UNIVERSITY OF AMSTERDAM

UvA-DARE (Digital Academic Repository)

Understanding Open Data Regulation: An Analysis of the Licensing Landscape

Giannopoulou, A.

DOI
10.1007/978-94-6265-261-3_6
Publication date
2018
Document Version
Accepted author manuscript
Published in
Open Data Exposed

Link to publication

Citation for published version (APA):
Giannopoulou, A. (2018). Understanding Open Data Regulation: An Analysis of the Licensing Landscape. In B. van Loenen, G. Vancauwenberghe, & J. Crompvoets (Eds.), *Open Data Exposed* (pp. 101-125). (Information Technology and Law Series; Vol. 30). Asser Press. https://doi.org/10.1007/978-94-6265-261-3_6

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: https://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

UvA-DARE is a service provided by the library of the University of Amsterdam (https://dare.uva.nl)

Download date:28 Sep 2021

Accepted manuscript
The final publication is available at B. van Loenen et al. (eds.), Open Data Exposed,
Information Technology and Law Series 30, https://doi.org/10.1007/978-94-6265-261-3_6

Understanding Open Data Regulation: An Analysis Of The Licensing Landscape

Alexandra Giannopoulou*

Abstract

Open data models have been developed to improve democratic participation, transparency, and innovation. There is a growing variation in open data policies and licenses that aim to maximize dissemination and reuse of data and databases. Diverse tools have been developed, tailored to the specific field of open data licensing. However, the normative framework of open data is nuanced, and intellectual property laws differ in the treatment of data and databases. The open data regulatory system created from the existing legal framework, the adopted open data policies, and the licenses developed presents an inherent complexity that impedes data reusability. The chapter investigates legal issues stemming from that system in the current open data environment. The objective is to showcase solutions to that complexity stemming from the evolution of open data licenses, policies, and ultimately, intellectual property laws.

KEYWORDS: Open data, copyright, sui generis right, public sector information, open data licenses, Creative Commons, Open Data Commons, Open Government Licence, Etalab

6.1. Introduction

In the digital era, the economic and social value of maximizing the dissemination, exchange, and reuse of data is growing. Today, open policy models are increasingly used both in the public and private sector for the dissemination, development, and maintenance of intellectual resources. According to the Open Knowledge Foundation, open data is data that "can be freely used, modified, and shared by anyone for any purpose- subject, at most, to measures that preserve provenance and openness."[1] The concept of open data emerged as a public policy concept due to the growing emphasis on the need for transparency, accountability, democratic participation, efficiency, and innovation. In 2013, the G8 Global Summit approved the Open Data Charter, stating that "the benefit of open data can and should be enjoyed by citizen of all nations."[2] However, open data does not have a uniform legal definition[3].

The Open definition created by the Open Knowledge Foundation lists a series of openness principles that can be used to assess different open data policies. The extent of freedoms granted and limitations imposed by the policies in question are justified on the basis of the main objectives of open data and the criteria set forth by each entity that defines the open data principles in question[4]. In the European Union, the Directive 2013/37/EU of 26 June 2013 amending the Directive 2003/98/EC on the re-use of public sector information[5] set the prerogatives for the development of practices and policies to make data held by the public

* Dr. Alexandra Giannopoulou, Associate researcher, Institute for Communication Sciences (ISCC), 20 rue Berbier-du-Mets, 75013, Paris, France, alex.giannopoulou@gmail.com.

[1] See the principles set out by the open definition at http://opendefinition.org/ Accessed May 2018.

[2] See the full text of the Open Data charter at opendatacharter.net Accessed May 2018.

[3] Zuiderwijk and Janssen 2014.

[4] For an overview of different open data principles: Lee 2017, p 212-213 and notes.

[5] Hereinafter the PSI Directive.

sector available for dissemination and reuse[6]. The development of public policies does not suffice in achieving open data, because the existing normative framework establishes a priori a restrictive environment of legal protection of data and databases. Thus, licenses are necessary tools in order to create open data.[7] These tools, however, are not explicitly granted in the Directive.

The PSI Directive and the development of an international open data movement has led to a cascade of open data policies on a local level including the creation or the integration of a variety of licenses and tools regulating the applicable rights. The establishment of a robust legal licensing framework constitutes "the cornerstone of open data"[8]. The aim of this chapter is to provide a comprehensive analysis of the framework surrounding open data regulation. In that respect, the study first lays out the potential legal rights in data that an open license must address. Then, it gives an overview of the most prevalent licenses and tools used on an international and national level, to provide an insight into the evolution of terms and policies by organizations such as Creative Commons and Open Data Commons or government-issued tools from the UK and France. In this regard, the study presents some key aspects of open data licensing in order to identify the major legal issues related to the application of the open data principles on the existing legal framework. Finally, it identifies the pitfalls in the current state of open data policies and licenses in order to find paths that overcome them.

6.2. The Network Of Rights Protecting Data And Databases

Many different layers of rights constitute the legal protection granted to data and databases. More specifically, applying the principles of open data to the existing legal framework invokes the rights of copyright for original content and the sui generis database rights.

6.2.1. Copyright In Data And Databases

The universal principles underpinning the protection of works of authorship stem from the Berne Convention[9]. Its flexible definition of "literary and artistic works" in article 2(1) permits the protection of different types of subject matter provided that the minimum qualitative standards are met[10]. Thus, data can only benefit from copyright protection when they pass the originality test as a literary work. Unoriginal raw data, as well as mere facts and ideas do not qualify for protection.

A collection of more data can also be considered for copyright protection separately from the copyright status of the individual component data according to article 2(5) of the Berne Convention. Protection of compilations of data is also granted by article 5[11] of the WIPO

[6] Janssen and Hugelier 2013.
[7] Even when principles of free re-use of data are added to the law in order to make licenses mostly redundant, their use is still recommended for clarity and educational purposes regarding the status of the data in question.
[8] Ubaldi 2013, p 37.
[9] The Berne Convention for the protection of literary and artistic works of 9 September 1886, last amended on 28 September 1979 is an international copyright agreement and it consists of a series of provisions ensuring minimum standards of protection for works and for authors' rights based on three founding principles: the principle of national treatment, the principle of automatic protection, and the principle of independence of protection. As of today, there are 175 contracting states adhered to the Berne Convention.
[10] See Goldstein and Hugenholtz 2013, pp. 191-197.
[11] "Compilations of data or other material, in any form, which by reason of the selection or arrangement of their contents constitute intellectual creations, are protected as such. This protection

Copyright Treaty of 1996[12] and article 10[13] of the Agreement on Trade Related Aspects of Intellectual Property Rights of 1994[14]. According to the European Directive 96/9/EC on the legal protection of databases[15], a database is defined as "a collection of independent works, data or other materials arranged in a systematic or methodical way and individually accessible by electronic or other means"[16]. The author[17] of the database holds exclusive rights over the use of the protectable expression - that is, the original selection or arrangement - of the database.

The European Court of Justice[18] expanded the application of the originality standard across all types of works of authorship in the Infopaq[19] decision by determining that "it is only through the choice, sequence and combination of those words that the author may express his creativity in an original manner and achieve a result which is an intellectual creation".[20] In the Football Dataco case[21], the ECJ ruled that a database is subject to copyright protection if it is the author's own intellectual creation expressing originality "in the selection or arrangement of that data".[22]

Even if users benefit from a series of exceptions limiting the scope of the owner's exclusive rights, the variety of the types of permitted acts in different legal systems and the lack of harmonization on a European level makes deciphering the legality of a use a difficult task. Moreover, access to data is necessary for a variety of reasons, some of them linked to public interest, to transparency as an essential aspect of the well being of a democratic society, to education, and to the development of businesses. Relying on the international patchwork of existing exceptions cannot ensure open data because their scope is very narrow and diverse according to applicable law.[23]

---

does not extend to the data or the material itself and is without prejudice to any copyright subsisting in the data or material contained in the compilation".

[12] The WIPO Copyright Treaty is a multilateral agreement expanding on aspects of protectable subject matter under the Berne Convention, which are of particular interest in the digital environment.

[13] "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 shall be protected as such. Such protection, which shall not extend to the data or material itself, shall be without prejudice to any copyright subsisting in the data or material itself."

[14] The Agreement on Trade Related Aspects of Intellectual Property Rights of 1994 (TRIPS) is an international agreement between members of the World Trade Organization (WTO), setting standards of intellectual property protection, dispute resolution, and enforcement in international trade between the signatory parties.

[15] Hereinafter the Database directive.

[16] Article 1(2) of the directive 96/9/EC of the European Parliament and of the Council of 11 March 1996 on the legal protection of databases.

[17] According to article 4(1) of the directive 96/9/EC of the European Parliament and of the Council of 11 March 1996 on the legal protection of databases, "the author of a database shall be the natural person or group of natural persons who created the base or, where the legislation of the Member States so permits, the legal person designated as the rightholder by that legislation".

[18] Hereinafter ECJ.

[19] ECJ Case C-5/08, *Infopaq International A/S v Danske Dagblades Forening*, [2009] ECDR 16.

[20] Id. para. 45.

[21] ECJ Case C-604/10, *Football Dataco Ltd et al v Yahoo! UK Ltd*, [2012] GRUR 2012, 386.

[22] Id. para. 42.

[23] Ensuring access to data for specific and limited purposes could be achieved through the existing exceptions in some countries. For example, data mining is a legal act in countries such as the United Kingdom for non-commercial uses (article 29A of the Copyright, Designs and Patents Act 1988) only and the United States as it is considered fair use.

Accepted manuscript
The final publication is available at B. van Loenen et al. (eds.), Open Data Exposed,
Information Technology and Law Series 30, https://doi.org/10.1007/978-94-6265-261-3_6

6.2.2. The Sui Generis Database Right

The Database directive introduced a new intellectual property right[24] in article 7[25], which is applied cumulatively to that of copyright in order to protect the economic investment[26] made for the creation of the database even if its contents are deemed to be unoriginal. It is not a copyright, but a sui generis right[27] since it does not fit any of the existing intellectual property right categories.

The sui generis right gives exclusive rights to the creator of the protected database for 15 years from its completion. As with the definition of the notion "database"[28] in the directive, the term "investment" is construed in an open-ended manner as well. However, the qualities of that investment remained unclear[29], other than the clarification that it consists of "the deployment of financial resources and/or the expending of time, effort and energy"[30]. The ECJ[31] has established a high standard of protection by maintaining that the investment dedicated to creating the data included in a database does not justify a sui generis right.[32]

The addition of an extra layer of protection to databases gives the database maker the exclusive right to "prevent extraction and/or re-utilization of the whole or of a substantial part, evaluated qualitatively and/or quantitatively, of the contents of that database."[33] The terms of "extraction" and "re-utilization" are to be interpreted widely.[34] Drawing on the rights of reproduction and communication to the public in copyright law, infringement is not limited

[24] Although not expressly qualified as an intellectual property right, the database right in question is considered as such because of its attributes. For example, article 7(3) clarifies that the right "may be transferred, assigned or granted under contractual licence."
See Derclaye 2007, p 3-4; Derclaye 2014, p. 320.

[25] "A right for the maker of a database which shows that there has been qualitatively and/or quantitatively a substantial investment in either the obtaining, verification or presentation of the contents to prevent extraction and/or re-utilization of the whole or of a substantial part, evaluated qualitatively and/or quantitatively, of the contents of that database."

[26] According to Goldstein and Hugenholtz, "judicial decisions in the Netherlands and the United States that copyright does not protect databases that result merely from economic investment or intellectual effort helped to spur the introduction of sui generis protection." Goldstein and Hugenholtz 2013, p. 240.

[27] "Of its own kind".

[28] Hugenholtz 2016.

[29] Derclaye 2005b.

[30] Recital 40, Database directive.

[31] In British Horseracing Board v William Hill[31], the ECJ ruled that "the expression 'investment in (…) the obtaining (…) of the contents' of a database must (…) be understood to refer to the resources used to seek out existing independent materials and collect them in the database, and not to the resources used for the creation as such of independent materials.": ECJ Case C-203/02, *British Horseracing Board Ltd v William Hill Organization Ltd*,[2004] I-10415, para 31. See also, ECJ case C– 338/02, *Fixtures Marketing Ltd v Svenska Spel AB*, [2005] ECDR 4.

[32] According to an author, "this interpretation is very important because a lot of so called spin-off databases, … This includes, for example, event schedules, television or radio programs, transport timetables, telephone subscriber data, stock prices, scientific data resulting from research or experimentation and sports results. If the substantial investment in the collection, verification or presentation of the materials is inseparable from the substantial investment in their creation, the right will not subsist".: Derclaye 2007, p. 7; Derclaye 2014, p. 320.

[33] Article 7(1) of the Database Directive

[34] Derclaye 2014a, p. 326.

to the creation of a competing database. All similar acts that result in the extraction or re-use of a "substantial part of the database"[35] are considered to be infringing.

The Database Directive does not clearly indicate the exclusion of public databases that fall under the PSI Directive from qualifying for the sui generis protection. In principle, since public sector databases are not excluded, branches of state power can benefit from the sui generis right protection when they fulfill the conditions[36]. Absent an ECJ decision, however, courts from some Member States have ruled against the possibility of public bodies asserting sui generis database rights. Namely, courts in Italy and Germany have held that even if public sector databases qualify for the protection, they should be exempt from it.[37] The highest administrative court in Amsterdam has held that the City of Amsterdam cannot hold sui generis rights on a database even if it has made a substantial investment towards its creation because the has not borne the risk for the investment in question.[38] Thus, it cannot impose limitations or charges in the reuse of that database. Finally, French law has been amended[39] to clarify that public bodies cannot invoke a sui generis right in order to refuse the reuse of their data.

6.3. Transnational Open Data Licensing Models

In spite of the current network of rights interlaced in the use of data and databases, a normative approach towards prioritization of open data was complemented if not preceded by transnational licenses. According to an author, when it comes to open data, "we need to be clear on what forms of re-use we expect or want to support."[40] The most prominent example in open licensing comes from Creative Commons, an organization that has dominated the field of open content[41] and has participated in the shaping of the open data licensing field worldwide. Another example of a transnational licensing of open data comes from the Open Data Commons project.

6.3.1. Creative Commons

Creative Commons is a non-profit organization created in 2001 in the United States as a reaction to the reservation-centered copyright regulation. According to Professor Lessig, copyright is a "permission culture" that is, "a culture in which creators can create only with the permission of the powerful, or of creators from the past"[42]. Creative Commons provides "free, easy-to-use copyright licenses to make a simple and standardized way to give the public permission to share and use your creative work–on conditions of your choice."[43] Inspired by the free software licenses[44], Creative Commons created a set of standardized licensing tools

[35] The content of term 'substantial' in order to qualify the infringement lacks clarity from both the Directive and the ECJ. See Derclaye 2014a, pp. 328-329; Masson 2006.
[36] Derclaye 2008; Sappa 2011.
[37] Derclaye 2008 ; Derclaye 2014a, p. 321 ; Sappa 2011.
[38] Ubaldi 2013.
[39] See article L321-3 of the code des relations entre le public et l'administration.
[40] Dodds 2010, p. 13.
[41] According to a report published in 2016 by Creative Commons, more than one billion works are currently licensed under a Creative Commons license. : State of the Commons, 2016, available online: https://stateof.creativecommons.org/ Accessed May 2018.
[42] Lessig 2004, xiv.
[43] See "What we do: What is Creative Commons?", available online: https://creativecommons.org/about/ Accessed May 2018.
[44] The innovation of the free software licenses does not only reside in the establishment of standardized

for non-software works[45]. The tools vary from the "some rights reserved" model to dedications to the public domain. [46]

6.3.2 The Creative Commons Licenses

The variety of the Creative Commons licenses is justified by the conviction that diversity in culture requires diverse tools: "Our view is that the necessary freedoms in different domains of creativity are not necessarily the same"[47]. The system of generating licenses is based on the combination of the four founding elements: Attribution (BY), No derivatives (ND), No commercial uses (NC), and Share Alike (SA), resulting in six different licenses.

6.3.3. License Core Elements And Formats

The **Attribution (BY)** element obliges the licensee to indicate the author(s) of the licensed work on each use and redistribution. Although it was introduced as an optional element in the licenses, approximately 97-98% of the users chose Attribution[48]. Creative Commons decided to make is an obligatory clause for all licenses in order to correspond both to a legal imperative and to a social one. The legal obligation is reflected in the moral rights regime of most countries, which recognizes the right of paternity of the work[49]. The social imperative corresponds to the use of attribution as a means of social recognition towards the author[50].

The **No Derivatives (ND)** element withholds the licensor's permission to create or distribute derivative material. When a license contains such an element, the work may be shared only verbatim or with format changes that do not result in an adaptation of the shared work.

The **Non-commercial uses (NC)** element of the license restricts the authorized uses of the shared work to the non-commercial sphere, defined[51] as one that is not "primarily intended for

licenses that provide the four essential freedoms to "run, copy, distribute, study, change and improve the software." Its success resulted from the creation of a veritable free movement, which consists of a community of people who share the same ideas and whose objective is to ensure access to free software and to make it evolve. Stallman 1999; Williams 2002, Shemtov and Walden 2013.

[45] According to Leonard Dobusch and Sigrid Quack, "the free/open source software movement did not only highlight the demand for nonsoftware licenses; it also functioned as a 'breeding ground' for the foundation of Creative Commons.": Dobusch and Quack 2008, p. 17.

[46] Creative Commons is presented as the mediating solution among polarized opinions expressed in the digital copyright debate because of the variety of licenses proposed to the public.: Rimmer 2007; According to Dulong de Rosnay, "unlike tailored copyright licenses written by lawyers for specific and unique needs comparable to 'haute couture', Creative Commons provides six 'prêt-à-porter' or 'ready-to- wear' texts aiming at answering most needs while minimizing the number of available 'sizes' or 'colors'.": Dulong de Rosnay 2010, pp. 9-10.

[47] Lessig 2005.

[48] "Our web stats indicate that 97-98% of you choose Attribution, so we decided to drop Attribution as a choice from our license menu — it's now standard.": Brown 2004.

[49] Giannopoulou 2014.

[50] Such recognition "as an end itself and/or as a means to obtaining financial rewards, is the common motivator among creators who use CC licenses to share their works". Katz 2006 ; According to Dulong de Rosnay, "Beyond fame and pride, it is a common feeling among creators to share their creation only in exchange of public recognition, and perhaps more visibility on their other activities": Dulong de Rosnay 2010.

[51] For the strategy-making process of defining the non-commercial uses element, see Dobusch and Kapeller 2017.

or directed toward commercial advantage or monetary compensation."[52] This restriction has been subject to criticism for posing an unnecessary restriction to free works[53] but also for the difficulties in finding a clear definition[54]. Besides these critiques, the potential of the NC element in positively affecting current business models towards more open standards has been recognized[55].

Finally, the **Share Alike (SA)** clause is the core element of the copyleft movement. By using this element, the rightsholders affirm their decision to allow modifications to the share work on the condition that the derivatives are released under an equivalent license, thus ensuring that the work will maintain the same degree of freedom throughout its evolution[56].

The combination of the aforementioned elements results in the following six licenses[57]:

Attribution (BY)

Attribution No Derivatives (BY- ND)

Attribution Non- Commercial No Derivatives (BY- NC- ND)

Attribution Non- Commercial (BY- NC)

Attribution Non- Commercial Share Alike (BY- NC- SA)

Attribution Share Alike (BY- SA)

[52] See Sect. 1(h) of the Creative Commons 4.0 "Attribution-Non commercial- No derivatives" and "Attribution- Non commercial- Share Alike" licenses.
[53] "To be free means to be open to commercial appropriation, since freedom is defined as the nonrestrictive circulation of information rather than as freedom from exploitation.": Nimus 2006; "Prohibiting commercial use except by special permission, on the other hand, puts you on the fringes *of* the free content movement, where the beer is free, but the philosophy is shallow.": Möller 2005; Mako-Hill 2005; Chen 2009; Grassmuck 2011.
[54] The interpretation of the non-commercial restriction varies but is subject to standardization from the relevant case law in different countries as well as from the norms created from users and authors. Giannopoulou 2016.
[55] Dulong de Rosnay 2010.
[56] According to the license, "if you remix, transform, or build upon the material, you must distribute your contributions under the same license as the original." See Sect. 3(b) of the Creative Commons 4.0 "Attribution-Non commercial- Share Alike" and "Attribution- Share Alike" licenses.
[57] Richard and Euan 2005.

The licenses exist in three different formats or 'layers'. The license is first delivered as a summary of its core elements, called the commons deed or human readable license. Then, the second layer is the legally binding license called the legal deed, and third is the machine-readable license[58], which describes the permissions and restrictions of the license in a form of digital-rights expression making it easier to identify and manage the shared work.

6.3.4. Evolution of the license versions

Version 1.0 of the Creative Commons licenses was published in 2002. Version 4.0 is the latest one, published in 2013. The modified incremental versions address issues identified during the use of the licenses and accommodate relevant legislations. A detailed analysis of the evolution of the licenses is outside of the scope of this chapter[59]. However, we will trace parts of this evolution that are relevant to the use of the licenses to open data.

The first significant evolution was the policy shift from "porting" licenses in different jurisdictions towards the establishment of one set of "jurisdiction-agnostic" licenses. Although every version of the licenses gradually became less US-centric, the expansion of the licenses to various countries involved not only their translation but also a process ensuring that the local licenses respected national laws. This process, called "porting"[60], created a "multilingual model of the licensing suite" that was "legally enforceable in jurisdiction around the world"[61]. However, version 4.0 adopted an international set of licenses that "operate globally, ensuring they are robust, enforceable and easily adopted worldwide"[62].

The policy shift was also possible because of the linguistic evolution of the terms of the licenses. For example, a modification that permitted the application of the Creative Commons licenses to data was the inclusion of data and databases in the subject matter of the license. The first mention of databases is found at ported licenses of version 2.0 in the Netherlands, Germany, Belgium, and France. However, it was not until version 4.0 that the international licenses explicitly included databases and their relevant rights. The shift in language had as a consequence the modification of the wording from 'author' and 'work' to 'creator' and 'material' in order to better reflect the requirements of databases.

The initial policy that advised against using the licenses for databases changed as the need for standardized licensing for open data became more evident. Addressing the sui generis right was the biggest challenge. In fact, excluding this right would make the licenses impossible to be applied to certain types of works such as Internet collections[63]. At the same time, incorporating the sui generis right to local versions of the licenses would "become especially problematic as they pose the danger that, through the use of a Creative Commons license, protection of the sui generis database right can be 'imported' to a jurisdiction without any sui generis database right protection".[64] Even if the unported version 3.0 of the licenses did not address the sui generis right, the ported versions in countries where that right existed chose to implement it. The implementation consisted of including the sui generis right in the rights

[58] This form of the license distinguished Creative Commons from other open content licenses because it was an effective expression of rights into code. See Dulong de Rosnay 2016
[59] Giannopoulou 2016; Dulong de Rosnay 2010.
[60] Maracke 2010.
[61] Ibidem, p. 70.
[62] Creative Commons goals and objectives.
https://wiki.creativecommons.org/wiki/4.0#Goals_and_objectives. Accessed May 2018.
[63] Maracke 2010, p. 79.
[64] Ibidem

covered by the license (when applicable according to applicable law) but at the same time excluding the application of the license elements when only the sui generis right was applicable. Finally, when the licensor was also the sui generis rightsholder, the right was considered waived. This method was justified in that the local teams did not want to see the database right be exported to other jurisdictions through the share alike provisions. Additionally, it was "demonstrated that applying license elements (BY, NC, ND, SA) to scientific databases is not recommended for science because the flow of information should be unrestricted and also because it is difficult even for specialized lawyers to distinguish what part is a database and assess what is a commercial use."[65]

Version 4.0 of the Creative Commons licenses fully integrates the sui generis database right[66]. Section 2 of the licenses includes the sui generis database right in the licensed rights. Section 4 describes the obligations of the licensee when using and extracting all or a substantial portion of a database in which the licensor holds a sui generis database right. All license conditions apply to the rights defined in Section 1. Thus, the licensees are required to respect the license elements even when only sui generis rights apply in the database.

Lastly, the prerogatives of applying the share alike restrictions have been subject to variations throughout the evolution of the versioning process of the Creative Commons licenses. Namely, while the share alike clause of version 1.0 of the licenses demanded that any adapted material be shared with the exact same license, consequent versions changed the wording in order to include future versions of the same licenses or other licenses with the same level of permissions and restrictions[67]. Furthermore, version 3.0 of the licenses changed the share alike clause in order to include the equivalent licenses in different jurisdictions, and a "Creative Commons Compatible License". Version 4.0 has simplified the wording but has kept the same terms in respecting the share alike clause when publishing adapted material.[68]

6.3.5. The Creative Commons waiver of rights (CC0)

Besides the main core of the six licenses, Creative Commons has also developed a tool destined to waive all existing rights in the work shared, which is called CC0. It is a "no rights reserved" tool, because the rightsholder waives all rights related to the work. CC0 is a "universal dedication that may be used by anyone wishing to permanently surrender the copyright and database rights they may have in a work, thereby placing it as nearly as possible into the worldwide public domain".[69] According to one of the accepted definitions, the public domain is comprised by "intellectual elements that are not protected by copyright or whose protection has lapsed, due to the expiration of the duration for protection"[70]. The dedication of databases to the public domain through the CC0 permits a free reuse of the database. The extent of the waiver can be limited according to applicable law because moral

[65] Dulong de Rosnay 2010, p. 95.
[66] Giannopoulou 2016; Artusio and Morando 2014.
[67] Melanie Dulong de Rosnay points out for example how "works under (Non Attribution) Share Alike (version 1.0) licenses may only breed derivatives under similar (Non-Attribution) Share Alike licenses" because of that difference in the share alike clause between version 1.0 and later versions of the licenses. : Dulong de Rosnay 2010, p. 66.
[68] These changes in the share alike conditions combined with the evolution of the subject matter and of the licensed rights touches upon the issue of compatibility between different licenses, which will be discussed section 4.1 of this chapter.
[69] Vollmer and Peters 2011; The regulation of the public domain varies according to jurisdiction and lacks clarity. Dusollier 2010; Guadamuz 2014.
[70] Dusollier 2010.

rights are inalienable in a lot of countries and because there is no generalized process of dedication to the public domain found in positive law.

The dedication of databases to the public domain with the use of CC0 was considered a solution that would circumvent the normative diversity in applicable laws. Indeed, the distribution of a European database using CC0 "would allow the product to circulate freely without any doubt about the systems of protection to be applied"[71]. Waiving all rights would bring the legal certainty needed for the open data.[72] However, the use of CC0 did not necessarily ensure respect of the principles of open data. For example, the free use of data did not accommodate the conditions of attribution and provenance in the use of databases even if the infrastructure of CC0 asks for information on the work in question. Applying CC0 on a dataset does not ensure a perennial state of openness because of the possibility of appropriation of the data. The application of a share alike clause, which would enable "an inclusive approach"[73] is only possible through the use of the licenses and not through the CC0 waiver. What's more, attaching restrictions to open data may derive from public policies more broadly. In these cases, the licenses seem more suitable than waiver.

6.3.6. Open data Commons

At the time of publication of version 3.0 of the Creative Commons licenses, the uncertainty surrounding the treatment of the rights in databases incited the creation of database-specific licenses. The Open Data Commons project was founded in 2007 in order to create specialized licensing tools[74], namely the Public Domain Dedication & License (PDDL) and the Open Database licenses.

6.3.7. The Public Domain Dedication and License (PDDL)

The first database-specific tool, which prompted the creation of the Open Data Commons project, was the Public Domain Dedication and License[75]. It consists of a waiver of all rights and claims in the database, placing the database in the public domain. The waiver serves the purpose of "maximizing the dissemination and the overall utility of data and databases—by reducing legal uncertainty and transaction costs to the minimum".[76] Applying the principles of CC0, the Public Domain Dedication and License is similar to both the Creative Commons structure and the terminology of database rights regulation.

The Open Data Commons project converged with the initial principles promoted by Creative Commons, that open data sharing "should come with the fewest possible restrictions and obligations (…) The PDDL, CC0, and other public domain dedications or copyright waivers, provide a far simpler, more consistent, and more benign approach that closely mirrors a long history and tradition of scientific, educational, and cultural sharing practices. For these communities, these norms converge on the public domain".[77] Although the language of CC0

[71] Aliprandi 2012, p. 11.
[72] "This is important since when a dataset is shared on the web, an American reuser, for instance, might wonder if he or she has to comply with an European right that he or she is not familiar with". Leucci 2014.
[73] Artusio and Morando 2014.
[74] The precursor was the Talis Community Licence created in April 2006 by the Talis firm, which specializes on developing Semantic Web solutions. See Miller, Styles and Heath 2008.
[75] Hereinafter PDDL
[76] De Filippi and Maurel 2015, p 20.
[77] Nguyen 2009.

remains broader, as it is not specific to one type of content, the PDDL applies the waiver for all rights that can be found in databases, with a fallback license for rights that are inalienable in some jurisdictions, such as moral rights.

The legal structure of the "two prong approach"[78] to open data consists of a waiver of sui generis rights and copyright for jurisdictions that allow for it[79] and a licensing of the rights that are not subject to a waiver in some jurisdictions[80]. However, the waiver and license do not cover for any patent or trademarks that may subsist in the database in question[81] because "it was important that the provider of the data under the PDDL be placed in the same position as anyone else using the data"[82]. The innovation of the PDDL is the voluntary addition of community norms in the use of open data. The soft law principles do not constitute a contract, but through the principles described they try to achieve a social responsibility towards the use of open data. Even though people are free to create their own community norms, the example provided by Open Data Commons incites the introduction of a reciprocal principle and the creation of attribution and citation norms in specific datasets. In maintaining a non-legally binding document, the project's objective was to avoid overburdening users and licensors and also to ensure malleability in data use.

6.3.8. The Open Data Commons licenses (ODbL)

The database-specific licenses created by Open Data Commons used the Attribution and Share alike elements thus creating the Attribution License (ODC-By) and the Open Database License (ODC-ODbL). The elements used in the licenses are already standardized by Creative Commons. The Open Data Commons licenses apply the same principles in order to avoid user confusion and to improve effectiveness of the end result licensing tools. At the same time and because the licenses are database-oriented, the elements are modified to cater for the requirements ensuring the legal dissemination of open data. The Attribution and Share alike elements are more detailed so as to include issues involving the creation of derivative databases, the extraction and reuse of data publicly, as well as the distinction between licensing the database and its content. Namely, the preamble of the licenses clarifies that "because databases can have a wide variety of types of contents, this document only governs the rights over the database, and not the contents of the database individually. You should use the Open Data Commons together with another license for the contents, if the contents have a single set of rights that governs all of them". As with the PDDL, only the copyright and sui generis database rights are covered by the license grant and are subject to the license conditions. However, the share alike license condition differs from the voluntary share alike soft law reciprocal principle that can be found with the application of the PDDL. While users are contractually bound to respect the former, they are free to ignore the latter. The adaptation conditions for the creation and use of derivative databases will depend on the nature of these elements[83].

6.4. National open data licenses

Some countries opted for the creation of a distinct set of licenses to accompany the dissemination of the public sector information. For example, the UK and France have

[78] Hatcher 2008.
[79] See Sections 3.1 and 3.2 of the license.
[80] See Sections 3.3 of the license.
[81] See Section 4 of the license.
[82] Ibidem.
[83] See Section 4.1 on compatibility issues between licenses.

produced national licenses adapted to the PSI Directive and covering rights related to the sharing and reuse of public data and databases.

6.4.1. The Open Government Licence (UK)

The UK Government Licensing Framework for Public Sector Information[84] was created to license "the use and re-use of public sector information both in central government and the wider public sector."[85] The Open Government License[86] replaced its precursor, namely the Click-Use license on 2010. The license removes barriers to the reuse of public sector information, requiring only "including or linking to any attribution statement specified by the Information Provider(s) and, where possible, provide a link to th(e) licence."[87]

The licensed rights include intellectual property rights and sui generis database rights. The copyright status of public sector information in the UK also includes the Crown copyright[88], which is a special copyright that vests in governmental works. The license ensures that the rightsholder of the Crown copyright can use the OGL to create open data.[89] Since public sector bodies can only license information for which they own the relevant rights covered by the license, it was especially important to ensure that information protected by Crown copyright can be made open under the terms of the license.

However, there are circumstances where public sector information can only be reused in a non-commercial environment. Recognizing the need to limit the reuse of that information without compromising access, the Non-Commercial Government License was created in the context of the UKGLF. The license permits the reuse of the subject matter according to the licensed rights but prohibits the exercise of these rights "in any manner that is primarily intended for or directed toward commercial advantage or private monetary compensation."[90] The restriction, similar to the non-commercial (NC) element of Creative Commons, creates an environment of limited reuse that gives the governing body the power to negotiate complementary agreements for commercial uses of the public sector information in question.

Finally, and complying with the PSI Directive, the Government can impose fees for the re-use of public sector information subject to certain conditions. Namely, the fee has to be limited to "the marginal costs incurred for their reproduction, provision and dissemination," and the charges "shall not exceed the cost of collection, production, reproduction and dissemination, together with a reasonable return on investment."[91] Towards that direction, the Charged License created by the UK government charges users for the use of public sector data in the exceptional cases for which these charges apply, in order to recompense for the making

[84] Hereinafter UKGLF.
[85] National archives, 2016.
[86] Hereinafter OGL.
[87] Open Government Licence, Version 3.0.
[88] According to Section 163 of the Copyright, Designs and Patents Act 1988, Crown copyright is attributed to works made by officers or servants of the Crown in the course of their duties.
[89] "The Controller of Her Majesty's Stationery Office (HMSO) has developed this licence as a tool to enable Information Providers in the public sector to license the use and re-use of their Information under a common open licence." Open Government License, Version 3.0.
[90] Non –Commercial Open Government Licence.
[91] Directive 2013/37/EU of the European Parliament and of the Council of 26 June 2013, 2013, O.J. (L 175) 1, 1 Amending Directive 2003/98/EC on the Re-Use of Public Sector Information, 2003 O.J. (L 345), art. 6., at 2

available of public sector information[92]. The charges are therefore permitted as long as they are in the public interest and as long as they are calculated based on objective, transparent and verifiable criteria.[93]

6.4.2. The Licence ouverte/Etalab (France)

France developed an open data policy in 2011 with the creation of Etalab and its respective license. The framework for open data permits the re-use of data for both commercial and non-commercial purposes provided that the source is properly attributed. Also, the licensor guarantees that no third-party intellectual property rights exist in the licensed data. Finally, the license also provides a non-endorsement clause and a license compatibility list[94].

Following the adoption of PSI directive, French law was modified[95] to establish that public sector data be open by default with limited exceptions. According to the new legal provisions, the licenses available for use by the public sector are listed in an official decree reviewed for renewal every five years[96]. This provision, along with other amendments, led to the creation of version 2.0 of the Etalab license. Published on April 2017, version 2.0 maintains the same open data principles regarding the reuse rights and obligations. It also adds a personal data provision in order to comply with the amended law that excludes the publication of sensitive data even with the consent of the concerned person. Henceforth, information containing personal data can be made available only through anonymization[97]. Finally, the license does not provide any warranties and alters the "compatible licenses" list to make the license compatible with all versions of the "Creative Commons Attribution" license.

The Etalab license version 2.0 is now the default license to use for the re-use of public sector information in France.[98] However, and according to the published decree of approved licenses[99], administrations can also use the Open Database License. The list in question does not include the Creative Commons licenses, even though their compatibility with the Etalab license is stated within the license terms. The absence of Creative Commons from the decree is noteworthy because they are the most widespread licenses on an international level in open scientific and public sector data.[100]

---

[92] The license is created in accordance with the Re-use of Public Sector Information (PSI) Regulations 2015: The Re-Use of Public Sector Information Regulations 2015, SI 2015/1415, (UK), available at legislation.gov.uk/uksi/2015/1415/contents/made. Accessed on 1 July 2017.

[93] The charges in question are not to be construed as means of opening data in order to finance the public sector. : Corbin 2010, Lee 2017.

[94] See Sect. 4.1

[95] Loi n° 2015-1779 du 28 décembre 2015 relative à la gratuité et aux modalités de la réutilisation des informations du secteur public, JORF n°0301 du 29 décembre 2015, p 24319. ; Loi n° 2016-1321 du 7 octobre 2016 pour une République numérique, JORF n°0235 du 8 octobre 2016.

[96] See article L.323-2 al. 4 of the code des relations entre le public et l'administration, as modified according to article 11, Loi n° 2016-1321 du 7 octobre 2016 pour une République numérique.

[97] In order to signify the change, the added provision clarifies that the license conforms to the n° 78-17 of 6 January 1978 regarding the protection of personal data.

[98] The decree (Décret n° 2016-1922 du 28 décembre 2016 relatif à la publication en ligne des documents administratifs) published in accordance to article 6 of the law n° 2016-1321 of 7 October 2016 specifies that all French administrations with more than 50 agents have to make their data open.

[99] See article D323-2-1 created by decree n°2017-638 du 27 avril 2017 - art. 1.

[100] The justification for not including the Creative Commons in the decree of open data licenses could be found on the fact that version 4.0 has not yet been translated in French thus risking user confusion if applied.

6.5. The interplay of diverse open data models

The diversity in policies, licenses, and intellectual property systems has nurtured the existence of a multifaceted open data environment. The applicable open data model is not uniform on a transnational level, creating a fragmented licensing matrix on top of an already diverse regulatory framework. The issues stemming from this diversity are not only theoretical but can constitute an obstacle in the development and maximization of the reuse of open data.

6.5.1. Compatibility issues among licenses

License compatibility is the most prominent issue stemming from the proliferation of open data licenses. Compatibility only exists when two works subject to different licenses can be legally fused together. Two licenses are compatible when all the rights granted by the absorbing license are included in all the rights conferred by the compatible license and if all the obligations imposed by the compatible license are included in the absorbing license.[101] The compatibility can be internal, meaning between the same licensing model, or external, meaning between different licensing models[102]. Compatibility can be one-way or two-way. According to Creative Commons, "one-way compatibility means that you may adapt work under one license (X) and apply a second license (Y) to your own contributions, but you may not adapt work under the Y license and apply the X license to your contributions."[103]

6.5.2. Compatibility issues in the same open data licensing model

The sources of incompatibility of licenses in the same model stem either from the difference in licensing terms or from the diversity arising from the evolution of the versions of the licenses. The Creative Commons licenses have created multiple sources of incompatibility varying from the difference in license elements to the differences in incremental versions and ported ones[104].

Incompatible restrictions are common within license families as well as between licenses of different models. For example, when the license restricts commercial uses, the derivative works created can only be published with the same license because additional permissions are required in order to obtain the right to permit commercial uses of the published content. Consequently, the license CC BY NC can only be combined with the same license or with an equally restrictive license such as CC BY NC SA. The only license that can be easily combined with works licensed under more restrictive terms is the CC BY license because it imposes no obligations as to the licensing conditions of the derivative content.

The Open Government license is incompatible with the two more restrictive licenses from the same licensing model. Similarly, the Open Data Commons permissive licenses are only one-way incompatible because of the share alike clause. For example, the database licensed with ODC By can freely be reused for the creation of a derivative database, which can be licensed with the ODbL, because nothing in the first license dictates the licensing terms of derivative databases. At the same time, the derivative created from the reuse of a database licensed under the ODbL can only be licensed under the same license or "a later version of this

[101] Benjamin Jean uses examples from set theory to demonstrate the complexities of license compatibility. Jean 2006
[102] Dulong de Rosnay 2010, p. 60 ; Giannopoulou 2016, pp. 118-126.
[103] See "What does one-way or two-way compatibility mean?" in ShareAlike compatibility. Available on: https://wiki.creativecommons.org/wiki/ShareAlike_compatibility. Accessed May 2018.
[104] Dulong de Rosnay 2010.

License similar in spirit to this License; or iii. a compatible license"[105]. The share alike obligation dictates the compatibility rules because it restricts the licensing conditions of the derivative databases. Etalab version 2.0 is the only license that declares backwards compatibility with content licensed with the prior version of the same license[106]. Finally, the public domain tools ensure only one-way compatibility with the licenses because the application of a waiver to a derivative database is only possible when its maker is the rights holder of all the rights involved.

The evolution of the terms of the licenses for databases creates incompatibility risks between different versions. For example, the ported versions 2.0 in France, the Netherlands, Belgium, and Germany are only compatible between each other because of the inclusion of the sui generis rights to the license. The ported versions 3.0 create an additional fragmented licensing field because of the waiver of the sui generis rights and the non-application of the license elements to the sui generis rights[107]. Creative Commons declares the licenses compatible with the equivalent version 4.0 for the purposes of the share alike element. However, the different prerogatives in database licensing would provoke a de facto incompatibility since in version 4.0 the license elements apply to sui generis rights too.

6.5.3. Compatibility issues across different open data licensing models

Creating interoperable licenses facilitates the reuse of information and strengthens the Open Data movement. Lately, "progress has been made to match some characteristic prescriptions featured in most of the open data licenses (particularly those developed by public authorities) by prohibiting explicitly to sublicense the material, as well as prescribing to mark or indicate modifications to the original data and not to assert any official status regarding the licensee's use of such data. In addition to that, the adoption of the Database directive terminology contributes to improve lexical accuracy and clarity".[108] The standardization tendency has harmonized a lot of prerogatives across the open data licensing field, but incompatibility risks remain.[109]

Versions prior to version 4.0 of the Creative Commons licenses are a priori not compatible with other open data licenses because of the diverse rules that apply to sui generis rights. The incompatibility across different open data licenses due to the license elements is relatively straightforward because, with the exception of the Creative Commons licenses, the rest of the open data models examined have developed a limited diversity in license elements. However, "the Share Alike compatibility is merely a political statement which must be validated by the facts"[110]. The process of validating compatible licenses is not only a parallel examination of the license terms in order to verify the similarities, but it is also a political consensus between organizations that the licenses produce the same effects. Creative Commons has made a list of compatible licenses, but there are no open data licenses on that list.

[105] See Sect. 4.4(a) of the ODbL.
[106] See Sect. "Compatibilité de la présente licence" : "La présente licence a été conçue pour être compatible avec toute licence libre qui exige au moins la mention de paternité et notamment avec la version antérieure de la présente licence". : Licence ouverte, version 2.0.
[107] The lack of clarity in compatibility issues concerning the licensing of the sui generis rights between different versions has been pointed out by Dulong de Rosnay. : Dulong de Rosnay 2010, pp. 94-96.
[108] Ibidem, p 279.
[109] On this issue regarding geodata: Loenen van, Janssen and Welle Donker 2012.
[110] Dulong de Rosnay 2010, p. 81.

The international open data models, such as the Creative Commons version 4.0 and the Open Data Commons licenses are considered the point of reference for the local open data licenses. In the spirit of standardization of the licensing terms and in order to maximize the potential of reuse of the open data, the local licenses intentionally declare compatibility with the larger open data models. For example, according to the Open Government license (UK), "these terms are compatible with the Creative Commons Attribution License 4.0 and the Open Data Commons Attribution License, both of which license copyright and database rights. This means that when the Information is adapted and licensed under either of those licences, you automatically satisfy the conditions of the OGL when you comply with the other licence." Similarly, the Etalab license declares compatibility with the Open Government Licence, the Creative Commons Attribution and the Open Data Commons Attribution. The compatibility declarations constitute a form of political consensus between the bodies that develop the licenses. They are also put in place in order to reduce user confusion and transaction costs. Thus the reuse and dissemination of open data is facilitated by the creation of legally interoperable databases.

6.5.4. Fragmentation and proliferation of open data policies

The economic and social value of open data has led to policies that prioritize transparency, democracy, and innovation. The multiplication of policies and the expansion of licensing models have nevertheless given rise to a quasi-universal set of principles for open data. However, even though there is a trend towards the establishment of open data, the policies implemented and the tools used to implement them are not as homogenous as they initially appear to be. At the same time, the interaction and engagement with open data transcends national border. Consequently, the diversity of policies poses an obstacle to the wide dissemination of data and hinders their reuse due to the fragmentation of the licensing options. The European Data Portal[111] demonstrates the extent of the license fragmentation by listing the different licenses applied across the EU. The license variation does not only depend on the body that created them. The license differences extend to various incremental versions, to the language used, and to type of restrictions applied varying from prohibition of commercial uses to that of derivative works.

This diversity is found in the restrictions applied on top of the freedom to access and reuse data, as entities often disagree on the scope of various limitations. A maximized dissemination of open data, however, can only be achieved when "minimal constraints" apply. Therefore, every restriction to the reuse of open data needs to be properly justified.[112] For example, open data policies rarely require the payment of fees for the reuse of open data since the free of charge sharing and reuse of open data is a founding principle of the open data movement. However, as mentioned earlier, there are cases where charges apply for the use of open data.[113] These fees risk creating a walled environment that, if provided without the appropriate safeguards, poses a hurdle in the dissemination and interoperability of open data.

---

[111] The European Data Portal is developed by the European Commission in order to harvest "the metadata of public data made available across Europe". There are more than 30 available licenses listed. See https://www.europeandataportal.eu/en/content/show-license. Accessed May 2018.

[112] "[O]pen data policies . . . encourage the wide availability and re-use of public sector information for private or commercial purposes, with minimal or no legal, technical, or financial constraints." Directive 2013/37/EU of the European Parliament and of the Council of 26 June 2013, 2013 O.J. (L 175) 1, 1 Amending Directive 2003/98/EC on the Re-Use of Public Sector Information, 2003 O.J. (L 345).

[113] The UK government created an exception to the freedom to use open data with the Charged license, justified on the "costs that arise from the reuse of information". The restriction of commercial uses of open data is also susceptible to create extra charges that may go beyond the initial policy justification

The share alike restriction is also a limitation that is not essential for the reuse of open public sector data. "A share-alike provision may impede new business models and innovative commercial uses of government data, which will eventually run counter to the policy goal of promoting economic development. (…) It is the responsibility of governments, rather than the private sector, to keep government data freely available." [114] Consequently, the choice of applying the reciprocal share alike limitation to open data may pose an unnecessary burden to the freedom to reuse and to disseminate the data in question. Finally, the attribution requirement is an important element of open data, whether as part of the license restrictions or as part of a contractual limitation on top of a waiver. It constitutes a restriction justified by open data policies since it contributes to the policy justifications of transparency[115]. In this respect, attributing the source of the data used could be qualified as one of the most common restrictions imposed among many open data policies applied[116].

Derived from the different legal systems, the open data policies reflect the jurisdictions for which they were created. The collaboration between different policy actors is necessary because the application of one set of international standardized licenses can be difficult to apply in the current legal framework. Thus, recognizing possible paths towards compatible licenses and interoperability constitutes an affirmative step in that direction.

6.5.5. Lack of harmonization of underlying legal framework

As we demonstrated in the first section, there is a big diversity in underlying positive law that applies to data and databases. This diversity adds to the legal complexity of creating open data that are accessible and re-usable on a transnational level, with the use of interoperable licenses and streamlined rules.

For example, the different applications of the originality threshold for copyright protection add to the confusion in distinguishing original from non-original databases according to applicable law. Applying an open data license is only possible when the database is protected by copyright or by a sui generis right. The application of the sui generis right in some jurisdictions for both original and non-original databases poses an extra layer of complexity to the creation of open data. The combination of the diversity in standards of copyright protection with the fragmented application of the extra layer of sui generis right can lead to either an extensive application of licenses or to the application of a license that insufficiently licenses the rights related to a database. Recognizing that the sui generis right can be an obstacle to the creation of open public sector data, French law precludes the sui generis rightsholder from invoking their right in order to impede the creation of open data. Consequently, the sui generis right cannot be used by administrations or other third party rightsolders to prohibit the application of a license creating open data in France.

---

of covering the necessary costs. What's more, the transposition of the PSI Directive in France brought a nuance to the principle of free access in applied open data policies. Namely, article 5 of the loi n°2015-1779 relative à la gratuité et aux modalities de la réutilisation des informations du secteur public du 28 décembre 2015 introduces charges that can be imposed by institutions to cover the costs of collection, production and the making available of data.

[114] Lee 2017, pp. 236-237.

[115] Ibidem, pp. 235-239.

[116] See for example the European Data Portal overview of different licenses applied to open data presenting different types of licenses, most of which include the obligation of attribution.

The application of a license inherently presumes the existence of property rights. According to some commentators, the creation of open data through the application of the licenses takes "for granted an established system made up of several layers of rights to exert control over information"[117]. Even though a lot of public sector data is not protected by copyright or sui generis rights, open data licenses and waivers are applied to databases with an unclear intellectual property rights status or who are already in the public domain.[118] The difficulty in ascertaining whether a database is eligible for protection or not leads to a tendency to over-regulate by applying open data licenses[119].

6.6. Conclusion

The development of open data policies has brought a new dynamism in the information digital age. Despite the evidence of social and economical advantages from the application of these policies, legal issues remain unresolved. Retracing the evolution of open data licenses is not only a theoretical question, because it provides the foundation for the improvement of the interoperability conditions maximizing the reuse possibilities of open data. The simultaneous development and parallel evolution of different open data licenses encourages the creation of licenses that are not only compliant with the existing legal framework and policies but also standardized and interoperable. Although it may be ideal from a user perspective[120], the diversity in data eligible to be included in the open data models cannot easily be subject to a uniform international license. Such a license would have to take into account the different needs of various types of datasets ranging from scientific data to public sector data, the variety of legal obstacles, and the linguistic barriers, in order to contribute to the improvement of open data quality. This study demonstrated that the variation in open data licensing schemes founded in the diversity of the underlying legal framework could be addressed with a combination of affirmative policies towards improvement of interoperability complemented by a standardization of the existing licenses.

Based on these conclusions, further research can be pursued on the relevance of the license infrastructure in improving open data quality. For example, the development of interpretation methods such as checklists across different dataset types could contribute towards the interpretation of restrictions such as share alike and commercial use in specific contexts. Also, research on the improvement of compatibility issues can focus beyond the practical and legal interplay of different open data licenses and towards the creation of harmonizing policies that provide external support to open data models. Finally, it seems promising to engage in studies on the development of open data in the current environment of mass production of data taking into account the case of machine-generated data and the rights created over them.

[117] De Filippi and Maurel 2016, p. 14.
[118] "[I]n countries where official texts are not protected by copyright, the issue of the need for access through licenses does not even arise".: Derclaye 2014b.
[119] De Filippi and Maurel 2016, p. 17 ; Lee 2017, p. 232 ; This issue was raised in France when new legislation made open data the default rule for most public sector data. It was however decided to continue the application of licenses for educational purposes and to avoid user confusion.
[120] See, for example, Van Loenen et al. 2012

References

Aliprandi S (2012) Open licensing and databases. International Free and Open Source Software Law Review, 4(1):5–18

Artusio C, Morando F (2014) Creative Commons 4.0 Licenses: A sui generis challenge? In: Parycek P, Edelmann N (eds) Proceedings of the international conference for e-democracy and open government. Donau- Universität Krems, pp. 269–282

Brown GO (2004) Announcing (and explaining) our new 2.0 licences. https://creativecommons.org/2004/05/25/announcingandexplainingournew20licenses/. Accessed May 2018

Bygrave L (2012) The Data Difficulty in Database Protection, EIPR 35(1):25–33

Chen SL (2009) To surpass or to conform – What are public licenses for? U. Ill. J.L. Tech. & Pol'y 1:107–139

Corbin C (2010) PSI Policy Principles: European Best Practice. In: Fitzgerald BF (ed) Access to public sector information volume 1. Sydney University Press, Sydney, pp. 161–167

De Filippi P, Maurel L (2015) The paradoxes of open data and how to get rid of it? Analysing the interplay between open data and sui-generis rights on databases. Sci. & Tech. L. Rev 23(1):1–22

Derclaye E (2005) Database sui generis right: what is a substantial investment? A tentative definition. IIC: Int'l Rev. Intell. Prop. & Competition L. 36:2–30

Derclaye E (2007) Database sui generis right: the need to take the public's right to information and freedom of expression into account. In: Macmillan F (ed) New directions in copyright law volume 5. Edward Elgar, Cheltenham, pp. 3–23

Derclaye E (2008) Does the Directive on the Re-use of Public Sector Information affect the State's database sui generis right? In: Gaster J et al. (eds) Knowledge rights – Legal, societal and related technological aspects. Austrian Computer Society, Vienna, pp. 137–169

Derclaye E (2014a) The Database directive. In: Stamatoudi I, Torremans P (eds) EU Copyright Law. A Commentary. Edward Elgar, Cheltenham, pp. 298–354

Derclaye E (2014b) The role of copyright in the protection of environment and the fight against climate change: is the current copyright system adequate? WIPO J. 5(2):156–158

Dobusch L, Kapeller J (2017) Open strategy-making with crowds and communities: Comparing Wikimedia and Creative Commons. LRP. https://doi.org/10.1016/j.lrp.2017.08.005

Dobusch L, Quack S (2008) Epistemic communities and social movements. Transnational dynamics in the case of Creative Commons. Max-Planck-Institut für Gesellschaftsforschung, Discussion paper, no08/8

Dodds L (2010) Rights statements on the Web of Data, Nodalities Magazine 9:13–14

Dulong de Rosnay M (2010) Creative Commons Licenses Legal Pitfalls: Incompatibilities and Solutions, Report. Institute for Information Law, Amsterdam

Dulong de Rosnay M (2016) Les Golems du numérique. Droit d'auteur et Lex Electronica. Presses des Mines, Paris

Dusollier S (2010) Scoping study on copyright and related rights and the public domain, WIPO (CDIP).

http://www.wipo.int/edocs/mdocs/mdocs/en/cdip_4/cdip_4_3_rev_study_inf_1.pdf. Accessed May 2018

Giannopoulou A (2014) The Creative Commons licenses through moral rights provisions in French law, Int'l Rev L Comp & Tech 27(1–2):60–80

Giannopoulou A (2016) Les licences Creative Commons, PhD Thesis, University of Paris II Panthéon Assas

Golstein P, Hugenholtz B (2013) International Copyright: Principles, Law, and Practice, 3rd edn. Oxford University Press

Grassmuck V (2011) Towards a social contract: free-licensing into the knowledge commons. In: Guibault L, Angelopoulos C (eds) Open content licensing: from theory to practice. Amsterdam University Press, Amsterdam, pp. 21–50

Guadamuz A (2014) Comparative Analysis of National Approaches on Voluntary Copyright Relinquishment. WIPO Report CDIP/13/INF/10

Hatcher J (2008) Implementing Open Data: The Open Data Commons Project. Open Source Business Resource, Technology Innovation Management Review. http://timreview.ca/article/119. Accessed May 2018

Hugenholtz PB (2016) Something Completely Different: Europe's Sui Generis Database Right. In: Frankel S, Gervais D (eds) The Internet and the Emerging Importance of New Forms of Intellectual Property. Kluwer Law International, pp. 205–222

Janssen K, Hugelier S (2013) Open data as the standard for Europe? A critical analysis of the European Commission's proposal to amend the PSI Directive, EJLT 4(3)

Jean B (2006), "Option libre": compatibilité entre contrats. Master thesis, University of Montpellier I

Katz Z (2006) Pitfalls of open licensing: An analysis of Creative Commons licensing, IDEA 46(3):391–413

Lee JA (2017) Licensing Open Government Data. Hastings Bus. L.J. 13(2):207–240

Lessig L (2004) Free culture: How big media uses technology and the law to lock down culture and control creativity. Penguin Press, New York

Lessig L (2005) CC in Review: Lawrence Lessig on important freedoms. https://creativecommons.org/2005/12/07/ccinreviewlawrencelessigonimportantfreedoms/. Accessed May 2018

Leucci S (2014) Preliminary Notes on Open Data Licensing. JOAL 2(1)

Mako Hill B (2005) Towards a standard of freedom: Creative Commons and the Free Software Movement. https://mako.cc/writing/toward_a_standard_of_freedom.html. Accessed May 2018

Maracke C (2010) Creative Commons International: The international license porting project-Origins, experiences, and challenges. In: Dulong de Rosnay M et al. (eds) Intelligent multimedia. Managing creative works in a digital work. European Press Academic Publishing. Florence, pp. 67–88

Masson A (2006) Creation of database or creation of data: crucial choices in the matter of database protection. EIPR 28(5):261–267

Miller P, Styles R, Heath T (2008) Open Data Commons: A license for open data, Proceedings of the 1st Workshop about Linked Data on the Web, pp. 1–5

Möller E (2007) The Case For Free Use: Reasons Not To Use A Creative Commons-NC License. https://freedomdefined.org/Licenses/NC. Accessed May 2018

National Archives (2016) UK Government Licensing Framework For Public Sector Information, 5.0 edition. http://www.nationalarchives.gov.uk/documents/information-management/uk-governmentlicensing-framework.pdf. Accessed May 2018

Nguyen T (2009) Comments on the Open Database License Proposed by Open Data Commons, Science Commons Reading Room. http://sciencecommons.org/resources/readingroom/commentson-odbl. Accessed May 2018

Nimus A (2006) Copyright, copyleft and the creative anti-commons, Subsol, Markets and immaterial labor. http://subsol.c3.hu/subsol_2/contributors0/nimustext.html. Accessed May 2018

Richard J, Euan C (2005) Full fat, semi-skimmed or no milk today - Creative Commons licences and English folk music. Int'l Rev L Comp & Tech 19(3):259–275

Rimmer M (2007) Remix culture: the creative commons and its discontents. In: Digital copyright and the consumer revolution: hands off my iPod. Edward Elgar, Cheltenham, pp. 261–295

Sappa A (2011) Public sector databases - the contentions between sui generis protection and re-use, CTLR 17(8):217–223

Shemtov N, Walden I (2013) Free and open source software. Policy, law, and practice. Oxford University Press, Oxford

Stallman R (1999) The GNU operating system and the free software movement. In: DiBona C, Ockman S (eds) Open sources. Voices from the open source revolution. O'Reilly, Sebastobol

Ubaldi B (2013) Open Government Data: Towards Empirical Analysis of Open Government Data Initiatives. OECD Working Papers on Public Governance, no. 22, OECD Publishing, Paris, https://doi.org/10.1787/5k46bj4f03s7

Van Loenen B, Janssen K, Welle Donker FM (2012) Towards true interoperable geographic data: developing a global standard for geo-data licences. In: Janssen K, Crompvoets J (eds) Geographic Data and the Law. Defining New Challenges. Leuven University Press, Leuven, pp. 19–36

Vollmer T, Peters D (2011) Creative Commons and public sector information: flexible tools to support PSI creators and re-users, EPSI Topic Report no 23. https://www.europeandataportal.eu/sites/default/files/2011_creative_commons_and_psi.pdf. Accessed May 2018

Williams S (2002) Free as in freedom. Richard Stallman's crusade for free software. O'Reilly, Sebastopol Zuiderwijk A, Janssen M (2014) Open data policies, their implementation and impact: A framework for comparison. Gov Inf Q 31:17–29