative Law, Multimedia- and Telecommunication Law and head of the Institute for Business Law at the University of Göttingen, Germany.

Philipp Schmechel, Dipl.-Jur., is a Ph.D. student and research assistant for the EU-PRACTICE project at Prof. Spindler's chair at the University of Göttingen. His doctoral thesis deals with "Big Data and European data protection law".

The research leading to these results has received funding from the European Union Seventh Framework Programme ([FP7/2007-2013]) under grant agreement number ICT-609611 (PRACTICE). The information in this document is provided "as is", and no guarantee or warranty is given that the information is fit for any particular purpose subject to any liability which is mandatory due to applicable law. The user uses the information at its sole risk and liability.

jipitec

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

www.jipitec.eu