



**HAL**  
open science

# The paradoxes of open data and how to get rid of it? Analysing the interplay between open data and sui-generis rights on databases

Primavera de Filippi, Lionel Maurel

## ► To cite this version:

Primavera de Filippi, Lionel Maurel. The paradoxes of open data and how to get rid of it? Analysing the interplay between open data and sui-generis rights on databases. *Columbia Science & Technology Law Review*, 2015, 23 (1), pp.1-22. <10.1093/ijlit/eau008>. <hal-01265200>

**HAL Id: hal-01265200**

**<https://hal.science/hal-01265200v1>**

Submitted on 2 Feb 2016

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization



# The paradoxes of open data and how to get rid of it? Analysing the interplay between open data and sui-generis rights on databases

Primavera De Filippi\* and Lionel Maurel†

## ABSTRACT

Open Data is an important public policy that contributes to achieving greater transparency and broader access to information, more citizen participation and engagement, while also supporting innovation and economic growth. The pace at which the Open Data movement is spreading in different fields of endeavour can be taken as an illustration that society is evolving towards greater openness, transparency and accountability. Yet, several constraints and legal uncertainties subsist beyond the façade of Open Data. This article investigates different layers of rights that regulate the use and re-use of data: from the copyright vesting in the content and/or structure of a particular dataset, to the sui-generis right protecting against the substantial reproduction and/or extraction of the content of a database. The objective is, ultimately, to illustrate the conflictual relationship that subsists between the underlying principles of Open Data, which purports to promote the free use and re-use of information, and the underlying legal system, whose provisions are increasingly relied upon to establish an exclusive right on public sector information.

**KEYWORDS:** open data, sui-generis right, copyright, public sector information, transparency

## INTRODUCTION

In a digital era, characterized by an increasing amount of data and information, the potential benefits that can be derived from the widespread dissemination, exchange and re-use of data are growing at an exponential rate. Defined by Neelie Kroes—European Commissioner for Digital Agenda—as the ‘new gold’ of the digital era,<sup>1</sup> data are regarded as a new economic and social resource,<sup>2</sup> the

\* CERSA / CNRS / University Paris II. Berkman Center for Internet & Society at Harvard Law School. E-mail: pdefilippi@gmail.com

† La Quadrature du Net; librarian, blogger and internet freedom activist. E-mail: calimaq@gmail.com

1 Such was the definition provided by Neelie Kroes in the Opening Remarks of the Press Conference on Open Data Strategy, held in Brussels, on 12 December 2011. The speech is available at <<http://europa.eu/rapid/pressReleasesAction.do?reference=SPEECH/11/872&type=HTML>> accessed 1 October 2014.

2 Berners-Lee T and others, ‘Tabulator: Exploring and Analyzing Linked Data on the Semantic Web’ (2006) *Proceedings of the 3rd International Semantic Web User Interaction Workshop* (Vol. 2006).

new ‘raw material’ which lies at the heart of social development, research and innovation.<sup>3</sup>

Yet, as opposed to material resources which are necessarily scarce and exclusive, data are a resource which is not affected by scarcity. Given that it can be reproduced and disseminated at virtually no cost, data can be used simultaneously by a large number of people, and potentially re-used for a variety of different applications. Besides, the value of data ultimately increases as it gets more widely disseminated and more tightly connected.<sup>4</sup>

Today, Open Data is considered to be an important public policy that could significantly contribute to promoting transparency and broader access to information, encouraging citizen participation and civic engagement, while also supporting innovation and economic growth.<sup>5</sup>

Yet, a legal definition of what actually constitutes ‘Open Data’ has not yet been provided in either statutes or case law. The principles of Open Data first emerged from the work undertaken by experts and practitioners from civil society or non-governmental organizations. The Open Government Data<sup>6</sup> initiative, launched in 2007, was the first to establish a list of eight principles describing the procedure of opening up public sector information—many of which are actually concerned with legal aspects.<sup>7</sup> Later, the Open Knowledge Foundation (OKF)<sup>8</sup> elaborated the ‘open definition’.<sup>9</sup> Inspired from the definition of Open source, the open definition identifies a list of 11 principles characterizing Open Knowledge—which can also be used to assess the ‘openness’ of data released by public sector institutions.<sup>10</sup> In 2010, it was the turn of Tim Berners-Lee who proposed a means to assess the quality of Open Data in five simple steps—one of which refers to the legal status of such data, which must be ‘available on the web (in whatever format) but with an open license’.<sup>11</sup>

3 Arzberger P and others, ‘Promoting Access to Public Research Data for Scientific, Economic, and Social Development’ (2004) 3 *Data Sci J* 135.

4 Bizer C, Heath T & Berners-Lee T, ‘Linked Data—the Story So Far’ (2009) 5 *IJSWIS* 1.

5 Davies T, ‘Open Data, Democracy and Public Sector Reform’ (2010), A look at open government data use from [data.gov.uk](http://data.gov.uk). Open Data Institute, UK.

6 A 2007 working group held in Sebastopol, California developed the 8 Principles of Open Government Data. These principles have become the de facto starting point for evaluating openness in government records. The principles define open government data as public data which is complete, primary, timely, accessible, machine processable, non-discriminatory, non-proprietary and license-free, and when compliance is reviewable. For more details, see <http://www.opengovdata.org/> accessed 1 October 2014.

7 Open Government Data Group, ‘8 Principles of Open Government Data’, 8 December 2007 <http://www.opengovdata.org/home/8principles> accessed 1 October 2014.

8 The Open Knowledge Foundation is a non-profit organization that promotes open knowledge, including open content and open data. It was founded on 24 May 2004 in Cambridge, the UK. More information on <http://okfn.org>

9 The Open Definition sets out principles that define ‘openness’ in relation to data and content. It makes precise the meaning of ‘open’ in the terms ‘open data’ and ‘open content’ and thereby ensures interoperability between different pools of open material. More information is available at <http://opendefinition.org> accessed 1 October 2014.

10 The full text of the Open Definition can be read at <http://opendefinition.org/okd/> accessed 1 October 2014.

11 Tim Berners-Lee, *Linked Data*, 18 June 2009 <http://www.w3.org/DesignIssues/LinkedData.html> accessed 1 October 2014.

According to these principles, Open Data seems to imply that the data held by public sector institutions be made available to the public—ideally on the web—under specific licenses allowing for anyone to freely re-use these data, including for commercial purposes.

Several States (such as USA, the United Kingdom, Australia, Austria, France) have been inspired by these principles and eventually decided to adopt more or less stringent Open Data policies, which have been to a certain degree accounted for by the law—such as, most notably, the new European Directive on the re-use of Public Sector information adopted by the European Parliament in April 2013.<sup>12</sup> Finally, even if the document does not have direct normative effect, the Open Data Charter<sup>13</sup> endorsed by the G8 on June 2013 further acknowledges and underlines these principles.

The pace at which the Open Data movement is spreading in many different fields of endeavour can be taken as an illustration that society is evolving towards greater openness and accountability. In the past few years, the opportunities that Open Data offer to both public and private institutions have become more and more apparent. In terms of transparency, Open Data allows for the public sector to establish a greater level of trust with its citizens, to promote democratic participation and civic engagement.<sup>14</sup> Open data also establish a climate of cooperation and collaboration—rather than of competition—between the public sector, the private sector and civil society.<sup>15</sup> Finally, in terms of innovation, Open Data constitutes an opportunity for the State to delegate some of its less crucial tasks to the private sector, to complement the public services offered by the State with new and innovative applications that will benefit society as a whole.<sup>16</sup>

This article does not, however, intend to analyse the benefits of Open Data, but rather to explore the paradoxes it conceals. Indeed, while digital technologies facilitated the practices of data mining and collection,<sup>17</sup> they also jeopardized the economic interests of large database producers to the extent that electronic datasets have become more vulnerable to copying and free-riding.<sup>18</sup> In Europe, intellectual property laws have therefore evolved as well, to provide an additional layer of protection for databases by means of a new *sui-generis* right.<sup>19</sup> In this regard, it is worth

12 According to the former PSI Directive from 2003, every Member State had the ability to select what kind of datasets should be freely reusable, and whether or not the reuse should be subject to the payment of a fee. The new PSI Directive from 2013 introduced an actual right for every citizen to use and reuse public sector information, with more restrictive conditions for the establishment of a fee.

13 See <<https://www.gov.uk/government/publications/open-data-charter/g8-open-data-charter-and-technical-annex>> accessed 1 October 2014.

14 Lathrop D and Ruma L, *Open Government: Collaboration, Transparency, and Participation in Practice* (O'Reilly Media 2010).

15 Conradie P, Mulder I and Choenni S, 'Open Data: Exploring the Release of Public Sector Information through Co-creation' (2010). *Engineering, Technology and Innovation (ICE), 2012 18th International ICE Conference on* (pp. 1-10). IEEE.

16 Atkinson R and others, 'Innovation Policy on a Budget: Driving Innovation in a Time of Fiscal Constraint' (2010) Working paper, available at SSRN 1722870.

17 Han J, Kamber M and Pei J, *Data Mining: Concepts And Techniques* (Morgan Kaufmann 2006).

18 Lemley MA, 'Property, Intellectual Property, and Free Riding,' (2004) 83 Tex L Rev 1031.

19 Brown MM, Bryan RM and Conley JM, 'Database Protection in a Digital World (1999) 6 Rich J L Tech 2.

considering whether we are indeed moving towards the establishment of a truly ‘open society’<sup>20</sup> or whether this openness is merely a reaction to the growing restrictions in the underlying code of society?

## BEHIND THE FAÇADE OF OPEN DATA

### The backstage of Open Data: a complex network of rights

Although regarded by many as an indicator for a new paradigm shift in society—which is evolving into an increasingly ‘open society’—ie a society where public data and information are truly open<sup>21</sup>—we regard Open Data mostly as a consequence of growing uncertainty and legal constraints.

The purpose of this section is to illustrate the opposition that subsists between the underlying principles of Open Data and the various legal norms regulating the use and re-use of data. Indeed, at the European level, we can identify at least four layers of rights that interact and complement (or supplement) each others: (i) copyright for the protection of original databases, (ii) sui-generis rights for the protection of non-original databases, (iii) the right of EU citizens to access and re-use public sector information, and (iv) user’s rights and obligations provided by the licenses. The intersection of these rules created a complex system of rules and exceptions—with some exceptions to these exceptions—that generated considerable legal uncertainty.

### *Copyright protection for original databases*

As intangible assets, European law<sup>22</sup> recognizes that databases can be protected like many other creative works under copyright law, but only on the condition that they can be regarded as original creations. Yet, although certain databases producers may benefit from the set of exclusive rights provided by the copyright regime, copyright protection only applies to the structure of the database, and not to the data it contains—which remain, in principle, in the public domain.

As regards databases, originality is difficult to assess because it is often not enough for a database to be a mere compilation of data to qualify as an original database. Courts have required that ‘originality’ be understood as ‘the imprint of the personality of the author’<sup>23</sup>—a classic criterion for many works of authorship, which is, however, difficult to implement in the case of databases. One must therefore consider whether the database producer has indeed developed an original architecture, specific ergonomics or search criteria. Originality may also result from the selection or

20 Defined from an epistemological perspective by the British philosopher Karl Popper as a society in which the government is purported to be responsive and tolerant, and political mechanisms are said to be transparent and flexible. See Popper K, *The Open Society and Its Enemies* (Routledge 1945, reprint 2006). We refer here to the concept of ‘open society’ as a society operates in a transparent manner and where most of the data and information are disseminated freely and are available for anyone to access and reuse.

21 Fioretti M, ‘Open Data, Open Society’ (2010) A Research Project about Openness of Public Data in EU Local Administration.

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

23 ECJ case C-145/10, *Eva-Maria Painer v Standard Verlags GmbH and Others*. More information is available at <<http://curia.europa.eu/juris/liste.jsf?language=en&num=C-145/10>> accessed 1 October 2014.

arrangement of particular sets of data, the tools used to access them, the way results are presented, etc.

A recent decision of the European Court of Justice clarified how the concept of originality should be applied to databases. In a trial involving the unauthorized reproduction of football fixture lists in the English and Scottish football leagues,<sup>24</sup> the Court noted that data are not protected ‘as such’ and that, in determining whether or not a database is eligible for protection under copyright law, neither the intellectual effort nor the expertise devoted to the production of its content should be taken into account. For the ECJ: ‘a database [...] is protected by copyright law [...] provided that the selection or arrangement of the data it contains is an original expression of the creative freedom of its author.’ Conversely, the Court held that no originality could be found ‘when the constitution of the database is dictated by technical considerations, rules or constraints which leave no room for creative freedom’.

The application of copyright law to databases is thus relatively restrictive and, given the vagueness surrounding the criteria of originality, considerable uncertainty is involved in determining, on a case-by-case basis, whether a particular database is protected or not.<sup>25</sup>

#### *Sui-generis right for the protection of non-original databases*

The European Directive on the protection of databases<sup>26</sup> (the Database Directive) introduced a ‘sui-generis’ right—a new exclusive right—on the actual content of a database. Yet, just like copyright, this right does not protect data ‘as such’, it merely protects databases producers against the substantial extraction and re-use of database content, but only insofar as the collection and organization thereof required a substantial investment.

Even though the scope of the ‘sui-generis’ right does not extend to data *per se*, given the broad definition of the right, it could nonetheless have an important impact on the use and re-use of data. The Database Directive introduced, in fact, only a very limited number of exceptions, and the duration of the right is potentially perpetual, since the term of protection is extended for a period of 15 years every time a new substantial investment is undertaken by the producer.

Indeed, database producers can benefit from the legal protection granted by the ‘sui-generis’ right whenever a substantial investment is made in obtaining, verifying or presenting the contents of the database. Investment can be made of physical, human, or financial capital, and the required amount of investment—although not precisely defined under the law—is usually relatively low when one looks at the case law.<sup>27</sup> Yet, judges nonetheless specified that only the capital going in the creation

24 ECJ case C-604/10, *Football Dataco Ltd and Others. vs Yahoo! UK Ltd and Others*, more details are available at <<http://curia.europa.eu/juris/liste.jsf?language=en&num=C-604/10>> accessed 1 October 2014.

25 De Filippi P, *Copyright Law in the Digital Environment : Private Ordering and the Regulation of Digital Works* (LAP Lambert Academic Publishing 2012).

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

27 See eg ECJ case C-203/02, *The British Horseracing Board Ltd and Others v. William Hill Organization Ltd*.

and organization of the database should be taken into account, and not capital going in the production of the data themselves.<sup>28</sup>

Mere consultation of a protected database is not affected by the ‘sui-generis’ right, which only covers the ‘substantial extraction of a substantial part’ of a protected database—assessed according to both quantitative and qualitative criteria. Determining the actual meaning of these concepts is a particularly complex task, which can only be achieved on a case-by-case basis, since ‘substantiality’ (in quantitative or qualitative terms) necessarily depends on the impact that the extraction may have on the investments of the database producer.<sup>29</sup> The result is a high degree of uncertainty, which can generally only be resolved by courts. This is the reason why the ‘sui-generis’ right potentially allows database producers to exert control over the use or extraction of data ‘as such’, even if the right is not supposed to apply over data *per se*.

Moreover, the ‘sui-generis’ right does not exclude other forms of protection—such as the principle of unfair competition, misappropriation, patent law, trademark law, or more generally, contract law—which can either complement or supplement the ‘sui-generis’ protection, and which may eventually led to an excessive level of protection for databases.

#### *Public sector information—an exemption to the exclusivity regime*

In Europe, access to and use of Public Sector Information (PSI) was first regulated by Directive 2003/98/EC, establishing a common legislative framework aimed at removing barriers that might hinder the re-use of PSI throughout the European Union (EU). The Directive was subsequently amended by the new PSI Directive of 2013,<sup>30</sup> which introduced a right for every European citizen to access and re-use public sector information. This right constitutes an exception to the exclusivity regime established by the sui-generis right on databases, which must—theoretically—be waived by (almost) every public sector institution.

However, the relationship between the ‘sui-generis’ right for databases and the right to re-use public sector information is actually a complex one. Although these two rights have been articulated in such a way that the former cannot neutralize the latter, sometimes, the ‘sui-generis’ right for databases might actually go counter the right to re-use public sector information.

In some cases, for instance, public administrations—which do actually fall within the scope of the PSI Directive—decided to retain exclusive rights over their databases. This happened, in France, with Legifrance<sup>31</sup>—a large government database of legislation and case laws which is freely accessible to the public, but which remains nonetheless protected under the ‘sui-generis’ right.

Similar tensions arose in other countries, such as the Netherlands, where the law explicitly precludes public authorities from exercising their ‘sui-generis’ rights to prevent the access to or the re-use of public datasets—unless otherwise stipulated by

28 ECJ case C-604/10, *Football Dataco Ltd and Others. vs Yahoo! UK Ltd and Others*, more details available at <<http://curia.europa.eu/juris/liste.jsf?language=en&num=C-604/10>> accessed 1 October 2014.

29 See ECJ case C-545/07, *Apis-Hristovich EOOD v Lakorda*.

30 Directive 2013/37/EU of the European Parliament and of the Council of 26 June 2013 amending Directive 2003/98/EC on the re-use of public sector information.

31 <<http://rip.journal-officiel.gouv.fr/>> accessed 1 October 2014.

law.<sup>32</sup> In the Dutch ‘Landmark’ case, for instance,<sup>33</sup> the Amsterdam District Court held that the Amsterdam City Council could not rely on the database right in order to impose restrictions and/or annual licensing fees for the use of its environmental database.

Conversely, in France, an administrative court ruled<sup>34</sup> that a department (as the holder of a ‘sui-generis’ right over a departmental registry database) could rely on the ‘sui-generis’ right to preclude a private company from performing a substantial extraction and re-use of the registry’s data. Such reasoning is disturbing to say the least, since—if it were to be generalized—it could potentially jeopardize the general principle of open PSI.

And yet, this ruling is line with the wording of the PSI Directive. Indeed, the exception to the exclusivity regime introduced by the Directive allows for public administrations to impose a certain number of conditions and restrictions over the re-use of public sector information—including, but not limited to, the requirement of proper attribution, non-alteration,<sup>35</sup> and, potentially, the imposition of a fee to cover the costs of making data available to the public.<sup>36</sup> These conditions could considerably restrain the right enjoyed by every European citizen to re-use public sector information, and may—in practice—eventually drift into some sort of proprietary rights.

Finally, the exception to the exclusivity regime established by the ‘sui-generis’ right is itself subject to a series of exceptions. Most relevant is the derogatory regime granted by the PSI Directive of 2003<sup>37</sup> to cultural heritage and research institutions, which were entitled to choose the conditions under which their data could be effectively re-used—the so-called ‘cultural exception’.<sup>38</sup> Although this derogatory regime has eventually been abandoned in the amended version of the PSI Directive,<sup>39</sup> it is—in fact—still possible for cultural heritage and research institutions to bypass the general principles of re-use of public sector information<sup>40</sup> (see below for a more detailed analysis).

32 See article 8 of the Dutch Database Act.

33 ABRvS 29 April 2009, n 07/786, AMI 2009-6 (College B&W Amsterdam/Landmark; m. nt. M. Van Eechoud).

34 <[http://www.legalis.net/spip.php?page=jurisprudence-decision&id\\_article=3598](http://www.legalis.net/spip.php?page=jurisprudence-decision&id_article=3598)> accessed 1 October 2014.

35 See eg Recital 17 of the 2013 PSI Directive: ‘In some cases [...] a licence will be issued imposing conditions on the re-use by the licensee dealing with issues such as liability, the proper use of documents, guaranteeing non-alteration and the acknowledgement of source.’

36 See eg Article 6 of the 2013 PSI Directive: ‘Where charges are made for the re-use of documents, those charges should be limited to the marginal costs incurred for their reproduction, provision and dissemination.’

37 See Article 1 of the PSI Directive of 2003: ‘This Directive shall not apply to [...] (e) documents held by educational and research establishments, such as schools, universities, archives, libraries and research facilities including, where relevant, organisations established for the transfer of research results; (f) documents held by cultural establishments, such as museums, libraries, archives, orchestras, operas, ballets and theatres.’

38 Breznik M, *Culture between “L’exception Culturelle” and “Cultural Diversity”* (Culture Ltd. 2005) 15.

39 See Article 1 of the new PSI Directive of 2013.

40 Moreover, public sector information from cultural heritage or research institutions is further exempted from the obligation of being freely made available to the public to the extent that, in the context of

*Open Data licenses*

In spite of these limitations, the regime established by the PSI Directive is not inconsistent with the general principles of Open Data. Yet, relying exclusively on the legislative framework for the re-use of public sector information is often not enough for an institution to actually comply with the requirements of an Open Data approach. This is why many governments and public administrations in Europe have decided to release their data under an Open data license<sup>41</sup>—thereby renouncing both to their ‘sui-generis’ right on databases and to their ability to impose a fee for the re-use of data.

Beyond USA, the United Kingdom was the first country to produce a specific instrument with the Open Government License.<sup>42</sup> This license—inspired from the Creative Commons Attribution license (CC-BY)<sup>43</sup>—allows for the free re-use of public sector information, including for commercial purposes, and not subject to any royalty fees, but only provided that the source is always and properly mentioned. The license also expressly stipulates the waiving of all copyright or sui-generis right that the public sector institution may hold.

Inspired from this example, France subsequently developed the Licence Ouverte/Open License<sup>44</sup> strongly anchored in the French law on public sector information<sup>45</sup> which transposed the PSI Directive into national law. Like the Open Government License in the UK, the French Open License allows for the re-use of public sector information, including commercial purposes, free of charge, and with as sole requirement the necessity of always mentioning the original source from which the data have been extracted, and the prohibition of improperly tampering with the meaning of the data. The license also maintains some of the conditions resulting from the underlying law, such as the requirement of preserving the integrity of the data and not tampering with its meaning.<sup>46</sup>

It is also worth noting that some countries, such as Australia, have opted for already existing Open Data licenses rather than developing their own national license. The portal data.gov.au uses, for instance, the Creative Commons Attribution license (CC-BY) for the dissemination of all datasets. Using a pre-existing license has the advantage of being more simple and of reducing the risks for an excessive proliferation of licenses. It might, however, raise a number of concerns insofar as the Creative Commons licenses were not specifically designed for the licensing of ‘sui-generis’ rights in non-original databases, but only for the licensing of copyright. While

academia, the copyright of professors and researchers automatically brings their work outside of the general regime of public-sector information.

41 In this regard, see the Open Definition of the Open Knowledge Foundation which specifies what can actually qualify as an ‘open’ license <<http://opendefinition.org/okd/>> accessed 1 October 2014.

42 <<http://www.nationalarchives.gov.uk/doc/open-government-licence/>> accessed 1 October 2014.

43 <<http://creativecommons.org/licenses/by/3.0/>> accessed 1 October 2014.

44 <[http://www.etalab.gouv.fr/pages/Licence\\_ouverte\\_Open\\_licence-5899923.html](http://www.etalab.gouv.fr/pages/Licence_ouverte_Open_licence-5899923.html)> accessed 1 October 2014.

45 Loi n° 78-753 du 17 juillet 1978 portant diverses mesures d’amélioration des relations entre l’administration et le public et diverses dispositions d’ordre administratif, social et fiscal, as amended as of today <<http://www.legifrance.gouv.fr/affichTexte.do?cidTexte=JORFTEXT00000339241>> accessed 1 October 2014.

46 De Filippi P, Bourcier D, ‘« Open Data » : l’ouverture des données’ (2014) in *La Semaine Juridique Administrations et collectivités territoriales (JCPA, Lexis-Nexis)*, n.28, Janvier 2014.

the older versions of these licenses might not be simple to articulate with the right to re-use of public sector information, the newest version the Creative Commons licenses<sup>47</sup> should, however, properly account for the ‘sui-generis’ rights on databases, so as to be more compliant with the licensing databases.

Open Data licenses contribute to clearly indicating that particular datasets are effectively and freely reusable. Yet, a comprehensive analysis of the overall legal framework shows that these licenses only intervene at the end of the supply chain, as the last ‘layer’ regulating the use of public sector information. Indeed, if the first layer of protection is established by the ‘sui-generis’ right for non-original databases, the right to re-use public sector information constitutes a particular exception to this right—implemented, ‘in fine’, by the contractual terms and conditions of Open Data licenses which stipulate the actual conditions for re-use.

The superposition of a variety of rights and obligations is likely to result in the establishment of a very complex and intricate legal framework regulating the use and re-use of public sector information—which is far removed from the original principle stipulating that data ‘as such’ should not be subject to any kind of appropriation.

### The offstage of Open Data

The most controversial aspect of Open Data is that the laws requiring public sector information to be made freely available to the public could be (and have been) used as a potential source of new exclusive rights.

Indeed, it is important to remember is that the ‘sui-generis’ right on databases does not actually cover data ‘as such’, but only the substantial taking of data whose collection and organization into a database (as opposed to the production thereof) required a substantial investment. This means that most of the data produced or stored by public sector institutions are not—in theory—protected by any underlying right.

In practice, however, the right to re-use public sector information (enshrined in the PSI Directive) has been ‘hijacked’ by public sector institutions pretending to have been endowed with a new exclusive right over public sector information (i). Indeed, the obligation to make data freely available to the public has been construed—albeit controversially—as implicitly acknowledging the subsistence of an exclusive right over data (ii). This interpretation has found the strongest support in the context of cultural heritage and research institutions, whose data are subject to a specific derogatory regime (iii). As a result, the Open Data approach to public sector information may—somewhat paradoxically—ultimately endorse an exclusive form of control over data *per se* (iv).

#### *From a re-use right to a new form of exclusive right*

The PSI Directive of 2003 first introduced a ‘right’ to re-use public sector information. Yet, the text of the Directive actually left it up to the member States to decide what kind of information shall or shall not be freely reusable. Only a few countries, such as France, have actually gone further in transposing the Directive, by implementing a general principle of re-use for all (or most) public sector information.

47 <<http://wiki.creativecommons.org/4.0>> accessed 1 October 2014.

This general principle has now been endorsed in the new version of the PSI Directive adopted in 2013, which actually recognizes a right for all European citizens to freely re-use public sector information.

Compared to the Database Directive, the approach adopted by the PSI Directive is revolutionary, to the extent that it does not introduce a new property right over public sector information (in favour of the public administration) but rather establish a new right for the benefit of all European citizens. As the right to re-use has become the rule, the restriction has become the exception. Yet, this innovative approach to data re-use is, in fact, relatively shallow, and actually fall short in overturning the ‘proprietary paradigm’ enshrined by the ‘sui-generis’ right.

In practice, even in those countries—such as France—which recognized the right to re-use public sector information as a general principle, the actual implementation of this right relies on specific mechanisms that bear a strong resemblance to an exclusivity right. Although public sector institutions cannot generally oppose the authorization requests made by citizens or corporate entities,<sup>48</sup> the directive nonetheless allows them to lay down specific conditions for the re-use of information, such as the payment of a fee<sup>49</sup> and/or the acceptance of a contractual license.<sup>50</sup>

While public sector bodies have not been (officially) granted an exclusive right to authorize or prohibit the re-use of their data—which is the prerogative of the exclusive rights enshrined in copyright law—in practice, conditions regulating the use and re-use of public sector information could have very similar effects. For example, the requirement to always mention the original source of data<sup>51</sup> does, in practice, implement some kind of paternity right over data; whereas the prohibition to alter or distort the meaning of public sector information<sup>52</sup> can be assimilated to some kind of integrity right.<sup>53</sup> The introduction of licenses and royalties for re-use constitute, however, the strongest restriction on the ‘right’ to re-use public sector information. Indeed, from the perspective of data re-users, what is the difference between paying for a license to re-use data from database protected by the ‘sui-generis’ right and paying public sector institutions for obtaining a license to re-use of their data ? Even

48 See Article 4 of the 2013 PSI Directive (Requirements applicable to the processing of requests for re-use).

49 See Article 6 of the 2013 PSI Directive (Principles governing charging), which stipulates that: ‘Where charges are made for the re-use of documents, those charges should be limited to the marginal costs incurred for their reproduction, provision and dissemination.’

50 See Article 8 of the 2013 PSI Directive (Licences), which stipulates that: ‘Public sector bodies may allow for re-use without conditions or may impose conditions, where appropriate through a licence. These conditions shall not unnecessarily restrict possibilities for re-use and shall not be used to restrict competition.’

51 See eg Article 12 of the French law on public sector information (Loi n° 78-753 du 17 juillet 1978), stipulating that: ‘Unless approved by the administration, the reuse of public information is subject to the condition that [...] their sources is properly indicated, together with the date they have been most recently updated.’

52 See article 12 of the French law on public sector information (Loi n° 78-753 du 17 juillet 1978), stipulating that ‘Unless approved by the administration, the reuse of public information is subject to the condition that they have not been distorted and that their meaning has not been altered.’

53 A right which might actually raise important challenges when applied to data, since is difficult to determine what constitutes an ‘alteration’ in the context of data which is meant to be manipulated for reuse.

though, legally speaking, there is no exclusive right on public sector information, in practice, the results are pretty much the same.

While the new PSI Directive significantly restricted the opportunities for governments to set the rates of these royalties, it did not, however, introduce a general principle of gratuity. Accordingly, even though citizens have been granted a right to re-use public sector information, this right appears to be, in fact, more like a *trompe l'oeil* (an illusion, a deceiving technique that fools the eye). In practice, very few citizens can actually appreciate the existence of such 'right', whereas the others only experience restrictions which are very similar to an exclusive right.

### *The emergence of a right over data per se*

Within the 1996 Database Directive, a clear distinction is made between the copyright over the structure of a database and the 'sui-generis' right over the use or extraction of the database content. Neither copyright nor the 'sui-generis' right do, however, provide protection over data *per se*—which cannot be subject to direct appropriation and remains (theoretically) in the public domain.

Through case law, courts have confirmed the fact that the 'sui-generis' is not concerned with data themselves, since the investment undertaken for the production of the database content should not be taken into account in determining the existence of this right.<sup>54</sup> Of course, the prohibition on extracting a substantial amount of data from a protected database introduced some form of control, on the basis of an exclusive right, over the re-use of data—but at least the principle stipulating that data shall belong to the public domain was preserved.

With the introduction through the PSI Directive of a new right to re-use public sector information, a shift has occurred, gradually revealing a new form of exclusive right directly related to data themselves. This shift first emerged on a semantic level, for the right 'to re-use' public sector information can easily be confused (sometimes willingly, sometimes not) with a right 'over' public sector information. This misconception of the right is useful to the extent that it enables public sector institutions to rely upon the legal regime established by the PSI Directive to justify the subsistence of a right 'over' public sector information.

As illustrated above, the obligation to mention the source and not to alter the data, or the requirement to pay a fee for the re-use of public sector information constitute the 'hidden' resurgence of an exclusivity right—which, as opposed to the 'sui-generis' right, whose restrictions relate to the substantial extraction or re-use of data, the exclusive right emerging from the PSI Directive actual relates to data *per se*.

The paradox inherent the PSI Directive is, therefore, that, as an attempt to expand the rights of European citizens to re-use public sector information, the Directive created a framework which provides public sector bodies with an unprecedented control over data. Obviously, whenever the public administration actually adopted an Open Data approach, citizens have the ability to re-use public data according to satisfactory

54 See eg the landmark ECJ case C-604/10, *Football Dataco Ltd and Others v Yahoo! UK Ltd and Others*, more details available at <<http://curia.europa.eu/juris/liste.jsf?language=en&num=C-604/10>> accessed 1 October 2014.

conditions—yet, to achieve such a result, one must first ‘open’ the barriers induced by regime stipulated by the PSI Directive.

While the strength of these barriers depends on the sector of activity, it is in the context of culture heritage and research institutions that emergence of a right ‘over’ public sector information is the most apparent.

*Cultural and research data: the exception that ‘affirms’ the rule*

According to the PSI Directive of 2003, public sector information produced or collected by cultural heritage or research institutions could avail themselves of a derogatory regime as they were—by definition—excluded from the general principles of re-use.<sup>55</sup> Even in a country like France, where the right to re-use public sector information has been endorsed as a general principle, a special regime has nonetheless be preserved in favour of cultural and research institutions—the so-called ‘cultural exception’.<sup>56</sup> Notwithstanding the general principle of re-use, this derogatory regime allows these institutions to autonomously establish the terms and conditions under which their data can effectively be re-used.<sup>57</sup>

It is generally said that the exception confirms the rule. Yet, instead of being regarded as an exception to the general principle of re-use of public sector information, the derogatory regime granted to cultural heritage and research institutions has been interpreted by some of these institutions as implicitly conferring them the right to decide what can or cannot be done with their data, regardless of whether or not they are actually entitled to a ‘sui-generis’ right over such data. In a sense, this interpretation of the cultural exception can be regarded as the most advanced stage of transgression: an attempt to move away from the original intent of the PSI Directive towards the establishment of a new exclusive right over data *per se*—and that is, therefore, much more restrictive than the rights they would have been granted with under the ‘sui-generis’ protection for databases.

In France, the cultural exception created a serious deadlock, as various cultural institutions interpreted the law in their favour, as giving them the opportunity to prohibit the re-use or to prevent the commercial re-use of their data. Most notable is the case of the departmental archives of Cantal that did not want their data to be re-used by private actors for the purpose of providing commercial services.<sup>58</sup> The

55 See Article 1 of the 2003 PSI Directive (Subject matter and scope):

This Directive shall not apply to [...] (e) documents held by educational and research establishments, such as schools, universities, archives, libraries and research facilities including, where relevant, organizations established for the transfer of research results; (f) documents held by cultural establishments, such as museums, libraries, archives, orchestras, operas, ballets and theatres.

56 Gare E, ‘Exception culturelle: les archives aptes à l’open data’ *Archimag* (September 2012) 257, p. 4.

57 See Article 11 of French law on public sector information (Loi n° 78-753 du 17 juillet 1978).

58 The most important case of misuse of the ‘cultural exception’ is found in the dispute between NotreFamille.com (a commercial company providing genealogical services to the public) and the Departmental Archives of Cantal, which refused to provide the census records they maintain on the ground that the cultural exception provided for in the French law on public sector information (Loi n° 78-753 du 17 juillet 1978) introduced a derogatory regime allowing for cultural institutions to establish themselves the conditions for reuse, including the prohibition thereof.

archives could have relied on the need to protect personal data to prohibit or regulate the re-use of such data, yet they actually decided to rely on the cultural exception as an attempt to completely neutralize the right to re-use public sector information.<sup>59</sup>

Independent authorities, and ultimately the courts had to intervene to confirm that the cultural exception could not be invoked in this way. After going through various tribunals and courts, in July 2012, the appellate court of Lyon confirmed that ‘documents and data held by public archives are subject to the general principle of re-use of public sector information enshrined in the law’.<sup>60</sup> This decision effectively brought this deadlock to an end, but it took many years to arrive at this result, during which several public institutions were able to exercise a power of control over data that should be available for all to access and re-use.

One way to preclude the emergence of a new exclusive right on data itself is to eliminate the derogatory regime granted to cultural heritage and research institutions. Indeed, as proposed by the European Thematic Network on Legal Aspects of Public Sector Information (LAPSI)<sup>61</sup> in its fifth recommendation,<sup>62</sup> cultural and research institutions should be included within the scope of the PSI Directive.

These recommendations have actually been adopted by the new PSI Directive from 2013, which reintegrated all data and documents of archives, libraries and museums into the general regime of public sector information re-use. This should—theoretically—result in the removal of the ‘cultural exception’, yet cultural and research data can now benefit from other and potentially stronger restrictions. It appears, in fact, that libraries, archives and museums have maintained the ability to decide the data that shall or shall not be freely reusable—which is equivalent to recognizing them a right to authorize or prohibit data re-use. Cultural institutions are also entitled to set higher re-use fees than other institutions, which are restricted to marginal costs. In addition, as regards the documents on which libraries, museums and archives hold intellectual property rights, these institutions may decide whether or not to allow their re-use—a choice which may be particularly problematic for all the information included in protected databases. Finally, it is important to note that the new PSI Directive does not eliminate the derogatory for research institutions, whose datasets could thus continue to be subject to specific conditions that do not favour re-use. All these elements contribute to the risk that the ‘cultural exception’—although theoretically removed from the text of the Directive—may well continue to

59 According to the Departmental Archives of Cantal, given that Article 11 French law on public sector information (Loi n° 78-753 du 17 juillet 1978) provides for a derogation to the general principle of data re-use for cultural institutions, departmental archives are not required to comply with any request for reuse, and—if granted—can independently set the terms and conditions regulating the use and reuse of their data.

60 See the decision of the appellate court of Lyon, clarifying the fact that the cultural exception cannot be relied up on to justify a refusal to provide access to public sector information or to prohibit the actual reutilization thereof, <[http://lyon.cour-administrative-appel.fr/media/document/CAA\\_LYON/11ly02325notrefamillemcom.pdf](http://lyon.cour-administrative-appel.fr/media/document/CAA_LYON/11ly02325notrefamillemcom.pdf)> accessed 1 October 2014.

61 <<http://www.lapsi-project.eu/>> accessed 1 October 2014.

62 <[http://www.evpsi.org/evpsifiles/lapsi\\_cultural\\_research\\_institution\\_policy.pdf](http://www.evpsi.org/evpsifiles/lapsi_cultural_research_institution_policy.pdf)> accessed 1 October 2014.

exist in practice, reinforcing the subsistence of an implicit exclusive right over data *per se*.

Hence, while the new PSI Directive of 2013 may—apparently—have abandoned the ‘cultural exception’ (since it now includes within its scope of application all public sector information held or produced by cultural and research institutions), the new Directive has nonetheless maintained several and significant exceptions that practically guarantee a derogatory status to these institutions.

### *Open Data: open, but at what cost?*

Of course, one should rejoice that many Open Data initiatives have been conducted in recent years to promote and facilitate the dissemination and re-use of data in several European countries. These approaches have shown that, despite an inadequate legal framework, pragmatic solutions can be found to properly implement the right to re-use public sector information enshrined in the European Directive. Legal instruments such as the various Open Data licenses are effective tools to regulate the free use and re-use of data, while ensuring interoperability with other datasets already available on the Internet.

However, in the end, it is important to note that—from a symbolic point of view—Open Data licenses can only effectively ‘open’ data by accepting the validity of many layers of rights, which contribute to regulating the use and re-use of information, often with a restrictive approach. Most Open Data licenses nowadays recognize the legal protection for original and non-original databases, including the ‘*sui-generis*’ right which protects the contents of these databases. The right to re-use public sector information is nothing more but an exception to the proprietary paradigm introduced by the ‘*sui-generis*’ right of databases. Besides, this right to re-use public sector information has itself has many limitations and exceptions—as we have seen in the case of cultural and research institutions. Since it does not introduce an obligation to make data freely available to the public, a true Open Data approach can only be implemented through the use of specific licenses designed to actually waive these right.

Finally, if one considers all the layers of rights accumulated over public sector information, Open Data licenses appear to be an exception to the exception to the proprietary principle of exclusive ownership, derived from the legal protection of databases. The complexity of this legal framework explains, by and large, the difficulties encountered in implementing a proper Open Data policies in various countries across Europe. Was it actually a good idea to implement an Open Data policy in this multi-layered framework? Was it not possible to find an easier way to open-up public sector information? The issue is particularly timely and relevant, as Open Data has recently become an important global policy aimed at improving State transparency and corporate governance. How can these objectives be realistically achieved in front of such complexity and legal uncertainty?

Open Data exists, but there is an important cost that comes with it: taking for granted an established system made up of several layers of rights to exert control over information. From this point of view, even if it presents considerable improvements, the new PSI Directive may not actually succeed in changing the rules of the game, as it does not change the legal framework that must necessarily be acknowledge for any effective Open Data approach to be successfully implemented. To

proceed further in the process of opening up public sector information, shall we not, perhaps, completely change the strategy, moving away from the current approach to Open Data licensing, to focus instead on a fundamental reform of the legal framework in which it is embedded?

## A NEW APPROACH TO OPEN DATA

### Promoting legal predictability by reducing legal complexity

Today, the Open Data movement is growing at an incredible pace as the benefits of Open Data are being acknowledged by both the public and private sector. As stipulated by the PSI Directive, a growing amount of data produced and/or collected by public institutions is made available to the public under a variety of Open Data licenses which facilitate the access, use and re-use of public sector information.

It is important to remember, however, that—as elaborated above—every contractual license constitute, in fact, an additional layer of regulation that sits on top of the legal framework for the protection of databases. Open Data licenses thus contribute to increasing the legal complexity of the overall framework, by introducing additional contractual provisions that stipulate new rights and obligations. The result is a complex system of norms—combining legal rules and contractual provisions—which might in fact discourage the practices of data sharing and re-use due to the additional transaction costs resulting from the legal and administrative burdens imposed on database users and producers.<sup>63</sup>

### Common licensing schemes

Providing a high degree of legal certainty as regards the rights and obligations of database producers, data providers and end-users is a precondition to the success of any Open Data approach.<sup>64</sup>

Yet, the legal framework regulating the protection of databases is made up of several layers of legal rules (copyright, neighbouring rights or sui-generis rights) and exceptions (fair use regime, public sector information regime)—whose content, scope and applicability significantly differ from one country to another. Precisely determining the extent to which a database is protected under copyright or other laws can therefore be a difficult task, especially when the database producers or contributors are located in different countries.<sup>65</sup>

To begin with, the sui-generis database right exclusively applies to database producers located in the EU. Indeed, under European law, the structure of original databases is protected under the copyright regime, while the content of both original and non-original databases is protected under the sui-generis rights to the extent that a substantial investment has been undertaken for the production thereof. Instead, in

63 Bourcier D, De Filippi P, 'Vers un nouveau modèle de partage entre l'administration et les communautés numériques' in Nicolas Matyjasik, Philippe Mazuel (eds), *Génération Y et gestion publique : quels enjeux ?*, Institut de la gestion publique et du développement économique (IGPDE 2012).

64 Höchtel J, Reichstädter P, 'Linked Open Data—a Means for Public Sector Information Management' in Kim Normann Andersen, Enrico Francesconi, Ake Gronlund, Tom M. van Engers (eds), *Electronic Government and the Information Systems Perspective* (Springer Berlin Heidelberg 2011) 330–343.

65 Fernández-Molina JC, 'The Legal Protection of Databases: Current Situation of the International Harmonisation Process' (2004) In *ASLIB Proceedings* (Vol. 56, No. 6), 325–334.

USA, databases are only eligible for protection under copyright law, insofar as their structure presents a certain degree of creative endeavour.

Besides, even amongst countries with a similar legal framework, different standards of protection apply. For instance, the authors' rights regime found in many countries of continental Europe provides legal protection only to 'original' works of authorship, whereas the copyright regime adopted in many common law countries allows for a lower threshold of originality by means of common law principles such as the 'sweat of the brow' doctrine.<sup>66</sup>

To ensure compliance with the specificities of domestic legislations, national governments and public administrations have developed their own set of licenses (drafted in their national language) with a view of pursuing similar objectives, albeit with a different mix of contractual terms.<sup>67</sup> This might lead to a process of data fragmentation—ultimately reducing the opportunities for aggregation and derivative reuse, as data released by one institution might be incompatible with that of another institution.<sup>68</sup>

This problem could (potentially) be resolved by means of a more standardized set of licenses. While international licensing schemes—such as the Open Data Commons licenses and dedications of the Open Knowledge Foundation,<sup>69</sup> or the more general set of licenses provided by Creative Commons<sup>70</sup>—have been developed with a view to provide a consistent system for the licensing of data by public or private institutions,<sup>71</sup> the terms of these licenses have, however, not been sufficiently harmonized. Many of these licenses present specific licensing terms—such as the *copyleft* or *share-alike* provisions<sup>72</sup>—which might potentially conflict with each other and preclude legal interoperability amongst different licensing schemes.

Data interoperability can therefore only be ensured by means of international collaboration aimed at establishing a consensus as regards the licenses that should—or could—be used by different institutions. This has been done, at the national level, in countries such as the UK with the Open Government license,<sup>73</sup> France with the

66 Blanke JM, 'Vincent Van Gogh, "Sweat of the Brow," and database protection' (2002) 39(4) *Am Business Law J* 645.

67 See eg the UK Open Government License, the French Licence Ouverte, the Italian Open Data License, etc.

68 Dulong de Rosnay M, 'Creative Commons Licenses Legal Pitfalls: Incompatibilities and Solutions' (2009) *IvIR Working Paper*.

69 Open Data Commons is a Open Knowledge Foundation project aimed at providing a set of legal tools to help people release and use open data <<http://opendatacommons.org>> accessed 1 October 2014.

70 Creative Commons provides a flexible range of licenses and dedications (with different degrees of protections and freedoms) for authors and right-holders to choose from. More details on the specificities of each one of these licenses or dedications can be found at <<http://creativecommons.org>> accessed 1 October 2014.

71 Dizon MAC, 'The symbiotic relationship between global contracts and the international IP regime' (2009) 4(8) *J Intell Prop L Pract* 559.

72 The copy left or share-alike provision requires that any derivative work be licensed under the same or compatible terms and conditions as the original work.

73 In 2010, the UK Government created the Open Government Licence (OGL), and public bodies can now opt to publish their Crown Copyright material under this license (although they are not obliged to). The OGL is a free, perpetual license, based on the CC-BY license from Creative Commons, without any other restrictions beyond attribution <<http://www.nationalarchives.gov.uk/doc/open-government-licence/> for more details> accessed 1 October 2014.

Licence Ouverte,<sup>74</sup> Italy with the Italian Open Data License,<sup>75</sup> Australia and New Zealand, where public sector information is automatically released under a Creative Commons license.<sup>76</sup>

Yet, in addition to the practical difficulties for a common licensing scheme to be unanimously adopted at the national and international level, there are also a few conceptual issues to address as regards the validity and effectivity of Open Data licenses. Indeed, to the extent that they rely on the subsistence of one or more underlying exclusive rights, contractual provisions designed to license these rights might ultimately result in a number of unintended consequences.

The problem arises from the fact that it is often difficult to assess whether or not a particular database is eligible for protection under the sui-generis right regime.<sup>77</sup> Besides, even after the subsistence of the sui-generis right has been established, significant analysis is required to determine the extent to which the extraction and/or reutilization of the database content will be 'substantial' enough as to actually trigger that right.<sup>78</sup> Thus, in view of the complexity and the necessary degree of subjectivity involved, neither the database producer nor the user will know, with any reasonable degree of certainty, what their corresponding rights and obligations are.<sup>79</sup>

As a general rule, a database that is not eligible for protection under any applicable law does not require any specific licensing agreement. Yet, given the uncertainty surrounding the subsistence of the 'sui-generis' right, most database producers will be tempted to regulate the use and re-use of their database's content by means of contractual provisions, regardless of whether or not the database is actually eligible for protection.<sup>80</sup> In order to avoid the risk of potential liability, users

74 The Licence Ouverte is a license created by the French Etalab committee to establish the framework for open data licensing in France. This license, published on October 2011, was designed to be compatible with the Open Government Licenses, the ODC-BY and the Creative Commons CC-BY 2.0 licenses. For more details, see <<http://ddata.over-blog.com/xxxxxy/4/37/99/26/licence/Licence-Ouverte-Open-Licence-ENG.pdf>> accessed 1 October 2014.

75 The Italian Open Data License (IODL) was launched in October 2010 as the very first official license for public data in Italy. Version 1.0 of the IODL contained a Share-Alike clause, directing producers of derivative works to release those works under similar licensing conditions. Version 2.0 of the IODL has subsequently been developed, where only attribution of the source is required. This is the version now in use on the Italian data portal [data.gov.it](http://data.gov.it)

76 Governments in Australia and New Zealand have adopted strong open licensing policies for all public sector information, released by default under the Creative Commons license CC-BY. See <<http://nzgoal.info/>> (accessed 1 October 2014) and <<http://www.ausgoal.gov.au/creative-commons/>> (accessed 1 October 2014) for more details on the Open data policy adopted in New Zealand and Australia, respectively.

77 Rowland D, 'The EC Database Directive: an Original Solution to an Unoriginal Problem?' (1997) 5 Web J Curr Legal Issues, available at <<http://webjcli.ncl.ac.uk/1997/issue5/rowland5.html>> accessed 1 October 2014.

78 Colston C, 'Sui Generis Database Right: Ripe for Review?' (2001) 3 J Inf Law & Tech, available at <<http://www2.warwick.ac.uk/fac/soc/law/lj/jilt/>> accessed 1 October 2014.

79 Grosheide FW, 'Database Protection—The European Way' (2002) 8 Wash U J L Poly 39.

80 See eg the landmark U.K. case of *The British Horseracing Board Ltd and Others v William Hill Organisation Ltd*, where a claim for infringement was dismissed by the Court—after the European Court of Justice clarified that the legal protection under the sui-generis right regime can only be sought when there has been substantial investment in collecting and organizing existing materials into a database, regardless of the investment undertaken in the production of such materials.

will, however, often be reluctant to breach the contractual provisions of the license—even when the database does not qualify for any underlying statutory protection.<sup>81</sup>

The same applies in the context of Open data licenses. A license—as open as it may be—is premised on an exclusive right (be it either copyright or the ‘sui-generis’ database right) which precludes or restrains certain usages of the underlying content or data. While these limitations might be either partially or totally waived through contractual provisions, the license does, however, implicitly endorse the subsistence of an underlying exclusive right insofar as its contractual provisions are explicitly tied to that right.<sup>82</sup> Hence, despite the intentions of the licensor, the mere fact of releasing data under an Open license might create the illusion of there being an exclusive right which, unless properly licensed, could preclude the use and re-use of the available data. Although aimed at promoting the dissemination and facilitating the use and re-use of data, Open Data licenses could, therefore, eventually discourage the practice of data sharing and exchange by introducing an additional layer of complexity and legal uncertainty.<sup>83</sup>

#### *Public domain dedication*

The complexity and the high degree of uncertainty surrounding the application of copyright<sup>84</sup> and other laws for the protection of databases—combined with the lack of international harmonization due to the difference in legal doctrines—makes it difficult to establish clear and predictable rules to regulate the use and re-use of data by contractual means. Many Open Data licenses are therefore made up of long and complicated provisions addressing the specificities of the underlying legislation.<sup>85</sup> Complex legal language might increase the transaction costs associated with the interpretation of these licenses by imposing a significant burden on users (due to, eg, legal costs or delays) and thereby contribute to ultimately discouraging the exchange or re-use of data. The problem has first been addressed by Creative Commons,<sup>86</sup> which provides, in addition to a legal contract, also a ‘human-readable’ and ‘machine-readable’ version of each license, along with a set of easily understandable logos to

81 Huse CC, ‘Database Protection in Theory and Practice: Three Recent Cases’ (2005) 20 Berkeley Tech L J 23.

82 Dusollier S, ‘Master’s Tools v. The Master’s House: Creative Commons v. Copyright’ (2006) 29 Columbia J L and Art. 29: 271.

83 Elkin-Koren N, ‘Creative Commons: a Skeptical View of a Worthy Pursuit’ (2006) Working Paper, Haifa University.

84 Gracz K, De Filippi P, ‘Regulatory Failure of Copyright Law through the Lenses of Autopoietic Systems Theory’ (2014) Int J L Inf Tech 1–33.

85 Copyright and database protection significantly differs across different legal systems, so that a license which is valid in one jurisdiction might be invalid or unenforceable in another because the contractual terms and conditions are incompatible with the underlying legislation. This problem has been addressed by Creative Commons—whose licenses were originally written for the U.S. legal system—through an initiative to ‘port’ their licenses into different jurisdictions, by translating them linguistically and adapting them to local copyright and private laws. All currently ported licenses are available at <[http://wiki.creativecommons.org/Jurisdiction\\_Database](http://wiki.creativecommons.org/Jurisdiction_Database)> accessed 1 October 2014.

86 De Filippi P, Ramade I, ‘Les licences Creative Commons: Libre Choix ou Choix du Libre?’ in Camille Paloque-Berges, Benjamin Jean, Christophe Masutti (eds), *Histoire du Libre* (Framabook, Mars 2013).

illustrate their options.<sup>87</sup> Yet, many other licenses do not provide any similar supports and remain therefore difficult to understand for many 'lambda' users.<sup>88</sup>

Many of the drawbacks inherent to Open Data licenses can be bypassed by means of public domain dedications<sup>89</sup>: simple legal tools that merely waive the rights potentially vesting in a database without introducing any of the rights and obligations generally found in a contractual license. As such, they can be drafted in a simple language and be readily understood by inexperienced users.

Public domain dedication schemes thus considerably reduce the transaction costs involved in extracting and/or re-using data from a given database. They also reduce legal uncertainty to the extent that the public domain dedication explicitly acknowledges the absence of rights in a database.<sup>90</sup> As opposed to Open Data licenses, which implicitly recognize the subsistence of a sui-generis right that will be subsequently licensed to the public, public domain dedications eliminate—at the outset—the subsistence of any potential right. This is likely to facilitate comprehension and encourage the use and re-use of data by a broader variety of people—who no longer have to comply with any specific requirements in terms of practices or formalities.

Public domain dedications can also resolve the issue of interoperability, since different dedication schemes can be used without incurring the risk of introducing incompatibilities between different datasets. As a result, both data providers and users could benefit from broader data exchange across different countries or communities, along with a higher degree of legal certainty as regards the use and re-use of such data.

As opposed to most Open Data licensing schemes, public domain dedications could, therefore, better serve the purpose of maximizing the dissemination and the overall utility of data and databases—by reducing legal uncertainty and transaction costs to the minimum.

### *Striking at the root*

Transaction costs subsist, nonetheless, insofar as most database producers are required to take action in order to waive the rights they have been automatically granted with under the law. While this situation might be regarded as acceptable by most private actors, who actually see a benefit in maintaining a certain degree of

87 Creative Commons licenses can be expressed in four different ways: (1) the Legal Code of the license which constitute the contractual agreement, (2) the Commons Deed—also known as the 'human readable' version of the license, (3) the actual Code—or the 'machine readable' version of the license—written into a format that software systems, search engines and other kinds of technology can understand, (4) the Logos (or badges) that represent graphically the various permissions and obligations incorporated into the license. For more details, see <<http://creativecommons.org/licenses/>>

88 In informatics, the expression 'lambda user' refers to the average user of a system: someone who does not have an extensive knowledge of the field, who does not seek to understand how the system works, and who generally does not benefit from any of its advanced features.

89 Most common examples of public domain dedications are the Open Knowledge Foundation ODC Public Domain Dedication and License (<http://opendatacommons.org/licenses/pddl/1-0/>) (accessed 1 October 2014) and Creative Commons CC-0 tool (<http://creativecommons.org/choose/zero/>) accessed 1 October 2014.

90 Miller P, Styles R and Heath T, 'Open data commons, a license for open data' (2008) *Proceedings of the 1st Workshop about Linked Data on the Web (LDOW2008)*.

exclusivity over the data they have collected and organized,<sup>91</sup> it is however at odds with the mission of many public sector institutions, whose objective is to maximize the dissemination of public data and research to the widest community of users.<sup>92</sup>

One opportunity to further reduce transaction costs is to address the problem ‘at the source’. Instead of imposing a burden on public sector institutions to license their rights under an open licensing scheme, the idea is to simplify the system by eliminating *ex ante*—instead of licensing *ex post*—the sui-generis protection on public sector information. Hence, rather than introducing a rule (according to which the producer of a database whose production required a substantial investment is entitled to an exclusive right over the extraction and/or reutilization of the database content) and an exception (according to which—in spite of this exclusive right—public sector information should be made freely available to the public), a better solution might be to reverse the standard of protection provided by the default under the law, turning the exception into the rule and the rule into the exception. We, therefore, suggest reforming the law so that public sector institutions are not granted a ‘sui-generis’ right to begin with—although they might, in certain special and exceptional situations, impose a series of limitations or constraints on the re-use of such data, if they can justify the need for such a derogatory regime.

A more drastic—but perhaps more effective—solution would be to entirely eliminate the ‘sui-generis’ protection for non-original databases, for both public sector and non-public sector information. This approach would ultimately restore the *status quo* (pre-1996) by eliminating the risk of establishing exclusivity over data *per se*, while ensuring that all factual data inevitably fall into the public domain—regardless of whether or not it belongs to or has been extracted from a particular database.

In addition to reducing the complexity of the legal framework, this approach would also eliminate the divergences that subsist between US and EU laws as regards the legal protection of data and databases. Indeed, although premised on the idea that the lack of a monopoly right over data would undermine the incentives to produce new and competitive databases, the efficacy of the sui-generis database right in Europe has yet to be demonstrated.<sup>93</sup> Indeed, in the USA—where no such protection exists—open and unrestricted access to a large quantity of data facilitated the development of a broad variety of cost-intensive databases and promoted the growth of many information intensive industries, which has been flourishing over the past few years.<sup>94,95</sup> By relying on antitrust laws and common law principles of misappropriation and unfair competition, USA successfully managed to protect the investment of database producers<sup>96</sup> without introducing a new exclusive right on information. It

91 Tomlin DH, ‘Sui Generis Database Protection: Cold Comfort for Hot News’ (2001) 19 Comm Law 15.

92 Reichman JH, Uhler PF, ‘Database Protection at the Crossroads: Recent Development and their Impact on Science and Technology’ (1999) 14 Berkeley Tech L J 793.

93 Lubens R, ‘Survey of Developments in European Database Protection’ (2003) 18 Berkeley Tech L J 447.

94 Maurer SM, ‘Across Two Worlds: Database Protection in the United States and Europe’ (2005) 1 Intell Prop Innov Know Based Econ at pp. 29–31.

95 Harison E, ‘Who Owns Enterprise Information? Data Ownership Rights in Europe and the US’ (2010) 47(2) Inf Manag 102–108.

96 Leaffer M, ‘Database Protection in the United States is Alive and Well: Comments on Davison’ (2006) 57 Case W Res L Rev 855.

might, thus, be worth considering whether such approach is viable and could easily be transposed into the European context.<sup>97</sup>

## CONCLUSION

The Open Data movement has brought a whole new dynamism the online and offline world. By facilitating the public disclosure of data by a variety of actors coming from both the public and the private sector, Open Data licensing schemes established the necessary legal framework for the successful deployment of innovative services and applications (developed by public institutions, market players or civil society organizations) which led to the emergence of a vibrant network of linked open data.<sup>98</sup>

One of the key successes of Open Data is that it has successfully brought the issues of transparency and accountability at the core of the public discourse. Moreover, after observing the benefits that can be derived from the adoption of an Open Data policy by several public sector institutions, a number of private actors also decided to experiment with an Open Data approach, even if they are not obliged to do so under the law.<sup>99</sup>

Hence, it might be tempting to say that, if Open Data is not the driver, it is nonetheless a strong indicator that we are moving towards an increasingly 'Open Society' characterized by increased transparency and democratic participation.

Yet, one must not forget what is going at the backstage (and at the offstage) of Open Data. The complexity of the legal system is hiding the fact that the underlying legal framework is actually evolving in a completely different direction, with stronger exclusive rights covering an increasingly broader range of subject matters—from the original structure of a database, to the factual content of a database, to end up with a new form of protection over data 'as such'.

Thus, while one might be tempted to think—at first glance—that society is evolving towards a greater degree of 'openness', the Open Data movement constitutes—in our eyes—a good indicator that something has actually gone wrong.

On the one hand, the need to rely upon a contractual license in order to 'open up' the data is an indication that the underlying legal framework is more restrictive than it should be. While the PSI Directive introduced the right to re-use public sector information, both the copyright regime and the Database Directive are likely to hinder the effective exercise of this right. Open Data licensing schemes might be regarded as a useful means to address the issue, yet they do not actually resolve the problem at the root. None of these licenses were necessary before introduction of the sui-generis right by the Database Directive, and none of them are today necessary in USA, which realized they do not need a 'sui-generis' right in order to enjoy a flourishing database industry.

97 Cardinale PJ, 'Sui generis database protection: Second thoughts in the European Union and what it means for the United States' (2006) 6 J Intell Prop 157.

98 Bizer C and others, 'Linked data on the web (LDOW2008)' (2008) Proceedings of the 17th International Conference on World Wide Web, ACM, 1265–1266.

99 See eg the SNCF in France, which although exempted—as commercial institutions—from the obligation to make all data publicly available for access and reused, nonetheless decided to experiment with an Open Data approach, in order to participate to the new ecosystem of open innovation and to bring more value and visibility to its services. For more details, see <[www.data.sncf.com/](http://www.data.sncf.com/)> accessed 1 October 2014.

On the other hand, and perhaps more critically, the use of Open Data licenses has the dangerous potential to implicitly establish an exclusive right on public sector information, since people naturally tend to assume that, where there is a license, there necessarily is an underlying right. Yet, many of the datasets produced or collected by public authorities do not actually require a substantial investment, and are thus not protected by the 'sui-generis' right on database. Releasing them under a specific Open Data license is thus likely to deceive people into thinking that they do actually have to abide to the conditions of these licenses, even if the data actually belong to the public domain. Besides, even in the case of a database that is actually eligible for protection, only the substantial reproduction and/or extraction of its content would actually trigger an infringement of the right. Yet, the subjectivity involved in assessing these operations might dissuade people from legitimately using (or reusing) a public dataset, to avoid the risk of potential liability.

We conclude, therefore, by claiming that the Open Data movement will only succeed after it has completely disappeared. Indeed, just like the 'free software' movement led by Richard Stallman, or the 'free culture' movement initiated by Lawrence Lessig—whose goal is to facilitate the free dissemination and re-use content by reducing the exclusivity granted by default under copyright law—the Open Data movement is aimed at encouraging the access to and the re-use of data by eliminating the restrictions granted under the sui-generis right regime.<sup>100</sup>

Yet, as opposed to the former two—which do not purport to eliminate copyright law, but rather to provide an alternative but complementary choice to authors, the Open Data movement distinguishes itself insofar as it was not conceived to provide a 'choice', but merely to facilitate the exercise of a 'duty'.<sup>101</sup> The case of Open Data is therefore a special case, which—we hope—will eventually lead to its own demise, as we enter a world where there will no longer be any restrictions over the use and re-use of data.

100 De Filippi P, Bourcier D, 'La double face de l'Open data' (2013) *Petites Affiches* 1.

101 Bourcier B, De Filippi P, 'L'Open Data : universalité du principe et diversité des expériences?' (2013) *La Semaine Juridique* N. 38, JCP Lexis-Nexis, 1.