

3. P. Bernt Hugenholtz

Against ‘data property’

I. INTRODUCTION

With the incessant growth of the ‘data-driven economy’¹ have come calls for the introduction of a novel property right in data. In response, in particular, to demands of the automotive industry,² and encouraged by a fair number of German lawyers and scholars,³ the European Commission has in its recent Communication on ‘Building a European Data Economy’ tentatively advanced the idea of creating at EU level a ‘data producer’s right’ that would protect industrial data against the world.⁴ The movement for ‘data property’ (in German *Dateneigentum*) has its champion in European Commissioner Günther Oettinger, who until 2016 led the directorate general that is responsible for the Communication, DG

¹ OECD, *Data-Driven Innovation: Big Data for Growth and Well-Being* (2015).

² C. Schulzki-Haddouti, ‘Wem gehören die Daten im Auto?’, *VDI Nachrichten* 15 April 2016, available at <http://www.vdi-nachrichten.com/Technik-Gesellschaft/Wem-gehoren-Daten-im-Auto>

³ See, inter alia, T. Hoeren, ‘Big Data and the Ownership in Data: Recent Developments in Europe’, *European Intellectual Property Review* 7 (2014) 751–754; A. De Franceschi & M. Lehmann, ‘Data as Tradable Commodity and New Measures for their Protection’, 1 *Italian Law Journal* (2015) 51–72; H. Zech, ‘A Legal Framework for a Data Economy in the European Digital Single Market: Rights to Use Data’, 11 *Journal of Intellectual Property Law & Practice* (2016) 460–470.

⁴ European Commission, ‘Building a European Data Economy’, Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions, 10 January 2017, COM(2017) 9 final, 13. See in particular: European Commission, ‘Staff Working Document on the free flow of data and emerging issues of the European data economy’, Brussels, 10 January 2017, SWD(2017) 2 final, 33–38.

Connect. An op-ed published by Mr Oettinger in the *Frankfurter Allgemeine Zeitung* reveals some of the thinking and the powerful forces behind this revolutionary legal concept. Data, writes Oettinger, are the 'gold of the future', particularly in the automotive sector where modern sensor-equipped cars automatically generate and collect large amounts of data – on traffic and road conditions, engine performance, etc. These machine-generated sensor data have enormous value, for example, for developing self-driving automobiles. But – writes Oettinger – it is as yet unclear who owns these data: the automobile manufacturer; the car owner; the producer of the sensor equipment; or no one at all? What we need, concludes the Commissioner, are rules at EU level that establish data ownership.⁵

Apparently inspiring this call for protecting industrial data is the fear – common to other recent policy initiatives – that valuable European assets are being misappropriated by large American companies. The specter of Google 'stealing' European news has already led to an ongoing EU initiative towards a neighboring right for news publishers, following similar rules previously introduced in Germany and Spain.⁶ The sui generis database producer's right introduced in Europe in 1996 was similarly inspired by European fears of dominance by the US database industry.⁷

Although the contours of the 'data producer's right' now being contemplated by the European Commission are still sketchy, as are its economic underpinnings, such a right would most likely bring the protection of industrial data in the EU to a much higher level than the – much maligned and still controversial – database right. Whereas the database right protects data on the double condition that the data are structured in a 'database' and the database is the result of 'substantial investment', the novel right would directly protect machine-generated data without any material prerequisite.

⁵ G. Oettinger, 'Wem gehören die Daten?', *Frankfurter Allgemeine Zeitung*, 14 October 2016.

⁶ European Commission, Proposal for a Directive of the European Parliament and of the Council on Copyright in the Digital Single Market, 14 September 2016, COM(2016) 593 final, art. 11 (Protection of press publications concerning digital uses).

⁷ P. B. Hugenholtz, 'Something Completely Different: Europe's Sui Generis Database Right', in: S. Frankel & D. Gervais (eds), *The Internet and the Emerging Importance of New Forms of Intellectual Property*, Kluwer (2016), 205–222.

As this chapter argues, introducing such an all-encompassing property right in data would seriously compromise the system of intellectual property law that currently exists in Europe. It would also contravene fundamental freedoms enshrined in the European Convention on Human Rights and the EU Charter, distort freedom of competition and freedom of services in the EU, restrict scientific freedoms and generally undercut the promise of big data for European economy and society. In sum, it would be a very bad idea.

This chapter starts (in section II) by briefly examining the background and stated aims of the proposed new right: why would there be a need for creating a property right in industrial data? And what would be its subject matter and scope? Section III looks at existing intellectual property regimes, inquires to what extent these extend to data, and speculates on how a data property right in data might affect these regimes. Section IV thereafter scrutinizes the data right from the broader perspective of fundamental rights and freedoms. Section V concludes.

Although creating a property right in data surely has additional ramifications outside these fields, in particular for the right of informational privacy (personal data protection), the focus of this essay will be on the law of intellectual property. We will therefore not examine whether the law of data protection might already imply a property right in personal data.⁸ Nor shall we query whether the civil law concept of private property might be extended – or already extends – to (recorded) industrial data, and thus offers alternate protection to data sets.⁹ We shall also avoid discussing other doctrines in potential support of ‘data property’,¹⁰ such as criminal law¹¹ or trade secret law,¹² and stay away

⁸ See N. Purtova, *Property Rights in Personal Data: A European Perspective*, reviewed by E. J. Dommering, *Maandblad voor Vermogensrecht* (2012) 22–26.

⁹ See H. Zech, ‘Daten als Wirtschaftsgut – Überlegungen zu einem “Recht des Datenerzeugers”’, *Computer und Recht* (2015) 137–146.

¹⁰ See for a comprehensive overview of possible doctrines M. Dorner, ‘Big Data und “Dateneigentum”’, *Computer und Recht* (2014) 617–628; Osborne Clarke LLP, *Legal study on Ownership and Access to Data*, Study prepared for the European Commission DG Communications Networks, Content & Technology (2016), available at: <https://bookshop.europa.eu/en/legal-study-on-ownership-and-access-to-data-pbKK0416811/>

¹¹ T. Hoeren, ‘Dateneigentum: Versuch einer Anwendung von § 303 a StGB im Zivilrecht’, *Multimedia und Recht* (2013) 486–491.

¹² See J. Drexler, ‘Designing Competitive Markets for Industrial Data – Between Propertization and Access’, Max Planck Institute for Innovation & Competition Research Paper No. 16-13 (2016), 22–24.

from the contract and consumer law related issues of 'trading' personal data for services, which have become moot in the light of the proposed EU Digital Content Directive.¹³

Finally, a general caveat is in order. Whereas the European Commission has now posited the issue of 'data property' as worthy of serious discussion, the policy arguments advanced in favor of introducing such a right are underdeveloped, and its contours remain sketchy at best. Criticizing a right of data property is therefore taking aim at a moving target.

II. *QUID DATA PROPERTY?*

The arguments advanced by proponents of introducing a right of 'data property' can be roughly summarized as follows. First, it is pointed out that industrial data represent enormous economic value. For example, according to the European Commission:

[T]he value of the EU data economy was estimated at EUR 257 billion in 2014, or 1.85% of EU GDP. This increased to EUR 272 billion in 2015, or 1.87% of EU GDP (year-on-year growth of 5.6%). The same estimate predicts that, if policy and legal framework conditions for the data economy are put in place in time, its value will increase to EUR 643 billion by 2020, representing 3.17% of the overall EU GDP.¹⁴

Studies by the OECD and the European Commission present similarly mind-boggling figures.¹⁵

The rapidly increasing value of machine-generated data is attributed to a variety of factors: the rise of 'smart manufacturing', which involves real-time exchanges of massive amounts of data between machines and robots; the economic potential of 'mining' Big Data (i.e. extracting

¹³ Proposal for a Directive of the European Parliament and of the Council on certain aspects concerning contracts for the supply of digital content, COM(2015) 634 final; see A. Metzger, 'Dienst gegen Daten: Ein synallagmatischer Vertrag', 216 *Archiv für die civilistische Praxis* (2016) 817–865.

¹⁴ European Commission, 'Building a European Data Economy' (n. 4), at 2.

¹⁵ OECD, *Data-Driven Innovation: Big Data for Growth and Well-Being* (n. 1); European Commission, 'Towards a thriving data-driven economy', Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions, 2 July 2014, COM(2014) 442 final, 2.

information by way of sophisticated large-scale data analysis);¹⁶ and the promise of the *Internet of Things*, the magical world where machines quasi-independently communicate, and exchange data, directly with other machines, such as the ‘intelligent’ energy meter that sends usage data directly to the energy company, or the axiomatic refrigerator that automatically orders milk and coffee from the online supermarket. As in Commissioner Oettinger’s op-ed, many of the examples used in the literature are taken from the automotive sector, where data have become essential input and valuable output in manufacturing and navigation.¹⁷ The specter of Google’s self-driving car potentially out-competing the European car industry is never far away.

Having thus demonstrated that data have tremendous and increasing value, proponents go on to point out that current legal regimes, such as the traditional civil law based property right and existing intellectual property regimes, do not, or do not adequately, protect these data. Admittedly, non-property regimes such as contracts and trade secret protection might occasionally do the job, as would technical protection measures that create de facto ownership positions. However, these regimes do not create rights *erga omnes*, so valuable data are at risk of being misappropriated and a market for using and trading (i.e. licensing) raw data will not develop.¹⁸ Ergo, what is needed is a novel property right that protects industrial data as such. As the European Commission tentatively suggests in its Communication on ‘Building a European Data Economy’, ‘[t]his approach would aim at clarifying the legal situation and giving more choice to the data producer, by opening up the possibility for users to utilise their data and thereby contribute to unlocking machine-generated data’.¹⁹

In light of these radical contentions, it is surprising to see how little economic evidence is brought forward in support of a property right in data. According to standard economic analysis, there are two main rationales that might justify the creation of a new IP right: (1) solving the

¹⁶ In addition, Big Data analysis may have numerous social benefits; see Federal Trade Commission, *Big Data: A Tool for Inclusion or Exclusion? Understanding the Issues* (2016), 5–8.

¹⁷ See A. Wiebe, ‘Protection of Industrial Data – a New Property Right for the Digital Economy?’, *GRUR Int.* (2016) 878.

¹⁸ H. Zech, ‘Information as a Tradable Commodity’, in: A. De Franceschi (ed.), *European Contract Law and the Digital Single Market*, Intersentia (2016), 51–79; see also M. Dorner (n. 10), *Computer und Recht* (2014) 617–628.

¹⁹ European Commission, ‘Building a European Data Economy’ (n. 4), at 13.

public good problem by creating an economic incentive for the production of data; and (2) facilitating the use and trade of data. As to the first rationale, Professor Kerber, a leading German economist, sees 'no evidence that there are generally too few incentives for producing and analyzing data in the digital economy'.²⁰ As to the second argument, Kerber observes:

Although it cannot be ruled out that the market for trading and licensing data can suffer from market failure problems, and empirically data markets are still developing and need more scrutiny, it seems that so far data producers and holders have sufficient possibilities for commercializing their data. The potentially most important market failure problem that the first buyer might resell data seems to be solvable through either contractual and technical restrictions or through the strategy of selling services based upon these data.²¹

Professor Kerber concludes: 'there are no convincing economic arguments for the introduction of such a new IPR'.²²

This chapter will not, however, engage in economic analysis of a possible data property right, but focus instead on its consequences for the existing system of intellectual property. In order to do so, it is important to gain some preliminary understanding of what such a right might entail. This is another challenge. Whereas its proponents are enthusiastically calling for protection, most fail to draft even the roughest contours of such a new right. Drawing from the sketch presented in the Staff Working Paper that underlies the European Commission's recent Communication,²³ which seems to be largely based on the work of Professor Herbert Zech,²⁴ we assume the features of a *data producer's right* to be as follows. The right would create a *right in rem* (i.e. a property right enforceable against the world) in respect of 'non-personal or anonymised machine-generated data'. It would comprise

the exclusive right to utilise certain data, including the right to license its usage. This would include a set of rights enforceable against any party independent of contractual relations thus preventing further use of data by

²⁰ W. Kerber, 'A New (Intellectual) Property Right for Non-Personal Data? An Economic Analysis', *GRUR Int.* (2016) 989, 997.

²¹ Kerber, *ibid.*, at 998.

²² Kerber, *ibid.*, at 989. See also M. Dorner (n. 10), *Computer und Recht* at 625 (2014); Drexl (n. 12), Max Planck Institute for Innovation & Competition Research Paper No. 16-13 at 30 ff., 66 (2016).

²³ European Commission, Staff Working Document (n. 4).

²⁴ Zech, 'Information as a Tradable Commodity' (n. 18), at 74–76.

third parties who have no right to use the data, including the right to claim damages for unauthorised access to and use of data.²⁵

Whereas the Commission remains vague on the issue of initial ownership, according to Professor Zech, the right would initially vest in ‘the economically responsible operator of equipment that generates the data (data producer)’.²⁶ As the European Commission concedes, thus allocating the right might be highly problematic in practice, since data-generating machines are often owned and operated – and corresponding investments done – by numerous different actors.²⁷

In view of its stated aims, the right would have to be fully transferable.²⁸ As to the term of protection, the Commission is silent, but according to Professor Zech ‘[a] short term of protection seems to be appropriate since using data by analysing them can be done relatively quickly’.²⁹ The Commission excludes from the scope of the new right personal data, ‘as the protection of the latter is a fundamental right in itself under which natural persons should have control of their own personal data’.³⁰ Otherwise, the Commission’s proposal seems to encompass all sorts of machine-generated data.

As to its precise subject matter, Professor Zech proposes: ‘A well-defined subject matter would be machine-readable coded information that is defined only by its representative characters (bits) irrespective of its content (data delimited on the syntactic level)’.³¹ This distinction is reflected in its intended scope. ‘The scope of protection would in particular include use by carrying out statistical analyses, but not the re-creation of the same data by independent measurement’.³² The Commission seems to embrace this distinction, perhaps in the hope that such a limitation might prevent undue information monopolies.³³ We shall examine the distinction between syntactic and semantic data in the following section (at section III.E).

In sum, both in terms of its intended subject matter (data, an immaterial good) and its scope of protection (reproduction and use of

²⁵ European Commission, Staff Working Document (n. 4), at 33.

²⁶ Zech, ‘Information as a Tradable Commodity’ (n. 18), at 75.

²⁷ European Commission, Staff Working Document (n. 4), at 35.

²⁸ Zech, ‘Information as a Tradable Commodity’ (n. 18), at 76.

²⁹ *Ibid.*

³⁰ European Commission, Staff Working Document (n. 4), at 33.

³¹ Zech, ‘Information as a Tradable Commodity’ (n. 18), at 74.

³² *Ibid.*, at 75.

³³ European Commission, Staff Working Document (n. 4), at 34.

data by third parties), the proposed data producer's right would probably qualify as a right of intellectual property.

III. DATA IN THE SYSTEM OF INTELLECTUAL PROPERTY

Before further scrutinizing the proposed data right and its possible impact on the system of intellectual property law, we first need to examine how existing legal regimes in the EU deal with machine-generated data. This section will focus on relevant laws of intellectual property, and not discuss other possibly relevant legal mechanisms, such as tangible property, contract and trade secrets. While the focus of this section will be on the two IP regimes most closely associated with protecting data structures, copyright law and a sui generis database right, we shall also make a brief excursion into the field of neighboring rights.

A. Copyright in Data

Is there copyright in data? The textbook answer is a resounding *no*. As US Supreme Court Justice Warren Brandeis famously stated in his dissent in the *INS* case, '[t]he general rule of law is, that the noblest of human productions – knowledge, truths ascertained, conceptions and ideas – after voluntary communication to others, are free as the air to common use'.³⁴ This axiom reflects what is called the *idea/expression dichotomy*: the dividing line between, as European copyright scholars prefer to say, protected form and unprotected content(s). The rule is generally codified in the US Copyright Act (s. 102 (b)), as well in the TRIPs Agreement (art. 9(2)) and the WIPO Copyright Treaty (art. 2). In EU law we find a similar rule, albeit limited to computer software, in the Computer Programs Directive (art. 1(2)).³⁵ Although these provisions do not expressly mention 'data', it is generally assumed, and uncontroversial – either on the basis of the 'dichotomy' or by way of direct application of copyright's requirement of authorship and creativity – that there cannot be copyright in data per se.³⁶

³⁴ J. Brandeis, diss. opinion, *International News Service v. Associated Press*, 248 U.S. 215 (1918).

³⁵ Directive 2009/24/EC of the European Parliament and the Council of 23 April 2009 on the legal protection of computer programs, OJEU 2009 L 111, 16 (recast).

³⁶ P. B. Hugenholtz, *Auteursrecht op informatie*, Amsterdam (1989).

Whereas data as such are thus excluded from copyright protection, copyright's treatment of data compilations is more complex. The Berne Convention protects 'collections of literary or artistic works such as encyclopedias and anthologies' (art. 2(5)), but does not mention collections of mere data, and expressly denies copyright to 'news of the day or to miscellaneous facts having the character of mere items of press information' (art. 2(8)). The TRIPs Agreement (art. 10(2)) more broadly protects 'compilations of data or other material, whether in machine readable or other form, which by reason of the selection or arrangement of their contents constitute intellectual creations'. The WIPO Copyright Treaty (WCT) contains similar language (art. 5). In line with these modern conventions the EU Database Directive instructs Member States to provide copyright protection to 'databases which, by reason of the selection or arrangement of their contents, constitute the author's own intellectual creation' (art. 3(1)).³⁷

In *Football Dataco and others* the Court of Justice clarified that the test of 'the author's own intellectual creation' (in short, originality) implies that the selection or arrangement of the data is the result of creative choices. Applying art. 3(1), second sentence ('No other criteria shall be applied to determine their eligibility for that protection'), the Court held that merely investing significant amounts of skill and labor does not justify a finding of originality. In other words, the Directive's originality standard preempts any (quasi-)copyright protection for databases that is merely based on investment or other criteria.³⁸ The Court's decision has brought to an end not only the United Kingdom's long-standing practice of according copyright protection to compilations of data based on 'skill and labour' (investment), but also similar doctrines in other Member States. For example, the Dutch protection of non-original writings ('geschriftenbescherming') that existed for over a century in the Netherlands as a vehicle for protecting non-original writings and compilations, was formally abolished in 2014 following *Football Dataco*.³⁹

³⁷ Directive 96/9/EC of the European Parliament and the Council of 11 March 1996 on the legal protection of data bases, OJEC 1996 L 77, 20

³⁸ *Football Association Premier League Ltd and Others v. QC Leisure and Others; and Karen Murphy v. Media Protection Services Ltd*, CJEU, 4 October 2011, joined cases C-403/08 and C-429/08, ECLI:EU:C:2011:631.

³⁹ See P. B. Hugenholtz, 'Works of Literature, Science and Art', in: P. B. Hugenholtz, A. A. Quaedvlieg & D. J. G. Visser (eds), *A Century of Dutch Copyright Law. Auteurswet 1912–2012*, deLex (2012), 54. See also P. B. Hugenholtz, 'Goodbye, Geschriftenbescherming!', Kluwer Copyright Blog, 6 March

Football Dataco also rules out copyright protection for data compilations that are generated by machines without any human intervention. This is in line with the general rule that copyright requires acts of human authorship. Note however that the UK Copyright, Designs and Patents Act appears to extend copyright protection to machine-created works: 'in the case of a literary, dramatic, musical or artistic work which is computer-generated, the author shall be taken to be the person by whom the arrangements necessary for the creation of the work are undertaken'.⁴⁰ Whether this rule can co-exist with *Football Dataco* remains to be seen. According to the European Court there is no originality 'when the setting up of the database is dictated by technical considerations, rules or constraints which leave no room for creative freedom'.⁴¹

Both TRIPs and WCT caution that copyright in compilations of data 'not extend to the data or material itself'. Similarly, the Database Directive (art. 3(2)) warns that database copyright 'shall not extend to their contents', thus ruling out copyright protection for the data compiled in a database. The scope of database copyright protection is limited to the original structure (selection or arrangement) of the database (Database Directive, recital 15). Extracting (parts of) the contents of the database without appropriating the selection or arrangement does not infringe the copyright in the database.⁴²

B. Sui Generis Database Right

Art. 7(1) of the Database Directive supplements the Directive's copyright regime by obliging Member States to protect databases that result from substantive – qualitative or quantitative – investment. This is the sui generis database right that has made the Directive internationally (in)famous. The substantial investment is to be made 'in either the obtaining, verification or presentation of the contents' of the database

2013, available at <http://kluwercopyrightblog.com/2013/03/06/goodbye-geschriftenbescherminng/>

⁴⁰ Copyright, Designs and Patents Act, UK 1988 (c. 48), c. 1, s. 9(3); see M. Perry & T. Margoni, 'From Music Tracks to Google Maps: Who Owns Computer-generated Works?', 26 *Computer Law & Security Review* (2010) 621–629.

⁴¹ *Football Dataco & Others v. Stan James Plc & Others and Sportdar GmbH & Others*, Court of Appeal (Civ. Division), 6 February 2013, [2013] EWCA Civ 27.

⁴² *The Newspaper Licensing Agency & others v. Meltwater & the PRCA*, High Court of Justice (Chancery Division), 26 November 2010, [2010] EWHC 3099.

(art. 7(1)). ‘Obtaining’ is the act of gathering the data, works or other materials to be included in the database. ‘Verification’ relates to the checking, correcting and updating of data already existing in the database. ‘Presentation’ concerns acts such as digitizing (scanning) analogue files, or creating a thesaurus. A decision by the German Federal Supreme Court suggests that the standard of ‘substantial investment’ is not very hard to meet. Any investment in a database that ‘viewed objectively ... is not wholly insignificant and easy to be made by anyone’ would suffice.⁴³ The European Court of Justice has yet to opine on the level of this threshold criterion.

In four landmark cases concerning the unauthorized use by betting companies of sports events schedules (‘fixtures’), the European Court held that the database right does not protect investment in generating the data or other contents of a database. According to the Court, ‘investment in the obtaining of the contents’ of a database ‘refers to the resources used to seek out existing materials and collect them in the database but does not cover the resources used for the creation of materials which make up the contents of a database’.⁴⁴ The main argument for this distinction, as is transparent from the decision, is that the Database Directive’s economic rationale is to promote and reward investment in database production, not in generating new data. According to the Court, ‘[t]he purpose of the protection by the sui generis right provided for by the directive is to promote the establishment of storage and processing systems for existing information and not the creation of materials capable of being collected subsequently in a database’.⁴⁵

Thus, investment in ‘creating’ data does not count towards investment. However, the European Court’s epistemological distinction between ‘creating’ and ‘obtaining’ data is not self-evident.⁴⁶ While the Court ruled out from sui generis protection such ‘invented’ data as horse racing

⁴³ Bundesgerichtshof (Federal Supreme Court), 1 December 2010, case I ZR 196/08, GRUR 2011, 724.

⁴⁴ CJEU, case C-46/02, 9 November 2004, *Fixtures Marketing Ltd v. Oy Veikkaus Ab*, ECLI:EU:C:2004:694; case C-203/02, 9 November 2004, *British Horseracing Board v. William Hill Organization*, ECLI:EU:C:2004:695; case C-338/02, 9 November 2004, *Fixtures Marketing Ltd v. Svenska Spel AB*, ECLI:EU:C:2004:696; case C-444/02, 9 November 2004, *Fixtures Marketing Ltd v. Organismos prognostikon agonon podosfairou AE (OPAP)*, ECLI:EU:C:2004:697.

⁴⁵ *British Horseracing* (n. 44), para. 31.

⁴⁶ M. J. Davison & P. B. Hugenholtz, ‘Football Fixtures, Horseraces and Spin-offs: The ECJ Domesticates the Database Right’, *European Intellectual Property Review* (2005) 113–118.

schedules and football fixtures, the Court of Appeal of England and Wales in a subsequent decision held that facts observed – such as the scoring of a goal in football – are not ‘created’ data.⁴⁷ But where in this spectrum between purely synthetic data and data ‘observed’ should we place machine-generated data? The answer depends on the type of data that the machine processes. For example, sensor data produced by a radar system or by an observation satellite are likely to qualify as data ‘observed’, and concomitant investments may thus be taken into account when applying the database right. Conversely, computer-generated airline schedule data squarely falls under the rubric of ‘created’ data excluded by the European Court.

Whereas the *sui generis* right comes close to a property right in aggregate data – and has been justly criticized for its potential to create harmful information monopolies⁴⁸ – the Directive’s recitals admonish that the *sui generis* right ‘does not in any way constitute an extension of copyright protection to mere facts or data’ (recital 45) and ‘should not give rise to the creation of a new right in the works, data or materials themselves’ (recital 46).

Indeed, the database right comes with several statutory limits in order to prevent the right from extending to the data in the database *per se*. The *sui generis* right protects database producers against ‘extraction’ and ‘reutilization’ of the whole, or a substantial part, of the database (art. 7). In other words, non-substantial takings of data are permitted without authorization. Moreover, as the European Court clarified in *British Horseracing*,⁴⁹ *sui generis* protection does not extend beyond misappropriation of data (contents) that result from substantial investment. In other words, the database right tolerates takings of (potentially valuable) data that are not the product of substantial investment.

Another delimiting factor is the notion of ‘database’. Art. 1(2) of the Directive defines this as ‘a collection of independent works, data or other materials arranged in a systematic or methodical way and individually

⁴⁷ *Football Dataco & Others v. Stan James Plc & Others and Sportradar GmbH & Others*, Court of Appeal (Civ. Division), 6 February 2013, [2013] EWCA Civ 27.

⁴⁸ See J. H. Reichman, ‘Database Protection in a Global Economy’, *Revue Internationale de Droit Economique* (2002) 455–504.

⁴⁹ *British Horseracing* (n. 44).

accessible by electronic or other means'. While the Explanatory Memorandum generally describes the contents of the database as "information" in the widest sense of that term,⁵⁰ the compiled data or materials must be 'independent', that is to say, 'materials which are separable from one another without their informative, literary, artistic, musical or other value being affected'.⁵¹ Therefore, a literary work, a musical composition or a sound recording does not qualify as a database, even if it can be conceived as a representation of data (recital 17). This reflects a clear intention on the part of the European legislature to avoid extensive overlaps between the database right and existing copyright and neighboring rights.⁵²

Finally, according to art. 1(2) of the Database Directive, the individual elements of the database must be 'arranged in a systematic or methodical way'. This squarely rules out protection – whether by copyright or by database right – of raw data.

C. Phonogram Protection

In addition to copyright and database right, the phonographic right – one of the four neighboring rights recognized at EU level – merits brief consideration. The rights of phonogram producers are harmonized by the Rental Right Directive⁵³ and the Information Society Directive⁵⁴ (Directive 2001/29/EC). These Directives leave defining the notion of 'phonogram' to the WIPO Performances and Phonograms Treaty (WPPT) of

⁵⁰ Explanatory Memorandum, Proposal for a Council Directive on the Legal Protection of Databases, Brussels (2005) 19.

⁵¹ *Fixtures Marketing Ltd v. Organismos prognostikon agonon podosfairou AE (OPAP)* (n. 44).

⁵² On the other hand, the European Court has held that the geographical data in a topographic map are sufficiently 'independent' for the map to qualify as a protected 'database'. According to the Court, 'geographical information extracted from a topographic map by a third party so that that information may be used to produce and market another map retains, following its extraction, sufficient informative value to be classified as "independent materials" of a "database" within the meaning of that provision', CJEU, case C-490/14, 29 October 2015, *Freistaat Bayern v. Verlag Esterbauer GmbH*, ECLI:EU:C:2015:735.

⁵³ Directive 2006/115/EC of the European Parliament and the Council of 12 December 2006 on rental right and lending right and certain copyright related rights in the field of intellectual property, OJEU 2006 L 376, 28.

⁵⁴ Directive 2001/29/EC of the European Parliament and the Council of 22 May 2001 on the harmonization of certain aspects of copyright and related rights in the information society, OJEC 2001 L 167, 10

1996. According to the WPPT (art. 2) 'phonogram' means 'the fixation of the sounds of a performance or of other sounds, or of a representation of sounds, other than in the form of a fixation incorporated in a cinematographic or other audiovisual work'. By including 'other sounds' and 'a representation of sounds' this definition apparently encompasses all sorts of raw audio data stored ('fixed') on a digital medium.

Whether there is a threshold criterion for the phonographic right that might delimit both the substance and scope of the right, is as yet unsettled under EU law. In its *Metall auf Metall* decision of 2008 the German Bundesgerichtshof extended neighboring rights protection to every single recorded note of a sound recording, because the record producer's investment is reflected in every – even a very minor – part of the recording.⁵⁵ This suggests that no threshold criterion (no investment minimum) applies. In a follow-up decision the German Constitutional Court has however held that a phonographic right of unlimited scope, as contemplated by the Federal Supreme Court, may collide with the 'freedom of art' that is constitutionally guaranteed in Germany (art. 5 of the Basic Act).⁵⁶

D. Assessment: Impact of Data Property on the System of Intellectual Property Law

As this section shows, both copyright and the database right do not extend to data per se. Both regimes do conditionally offer protection to data compilations that result from creative selection and arrangement (copyright) or substantive investment (database right). Both regimes deny protection to machine-generated raw data. For copyright, this follows from the axiom that only acts of authorship conducted by human beings are protectable. For the database right, this is a consequence of the sui generis right's categorical delimitation: only data structured in a 'database' qualify for protection. Moreover, the sui generis right's substantial investment test sets an – admittedly fairly low – minimum threshold. If operating a machine that records sensor data does not require substantial investment (for example, a low-cost digital weather station or a bicycle computer), then this will not result in a protected database. The CJEU's 'Fixtures' decisions pose an additional hurdle to sui generis protection for

⁵⁵ *Metall auf Metall*, Federal Supreme Court, 20 November 2008, case I ZR 112/06, *GRUR* 2009, 403.

⁵⁶ German Constitutional Court, 31 May 2016, case 1 BvR 1585/13, *GRUR* 2016, 690.

machine-generated data by excluding ‘created’ data from protection, thus ruling out machine-generated synthetic data.

In sum, introducing a right in raw, machine-generated industrial data, as envisaged in the Commission’s Communication, would go far beyond the main IP regimes presently existing in Europe in the field of data and information, copyright and database right. The only remotely comparable regime would be the neighboring right of phonogram producers, which does seem to extend to raw (audio) data.

E. Disruptive Overlaps

How would this affect existing IP law? In the first place, creating a new layer of rights in machine-generated data would cause broad and disruptive overlaps with copyright and the sui generis right in productions made with the aid of digital machines. For example, a film shot with a digital camera would qualify not only as a work protected by copyright, but also as machine-generated (sensor) data subject to a ‘data producer’s right’. Similarly, the aggregate stock market data in a financial database would be protected both by the sui generis right and the ‘data producer’s right’, since the data are recorded automatically by the computerized stock exchange.

Whereas the EU legislature has clearly intended to prevent the database right from spilling over into the realms of copyright and neighboring rights, the ‘data producer’s right’ would lead to extensive overlaps. As a consequence, the new right might give rise to multiple competing claims of ownership in the same content. To continue with our first example, while the creators of the film (e.g. the director, screen writer, and other creators of the film) could claim authorship in the cinematographic work, the owner or operator of the camera might claim ‘data property’ in the photographic data (i.e. the digital representation of the film) – surely to the unpleasant surprise of the film’s producer. Similar examples might be given with regard to digital photographs or e-books. In the second example, the database producer might be confronted with ‘data property’ claims by the stock exchange, or the exchange’s computational services company.

Another consequence of this wide-ranging overlap would be that statutory limitations and exceptions under copyright, neighboring rights or the database right are ‘trumped’ by the data producer’s right. For example, both copyright and the database right in the EU presently allow users to copy or extract data from databases for non-commercial research purposes. Unless the ‘data producer’s right’ was to replicate all relevant

existing exceptions in the field of copyright, neighboring rights and the database right, it would undercut these essential user freedoms.

This is especially true for *data mining*. Bizarrely, while the Commission's Communication on 'Building a European Data Economy' ponders the introduction of an exclusive right in machine-generated data, one of the highlights of the Digital Single Market (DSM) Directive proposal that is currently being debated in the European Parliament is a mandatory exception, both under copyright and the database right, for text and data mining by non-commercial research organizations.⁵⁷

In line with Professor Zech's suggestions, the European Commission in its Staff Working Paper attempts to alleviate concerns of wholesale overlap by distinguishing *syntactic* from *semantic* data. The proposed 'data producer's right' would be conceived in such a way that 'only the syntactical level of information is protected, not the semantic level'.⁵⁸ What is probably meant here is that the raw data would be protected only as regards its digital representation (the machine-readable bits and bytes, the 'ones and zeros' in the digital file), not the informational content that these data convey. Thus, the European Commission hopes, the new right would not extend to ideas and information, and the new right would not become a 'super-IP right'.⁵⁹

But would such a distinction really prevent the new right from extensively overlapping with existing IP rights? I do not believe so. The problem here is that digital data are commonly coded and interpreted following standardized rules and protocols. In other words, there will usually be a one-to-one relationship between the (syntactic) data substrate and the (semantic) content layer. Returning to our example of the digitally produced film, any copy of the film's digital file (the syntactic data) would by necessity also reproduce the copyright protected work (the semantic layer). Thus, the new data right could be invoked against any digital copying (or streaming) of the digitized copyright work. For

⁵⁷ European Commission, Proposal for a Directive of the European Parliament and of the Council on Copyright in the Digital Single Market, Brussels, 14 September 2016, COM(2016) 593 final. Art. 3(1) of the proposed Directive provides: 'Member States shall provide for an exception to the rights provided for in Article 2 of Directive 2001/29/EC, Articles 5(a) and 7(1) of Directive 96/9/EC and Article 11(1) of this Directive for reproductions and extractions made by research organisations in order to carry out text and data mining of works or other subject-matter to which they have lawful access for the purposes of scientific research'.

⁵⁸ Zech, 'Information as a Tradable Commodity' (n. 18), at 74; European Commission, Staff Working Document (n. 4), at 34.

⁵⁹ European Commission, Staff Working Document (n. 4), at 34.

the same reason, the new right would broadly overlap with the database right, even if its scope were confined to the syntactic layer. The phonographic right discussed above illustrates this point. Whereas its subject matter, like the proposed ‘data producer’s right’, is limited to the recorded signal (i.e. syntactic audio data), its scope extends into the semantic realm. Reproducing a CD recording of a musical performance will, by necessity, result in the reproduction of the underlying musical work and performance.

The only way to prevent the data right from becoming an all-encompassing ‘super-IP right’ would be to categorically exclude all data that (possibly) represent subject matter protected under traditional IP regimes: not only copyright, the database right and neighboring rights, but also design right and perhaps patent law. But even a non-overlapping data right would have seriously corrosive effects on the system of IP, for various reasons. First, it would undermine the economic incentives that underlie existing IP rights. For example, the main rationale of the *sui generis* right is to promote and reward investment in the building of databases from pre-existing data and other materials. This incentive is clearly undercut if a lower-tier, no-threshold right in machine-generated data were to exist in parallel. Second, and more importantly, it would compromise the general principle of intellectual property – whether utilitarian or grounded in natural law theory – that protection ought to be reserved to creation, innovation or otherwise meritorious investment. A data right in all data produced by machines might, on occasion, protect assets of considerably economic value, but nothing of merit. This has ramifications, in particular, at the political level. With intellectual property laws under increasing fire, legislatures – at EU and national level – need powerful and convincing arguments to defend existing regimes and introduce new rights. In this volatile political climate, proposing a data producer’s right with the sole aim of (better) protecting the economic assets of the automotive industry will surely backfire. Not only is such an initiative likely to fail in the legislative process, but it will also (re)ignite broader discussions on the legitimacy of intellectual property law.

F. No Legal Certainty

Another, more mundane objection against a property right in data lies in its inherent lack of legal certainty. Although it is still not fully conceptualized, it is difficult to imagine a data right that would be sufficiently stable – in terms of subject matter, scope and ownership – to be admitted to the ranks of IP. As to subject matter, if the right vests in data generated by machine processes, which data would it protect: all the data that the

machine produces within a given time frame (e.g. an hour, a minute or a second); or all the data that result from a finite machine process (e.g. all the data gathered by a satellite that sensors the earth)?

Admittedly, the sui generis database right has already raised similar questions. With data in a database constantly being updated, what exactly constitutes the protected database? But in database law the definition of 'database' and of the requirement of substantial investment create at least some measure of permanency in the subject matter and scope of the right. This stability is, however, completely absent from the data producer's right. The problem here is that industrial data generation occurs in real time. The 'velocity' – the dynamic nature – of big data makes it very difficult, if not impossible, to identify a stable object of protection.⁶⁰ The subject matter of the right is simply too fluid. If this is to become a fully-fledged right of intellectual property that is enforceable against the world, it should be possible to ascertain its subject matter – and, by implication, its scope of protection – with sufficient legal certainty.

A related problem is allocating ownership of the right. Since the right would be sparked by machine operations, no 'natural' ownership connection with a natural person as, for instance, in copyright, exists. As Professor Zech and the European Commission suggest, ownership might be vested in the person owning or operating the machine that generates the data. This, however, is hardly a reliable rule. As the OECD points out in its groundbreaking study on 'big data', multiple actors/stakeholders might claim ownership of the data, both upstream and downstream in the process of generating and processing data.⁶¹

In sum, the proposed 'data producer's right' would most likely seriously affect, or even distort, existing copyright as well as the database right and its underlying incentives. Moreover, in the absence of clear and predictable rules circumscribing its subject matter, scope and ownership, this would lead to gross legal uncertainty. This conclusion in itself justifies serious restraint on the part of the EU legislature, even without considering the adverse effect the new right might have on the free flow of information, one of the cornerstones of the emerging information society.

⁶⁰ Drexl (n. 12), Max Planck Institute for Innovation & Competition Research Paper No. 16-13 at 15 (2016).

⁶¹ See OECD (n. 1), at 195–196; Drexl (n. 12), Max Planck Institute for Innovation & Competition Research Paper No. 16-13 at 6 (2016).

IV. DATA PROPERTY AND THE FREE FLOW OF INFORMATION

The exclusion of data per se from the scope of existing IP regimes is not merely ontological. Although old-school author's-right scholars might argue that data are not copyright works, *because* data are not 'created', this is at best a partial explanation for this exclusion.⁶² Rather, IP law's abhorrence of protecting data reflects implicit or explicit information policies not to protect data.⁶³ These policies are, in turn, informed by a variety of public interest values and concerns; in the first place, of course, freedom of expression and information – the fundamental freedom enshrined in the European Convention on Human Rights (art. 10 ECHR) and the EU Charter (art. 11).

As case law and doctrine regarding the Convention demonstrate, this fundamental freedom is to be interpreted broadly. It is phrased in media-neutral terms and thus applies to old and new media alike. The term 'information' (in French: 'informations') comprises, at the very least, the communication of facts, news, knowledge and scientific information. It also, undoubtedly, extends to syntactic data; the scope of art. 10 ECHR is not limited to (semantic) speech, but extends to the means used for communication purposes. To what extent art. 10 ECHR protection extends to *commercial* speech has been a matter of some controversy. However, the European Court of Human Rights has made it clear that information of a commercial nature is indeed protected, albeit to a lesser degree than political speech.⁶⁴

Art. 10 ECHR prevents states from creating restrictions on the free flow of information unless such restrictions 'are prescribed by law and are necessary in a democratic society ... for the protection of the ... rights of others'. From this perspective, data and information must flow freely, uninhibited by property rights or other state-created restrictions,

⁶² See generally P. B. Hugenholtz, *Auteursrecht op informatie* (n. 36).

⁶³ Note that patent law also does not protect data; see Drexl (n. 12), Max Planck Institute for Innovation & Competition Research Paper No. 16-13 at 24–25 (2016).

⁶⁴ See e.g. *Hertel v. Switzerland*, ECHR 25 August 1998, Publications of the ECHR, Reports 1998-VI.

unless a compelling societal need for protection ('necessary in a democratic society') can be established. Freedom of expression and information, in other words, makes IP rights in data the exception to the default rule of freedom.⁶⁵

This brings us back to the question of expediency. The EU legislature would bear the burden of proving that a property right in machine-generated data is a socially and economically justifiable ('necessary') interference in the freedom of European citizens and companies to freely access and reutilize machine-generated data. In light of the abundant praise in political literature of 'big data'⁶⁶ and big data mining as drivers of progress and prosperity, and the absence of convincing evidence supporting a property right in machine-generated data, this burden of proof seems insurmountable.

In particular, freedom of expression and information militates strongly against any new IP right that would restrict scientists' access to data – a freedom that the EU legislature expressly wishes to preserve as regards 'text and data mining' by non-commercial research institutions. Note that this freedom finds additional support in art. 13 of the EU Charter ('The arts and scientific research shall be free of constraint. Academic freedom shall be respected'). Another area where a data right would patently conflict with freedom of expression and information is journalism, where mining data has become an essential tool for investigative reporting.⁶⁷ Surprisingly, data mining for journalistic purposes seems to be overlooked in the proposed TDM exception of the DSM Directive.

A second over-arching policy consideration underlying IP law's reluctance to protect data per se is freedom of competition (enshrined in art. 16 of the EU Charter as the 'freedom to conduct a business'). This economic

⁶⁵ See (for copyright) European Court of Human Rights 10 January 2013, No. 36769/08, *Ashby Donald and Others v. France*, ECLI:2013:0110JUD00367690.

⁶⁶ OECD, *Data-Driven Innovation: Big Data for Growth and Well-Being* (n. 1); see also European Commission, 'A Digital Single Market Strategy for Europe', Communication from the Commission to the European Parliament, The Council, The European Economic and Social Committee and the Committee of the Regions, 6 May 2015, COM(2015) 192 final, p. 14: 'Big data, cloud services and the Internet of Things are central to the EU's competitiveness'. I. Hargreaves et al., *Standardisation in the Area of Innovation and Technological Development, notably in the Field of Text and Data Mining*. Report from the Expert Group to the European Commission (2014), doi: 10.2777/71122.

⁶⁷ See *Dammann v. Switzerland*, ECHR 25 April 2006, no. 77551/01. The Court opines that 'the gathering of information was an essential preparatory step in journalism and an inherent, protected part of press freedom'.

freedom traditionally sets limit to intellectual property rights and is one of the rationales underlying the idea/expression dichotomy. As the abundant literature on the economic potential of ‘big data’ demonstrates, machine-generated data are both input and output to innovative manufacturing processes and value-added services, and thus a major driver of economic growth. This calls for measures promoting access to data and fostering data mining rather than commodification of data by creating property rights in data. Unless equipped with wide-ranging exceptions and safety valves, introducing a new property right in data might create undesirable data monopolies that could impede, rather than foster, competition in this rapidly evolving European ‘data market’.⁶⁸ At the global level, introducing data property rights in the EU might well lead to anti-competitive distortions as well, in cases where European data users are obliged to purchase licenses for usage of data freely available to their competitors in the United States.

Finally, a novel data right would also create new barriers to the freedom of services, one of the four freedoms of the EU Internal Market. In its Communication on ‘Building a European Data Economy’, the European Commission interprets this freedom, together with the freedom of establishment, as implying a ‘principle of free movement of data within the EU’.⁶⁹ It is hard to see how a novel property right in machine-generated data would square with this freedom.

V. CONCLUSION

This chapter makes the case against introducing a data property right. As we have seen, there are abundant reasons to reject this idea. A ‘data right’ in machine-generated data would ride roughshod over the existing system of intellectual property. It would violate one of the main maxims of IP that data per se are ‘free as the air for common use’, and that only creative, innovative or other meritorious investment is protected. It would corrode IP’s existing mechanism of incentives by creating an under layer of rights that automatically protects all data produced with the aid of machines. This parallel layer of rights would, most likely, extensively overlap with other IP regimes, and thus create undue impediments for the

⁶⁸ Max Planck Institute for Innovation and Competition, ‘Data Ownership and Access to Data’, Position Statement of the Max Planck Institute for Innovation and Competition, Max Planck Institute for Innovation & Competition Research Paper No. 16-10 (2016) 2.

⁶⁹ European Commission, ‘Building a European Data Economy’ (n. 4), at 7.

exploitation of existing IP rights, such as copyright and database right, and endanger user freedoms guaranteed under these regimes. It would also give rise to gross legal uncertainty, since the 'velocity' of real-time data generation makes it difficult, or even impossible, to circumscribe its subject matter, scope of protection and ownership. More generally, a property right in machine-generated data would contravene freedom of expression and information, and pose new obstacles to freedom of competition, freedom of services and the 'free flow of data'.

The great promise of big data – for the economy, for science, for society at large – is that this resource can be freely exploited. Introducing a 'data right' preventing unauthorized access to big data would directly contradict this. Indeed, it is hard to understand how the proposed new right would square with the text and data mining exemption proposed by the European legislature in the current EU copyright reform package.

If, as the European Commission rightly believes, 'big data, cloud services and the Internet of Things are central to the EU's competitiveness',⁷⁰ one would have expected supporters of a novel data producer's right to present powerful and convincing arguments in support of this revolutionary proposition. So far, the case for a property right in machine-generated data has yet to be made. As Professor Drexl and others have pointed out, the existing toolkit of trade secret protection, contract and technological protection measures offers data producers ample means of securing de jure or de facto exclusivity.⁷¹ Rather than wasting time and effort on inventing a data producer's right, the focus of the European Commission's possible interventions should be on fostering access to big data.⁷²

In sum, there are innumerable reasons for the European Commission not to go down this road. If nothing else, Europe's recent experience with the sui generis database right should give reason for extreme caution. In 2005, less than ten years after it was introduced at EU level, the European Commission published its first (and still only) review of the Database Directive, a remarkably self-critical assessment. According to the Commission, '[t]he economic impact of the 'sui generis' right on database production is unproven. Introduced to stimulate the production of databases in Europe, the new instrument has had no proven impact on

⁷⁰ DSM Strategy (n. 66), at 14.

⁷¹ Drexl (n 12), Max Planck Institute for Innovation & Competition Research Paper No. 16-13 at 66 (2016).

⁷² See Drexl (n 12), Max Planck Institute for Innovation & Competition Research Paper No. 16-13 at 41 et seq. (2016); European Commission, Staff Working Document (n 4), at 36 et seq.

the production of databases'.⁷³ The Commission's report also suggests that the sui generis right has not helped the European industry to overcome its productivity gap vis-à-vis the United States.⁷⁴ It points to several other deficiencies of the sui generis right, such as its uncertain contours, and its proximity to a property right in data that might negatively affect innovation and growth. The report juxtaposes the legal situation in the EU with that in the United States, where since the Supreme Court's landmark *Feist* decision⁷⁵ no legal protection for 'sweat of the brow' based databases exists. Nevertheless, as the Commission wryly observes, 'there has been a considerable growth in database production in the US, whereas, in the EU, the introduction of 'sui generis' protection appears to have had the opposite effect'.⁷⁶

The 2005 evaluation report concludes by offering four possible ways forward: (1) repeal the whole Directive; (2) withdraw the sui generis right; (3) amend the sui generis to clarify its scope; and (4) maintain the status quo. Despite these harsh conclusions, the database right has yet to be amended or repealed. The problem is that removing (parts of) a directive is, politically and legislatively, even more complex than substantive harmonization. Repealing the database right would require a new directive not only rescinding major parts of the existing Directive, but also – rather absurdly – instructing Member States to *abolish* sui generis database protection. Unsurprisingly, the only option that has materialized from the Commission's assessment is number 4: 'do nothing'.

The lessons of the EU's database experiment⁷⁷ are not to be forgotten. Introducing a novel right of intellectual property should of course never be done in the spur of the moment. Any new right should be contemplated only after conducting thorough economic, evidence-based research that demonstrates a real need for the right and predicts its consequences for information markets and society at large. Assuming a convincing case in support of the right can indeed be made, this should then be followed by systematic legal analysis of the new right's contours and scope, and of

⁷³ European Commission, 'First evaluation of Directive 96/9/EC on the legal protection of databases', DG Internal Market and Services Working Paper, Brussels, 12 December 2005, p. 5.

⁷⁴ *Ibid.*, 22–23.

⁷⁵ *Feist Publications, Inc. v. Rural Telephone Service Co.*, 499 U.S. 340 (1991).

⁷⁶ European Commission, 'First evaluation of Directive 96/9/EC on the legal protection of databases' (n. 73), at 24.

⁷⁷ S. M. Maurer, P. B. Hugenholtz & H. J. Onsrud, 'Europe's Database Experiment', 294 *Science* (2001) 789–790.

Against 'data property'

71

its impact on the existing system of IP. The two-tiered structure of the Union does not allow for legal experimentation at the EU level. Like the database right, a 'data right' would be here to stay – a most unwelcome guest in the house of European intellectual property.