

CHAPTER 7. LEGAL ASPECTS OF OPEN ACCESS TO PUBLICLY FUNDED RESEARCH

Internet growth, content digitisation, and expanding “big data” and data analytics capabilities have affected the ways in which publicly funded research results are accessed, disseminated and used. While these technological advances have made sharing and processing information easier, that does not change the fact that the information may be protected by IP laws. Open access efforts, which aim to make the outputs of publicly funded research more widely accessible in digital formats, therefore raise a number of IP policy questions. To explain the interplay between open access and IP laws, this chapter provides an overview of the IP regimes that protect research outputs in a sample of OECD jurisdictions. It then reviews the open access policies that are in place in some of those jurisdictions and examines two contexts in which IP questions can arise when open access principles are applied: public/private partnerships and text and data mining.

The statistical data for Israel are supplied by and under the responsibility of the relevant Israeli authorities or third party. The use of such data by the OECD is without prejudice to the status of the Golan Heights, East Jerusalem and Israeli settlements in the West Bank under the terms of international law.

It should be noted that statistical data on Israeli patents and trademarks are supplied by the patent and trademark offices of the relevant countries.

EXECUTIVE SUMMARY

This chapter gives an overview of the main legal issues and policy considerations involved in accessing, disseminating or using both publicly funded scientific publications and research data under open access (OA) conditions. The term open access is used to denote public availability of scientific output (publications and data) without payment and re-use restriction. This paper focuses on the legal aspects of OA to publicly funded research results and, in particular, it examines whether copyright law and database protection rights support, impede, enable or are neutral towards the implementation of open access principles for the dissemination of publicly funded scientific results.

After a general overview of intellectual property protection at the international level, Section 2 examines the laws of different jurisdictions (including countries from the European Union, the United States of America, Australia, Canada, Israel, Japan and South Korea) as they relate to the protection of scientific publications and research data. The brief overview of the copyright and, where applicable, the database legislation in force in the different jurisdictions shows that the scope of protection granted to research output varies significantly from one country to another. Although the implementation of OA principles is based on contractual arrangements between authors, publishers, universities, and funding organisations, the framework set by copyright and database protection regimes, where relevant, is an important factor in how those arrangements are formed.

The way copyright law defines the scope of rights and recognises limitations and exceptions on these rights serves as the backbone to the licensing agreements. If it is to support OA, whether granted through the ‘Golden Road’ (where articles published in OA journals are provided directly and for free) or the ‘Green Road’ (where articles are published through traditional channels and subsequently deposited in institutional repositories), the copyright regime should create a favourable environment for the dissemination and re-use of publicly funded scholarly publications.

Section 3 follows by highlighting the main characteristics of OA principles and policies in certain jurisdictions. Section 4 draws attention to two as yet unresolved issues, namely the problem of OA in the context of public-private partnerships and that of text and data mining (TDM). OA principles entail more than just granting access to research output free of charge. Where the funding of a research project is partly realised through external private sources, different rules of ownership may apply. What are the implications of applying OA principles to research results that are not funded entirely by public money? The uncertainty in current legal frameworks regarding the scope of protection of works and databases could create obstacles for TDM activities. A system resting solely on licensing agreements might be insufficient to allow TDM to take place in all instances where it would be socially desirable.

Although no generalisation can be made, countries that actively encourage compliance with OA principles for the publication of publicly funded research results seem to steer copyright reform in a more flexible and research-friendly direction. The United Kingdom is one good example of this: while the Research Council has adopted a ‘Golden Road’ policy, mandating researchers to publish results under a Creative Commons Attribution Licence 4.0, the legislator has also proceeded with the adoption of new exceptions on copyright, including a specific exception for text and data mining. The German research council may not have officially opted for an OA mandate on its grant recipients, but the legislator did modify the copyright act to make it easier for authors to comply with the contractual arrangements with publishers. The general framework of United States copyright law, which excludes Federal Government

works from protection and admits a fair use defence, is also very conducive to the application of OA principles.

The complexity of the status of research data in particular as protected by intellectual property in Europe and other jurisdictions arguably has the potential to adversely affect re-use opportunities, given the difficulty – both for research institutions making the database available and for prospective re-users – in determining each time whether a certain database is covered by a *sui generis* right and in which measure re-utilisation and extraction can take place freely. Whether the use of compilations or databases for purposes of research and private study in general, and text and data mining in particular, is covered by any relevant exception on copyright or the database right is not always clear. The use of Creative Commons licences 4.0 may alleviate the uncertainty by clearly stating what can and cannot be done with the licensed material.

Finally, four more conclusions can be drawn:

- The exponential growth in data should have little effect on copyright protection, e.g. should not make it either more or less relevant than it currently is. As long as the criteria for copyright protection are applied strictly (either in the form of an ‘originality’ requirement or that of being an ‘author’s own intellectual creation’), there is no reason to think that the copyright regime will lose its relevance.
- On the other hand, in those countries that recognise a *sui generis* regime of protection for databases, the exponential growth in data may entail a greater tendency towards private appropriation of databases. In this case, the application of OA principles to publicly funded scientific output is particularly important.
- The increase of machine-generated data in science (e.g. scientific sensors) may pose particular challenges, especially in trying to determine whether the result qualifies as a subject matter susceptible of intellectual property protection and whether it meets the criteria for protection.
- A related challenge might be the use of machine-generated data to identify a rights owner. This analysis would need to take place on a case-by-case basis. In the best case scenario, the machine-generated data will not qualify as protectable subject matter and may also be devoid of any originality so that it would in principle be free for use by everyone.

Introduction

Innovative scientific research plays a crucial role in addressing global challenges, such as healthcare, environmental, and security issues, while research in social sciences and the humanities occupies a key function in understanding emerging social phenomena²⁸⁴. Evidence shows that access to research data not only increases the returns from public investment in this area, but it also reinforces open scientific inquiry, encourages diversity of opinions, promotes new areas of work and enables the exploration of topics not envisioned by the initial investigators²⁸⁵. Collaborative exchanges help avoid unnecessary duplication of research and give insight into the methodology followed. Timely and cost efficient access to publicly funded scientific research therefore contributes to increasing the general economic and social welfare²⁸⁶. Open access principles are gaining ground among policy makers, research funding agencies, high education institutions and researchers alike as the way forward. ‘Openness’ in science involves granting ‘access on equal terms for the international research community at the lowest possible cost, preferably at no more than the marginal cost of dissemination. Open access to research data from public funding should be easy, timely, user-friendly and preferably Internet-based’²⁸⁷.

Copyright and other intellectual property rights play a decisive role in the way scientific output is being disseminated and used by the scientific community because they underpin the relevant licensing practices. The expansion of open access policies to publicly funded research data raises a number of legal and policy issues that are often distinct from those concerning the publication of scientific articles and monographs. Since open access of research data – unlike publications – is a relatively new policy objective, less attention has been paid to the specific features of research data. This paper therefore gives an overview of the main legal issues and policy considerations involved in licensing both scientific publications and research data under open access conditions. It examines whether copyright law and database protection rights support, impede, enable or are neutral towards the implementation of open access principles for the dissemination of scientific results.

Generally speaking, intellectual property licenses are permissions to use protected subject matter, without which such use would constitute an act of infringement on the owner’s rights²⁸⁸. Parties to a license tailor their contractual arrangements on the basis of the protection granted by the law. A license may therefore be more or less permissive for the licensee, depending on the circumstances and the goals pursued by the parties. Where the law is already permissive, a restrictive license will be deemed acceptable to an informed and equal licensee only if the reduction in permission is compensated in another way (reputation, income, etc.). Where the law is less permissive, a more liberal license will procure clear advantages to the licensee. The degree of permissiveness of an open access license is intrinsically connected to the scope of rights granted in each jurisdiction to scientific output, and the exceptions to those rights. While the implementation of OA principles may be based on contractual arrangements between authors, publishers, research institutions and funding organisations, the legal framework represents an important factor in how these arrangements are formed.

Other elements like the definition of protectable subject matter (e.g. the question of what is or is not protected by an IP right) and the duration of protection do not directly influence the permissiveness of an OA license, but are likely to affect the enforceability of a license. In principle, unprotected or no longer protected subject matter belongs to the public domain and should remain freely available for re-use by anyone and for any purpose. The failure to abide by the terms of a license attached to public domain material does not give rise to an infringement of rights. At most, it amounts to a breach of contract, which generally entails less far-reaching implications for licensees and more limited redress possibilities for the licensor. Because the qualification of the object licensed does not immediately affect the permissions granted under the OA license, this issue will essentially remain outside the scope of this chapter. References to the criteria for protection in the jurisdictions surveyed below are meant to give a rough

indication of the likelihood that research data is protected independently from the scientific publications, and consequently whether it could be the object of a separate open access license.

This chapter is further divided in three parts after this introduction. Section 2 examines the intellectual property protection of scientific output, understood as including scientific and scholarly publications as well as collections of any type of data. After an introduction of the general legal framework, which gives an overview of the main existing intellectual property rights that may attach to scientific output, section 2 further gives a portrait of the laws of several jurisdictions applicable to research results, with a particular focus on copyright law and, where applicable, on the protection of databases. Section 3 follows by giving a brief overview of the open access policies in some of these jurisdictions, making a link with the state of the IP legislation in those countries. Section 4 draws attention to two as yet unresolved issues, namely the problem of open access in the context of public-private partnerships and that of text and data mining.

It is important to realise at the outset, that the promotion of OA in science is essentially a matter of policy on the part of the government or the parties involved in the licensing of research output. The basic rationale behind promoting OA in science is generally that the gains made by society as a whole in having free access to and re-using the output of publicly funded research outweigh the loss of control of the licensor over the material. To demonstrate the soundness of this rationale as a policy choice would have required economic evidence and sociological analysis far beyond the means available for this study.

Scientific output and Intellectual property protection

General legal framework

This section will give a general overview of the international intellectual property regimes that are or may be applicable to the protection of scientific output. Scientific output, understood as including scientific and scholarly publications as well as collections of any type of data, can be protected by a number of rights within the category called intellectual property, chiefly copyright and database protection. These two rights are granted to their owners without any formality, at the moment an original work or qualifying database is created. Many other rights exist in the area of industrial property, such as trademarks, patents, topographies of integrated circuits, design rights, or protection of plant varieties. However, because industrial property rights accrue to their owner only once an active step towards registration has been made, the fulfilment of an open access policy can easily be achieved by not applying for registration. These rights will therefore not be discussed. Before turning to a jurisdiction-specific analysis, let us first provide a general overview of the selected rights at the international level in order to determine their ambit and the scope of their protection.

Copyright

Copyright protects literary and scientific works such as articles, papers, and other types of publications in the scientific field. However, copyright protects the original expressions of ideas, and not ideas themselves. This means that the article, paper, or other type of scientific work has to be expressed in some original form. This basic principle is implicitly enshrined in the oldest and most influential international agreement in the field of copyright, the Berne Convention of 1886 which states that: “The expression “literary and artistic works” shall include every production in the literary, scientific and artistic domain, whatever may be the mode or form of its expression, such as ... ”²⁸⁹.

A more explicit statement can be found in Art. 2 of the WIPO Copyright Treaty (WCT) of 1996, and in the specular Art. 9(2) of the Agreement on Trade Related Aspects of Intellectual Property Rights (TRIPS) of 1994, which respectively establish that: “Copyright protection [shall] extend[s] to expressions and not to ideas, procedures, methods of operation or mathematical concepts as such”²⁹⁰. According to this basic principle of copyright law, it is not sufficient that an idea takes shape in the mind of the researcher or scientist, but that idea needs to be expressed in a form that can be perceived (but does need to be perceived, as un-published works are protected as well) by others; this principle is also known as the idea-expression dichotomy²⁹¹.

The rights granted by copyright law at the international level are quite broad and traditionally encompass the right of reproduction of the work, the right of distribution of the work and of copies thereof, the right of communication of the work to the public including by public performance, the right of rental, and the right of translation, adaptation, arrangement and alteration of the work²⁹². Only the rights relevant for the scope of this study will be dealt with in detail in the next paragraphs.

The right of reproduction is one of the most fundamental rights granted by copyright laws worldwide and gives rights holders the right to reproduce the work in any manner or form²⁹³. The right of reproduction generally covers verbatim copies of a protected work as well as partial or complete reproductions of the work in other formats. In other words, methods of reproduction such as design, engraving, lithography, printing processes, typewriting, photocopying, mechanical or magnetic recording, and all other processes known or yet to be discovered are acts covered by the reproduction right²⁹⁴. The extent to which these acts also include specific forms of adaptation is not entirely clear, and both the Convention and national implementations protect adaptations – or specific types thereof such as translations – either independently or as part of the reproduction right. This is usually due to the

indeterminateness of the boundaries of the reproduction rights, especially in the case of borrowings that are far from literal copies of works²⁹⁵.

The right to distribute copies is not present in the Berne convention²⁹⁶. In 1996, the WCT eventually filled the gap by establishing in art. 6 that “authors of literary and artistic works shall enjoy the exclusive right of authorising the making available to the public of the original and copies of their works through sale or other transfer of ownership”²⁹⁷. Accordingly, all WCT signatories that did not already provide for a general distribution right in their domestic legislations have implemented this provision. The InfoSoc directive of 2001 expressly prescribes this right at the EU level, something that was previously required only for specific subject matter such as computer programs, databases, and some neighbouring rights²⁹⁸. Accordingly, the distribution of copies of scientific articles, books, monographs, studies, or computer programs or databases, are all acts that need to be authorised by the right holder (e.g. scientist, department, university, research centre).

An important limitation to the distribution right is the principle of “exhaustion” or “first sale doctrine”. This principle allows the resale of tangible copies of works (e.g. books) without authorisation from the right holder once the copies have been put on the market with his consent. In absence of this limitation, an author would be able to exercise his exclusive right every time a copy incorporating his work changed owners, thus enabling the author to control not only the first commercialisation of the work but also the derived, second-hand market with consequences on the international movements (imports and exports) of goods²⁹⁹.

At the international level, a technology neutral formulation of the right of communication to the public constitutes a relatively recent addition, as it was introduced in 1996 in Art. 8 WCT. A right to communicate to the public was however already present in some jurisdictions, while others offered individually defined rights such as the right of public performance, broadcasting, and recitation, in accordance with Berne provisions (i.e. Arts. 11, 11bis, and 11ter)³⁰⁰. Art. 8 WCT, however, brings the new concept of making available works “in such a way that members of the public may access these works from a place and at a time individually chosen by them” within copyright control³⁰¹. The main field of application of the new addition is in online and interactive environments, such as the Internet. Accordingly, to upload a work such as a scientific article or a database and offer it for download to Internet users constitutes an act of making available to the public, which needs to be authorised by the right holder. The case law of some jurisdictions has recognised that posting hyperlinks to works already available on websites, however, is not an independent act of communication, unless the link is clickable and it provides access to the work to a new public³⁰².

Acts such as the publication of articles, the making of reproductions, both analogue and digital, as well as the creation of derivative works - for instance a computer program or a database developed from a pre-existing one – are all acts that, at least in general terms, require the authorisation of the right holder unless an exception applies. It must be stressed that the international agreements of reference usually establish a “minimum level of protection” which means that signatory countries are allowed to offer more protection than that which is contained in the international agreements.

Another key characteristic of copyright law is found in the balance it strikes between uses reserved for right-holders and uses that are free, or in other words uses for which no authorisation is required. Two typologies can be found: uses that are not covered by copyright law and uses that are covered but exempted from authorisation, although sometimes “fair compensation” is required.

The former category enlists cases such as the possible requirement of fixation³⁰³; official texts of a legislative, administrative and legal nature, and official translations of such texts³⁰⁴; applied art and industrial designs and models³⁰⁵; political speeches and speeches delivered in the course of legal

proceedings³⁰⁶; and news of the day or miscellaneous facts having the character of mere items of press information³⁰⁷. Except in the latter case, for which an absolute exclusion from protection is mandated, in all the other cases Berne countries can decide whether or not to offer protection to such categories of works.

The second category, “exceptions and limitations to copyright”, generally permit the use of a copyrighted work without the permission of the author or copyright owner. Such exceptions and limitations may be provided by statute or case law, including uses covered by “fair use” or “fair dealing” provisions. The policy rationales for exceptions and limitations in national law vary, including the protection of constitutional and/or fundamental rights, the regulation of industry practice and competition, the dissemination of knowledge, or finally market failure considerations³⁰⁸. Examples of this second category can be seen in cases such as quotation³⁰⁹; illustration for teaching³¹⁰; certain articles on current economic, political, or religious topics³¹¹; and the reproduction of works for the purpose of reporting current events³¹². Once more, given the minimum level of protection approach of the conventions, signatories are free to enact other exceptions and limitations, as long as these apply only to certain special cases that do not conflict with a normal exploitation of the work and do not unreasonably prejudice the legitimate interests of the author³¹³.

Database protection

Databases represent a peculiar subject matter that is protected by copyright under certain circumstances, but in some countries – namely in the EU, Japan, South Korea – is also protected by a so called sui generis database right (SGDR) aimed at the protection of the investment.

The Berne Convention does not mention databases, but provides protection for collections of literary or artistic works such as encyclopedias and anthologies that, by reason of the selection and arrangement of their contents, constitute intellectual creations³¹⁴. The plain meaning of the provision seems to exclude from protection collections that do not consist of works, that is to say, collections of data (databases) are not covered by Art. 2(5). It has been argued that collections of data are in fact covered by the general provision of art. 2(1) as “literary and artistic works”.

In any event, nowadays the protection afforded to databases (as collections of data or other elements) is established – or confirmed – by both Art. 10(2) TRIPS and in the almost identical Art. 5 WCT.

“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 creation shall be protected as such..”³¹⁵

As anticipated, an additional layer of protection is found in some countries and is afforded to databases regardless of the intellectual creation (i.e. “selection or arrangement”) that may or may not be present. What is protected instead is the investment in making the database, i.e. in the obtaining, verification or presentation of the data. This type of right, also known as Sui Generis Database Right (SGDR), is typical of the EU Database Directive, and of the laws of a number of other countries and will be dealt with below. It should be borne in mind that while the protection afforded to original databases focuses on the arrangement or selection without extending to the content of the database, the SGDR offers protection against the copying of substantial parts of the database, that is to say it extends, at least to some extent, to the data themselves.

National legislative framework

This section examines the laws of different jurisdictions as they relate to the protection of scientific publications and research data. In fact, scientific publications, like articles, monographs and reports, are in the vast majority of cases covered by copyright law, as original works of authorship. The question becomes

more problematic in the case of research data such as survey results, datasets etc. This type of information is most of the time excluded from copyright protection in so far as the compilations in which they are gathered do not meet the requirement of originality. However, the compilations or databases may be protected in those countries that recognise a *sui generis* database right. This explains why, in the pages below, greater attention is put on research data as objects of protection. The jurisdictions examined below are a sample of the Contracting parties of the OECD and include countries from the European Union, the United States of America, Australia, Canada, Israel, Japan and South Korea.

EU legal framework

Copyright protection

Whereas scientific publications, such as articles and monographs, virtually always attract copyright protection under the copyright laws of the Member States of the EU, the individual research data and the datasets containing them may not so easily fall under the copyright regime. Since copyright does not protect mere facts and ideas, but rather applies to the original expression of ideas, research data is not likely to qualify as protectable subject matter for lack of originality. The concept of originality in copyright law has been harmonised at the European level with respect to software³¹⁶, databases³¹⁷ and photographs³¹⁸, and more recently extended to all kinds of works through the interpretation of the Court of Justice of the European Union (CJEU), which established that a work is original if it is the ‘author’s own intellectual creation’³¹⁹.

To be eligible for copyright protection, collections of data, tables and compilations must therefore show a sufficient degree of originality in their selection or arrangement, that is to say if through this selection or arrangement the author was able to express his free and creative choices³²⁰. Whether collections of scientific research data meet the criterion of originality is a question of fact to be determined on a case-by-case basis. However, if the selection or arrangement of the contents of a scientific database is dictated by technical factors or imperatives of accuracy and exhaustiveness, then the author can exercise little to no creativity or discretion in the choice, sequence and combination of the data in the collection. Scientific databases are therefore in most cases not likely to meet the threshold for copyright protection.

The Information Society Directive confers on rights owners the exclusive right to reproduce, communicate to the public and distribute their works. Member States must ensure that the reproduction right covers both the direct and indirect and the temporary and permanent making of reproductions by any means, in whole or in part, both to authors of works and to owners of neighbouring rights³²¹. This broad definition is however limited by the mandatory exception for transient and incidental reproductions contained in art. 5(1)³²². The Directive also contains a list of exceptions on these exclusive rights, the most relevant of which is article 5(3)(a) that allows Member States to provide for exceptions in the case of ‘use for the sole purpose of illustration for teaching or scientific research, as long as the source, including the author’s name, is indicated, unless this turns out to be impossible, and to the extent justified by the non-commercial purpose to be achieved’. Like the vast majority of other exceptions in the Directive, this exception is optional; Member States may decide whether to implement it or not³²³. As a result, Member States have different rules and regulations in this context, where some countries recognise no research exception at all (like The Netherlands and Spain). As a result the research exception is generally vague and unevenly implemented at national level, which may put some researchers at a disadvantage³²⁴.

In its 2008 Green Paper on Copyright in the Knowledge Economy³²⁵, the Commission initiated a round of consultations among stakeholders to discuss whether an approach based on a list of non-mandatory exceptions was still adequate in the light of evolving Internet technologies and the prevalent economic and social expectations³²⁶. This consultation resulted in the publication of a Communication to the European Parliament and the Council on Copyright in the Knowledge Economy³²⁷. As the Green Paper

that preceded it, the Communication addressed several aspects of copyright in the knowledge economy, putting particular emphasis on the exceptions for the benefit of libraries and archives, including the issue of orphan works, teaching and research, persons with disabilities, and user-created content (UCC).

At the beginning of 2014, the Commission launched another round of consultations on the reform of the European copyright regime, asking again about the adequateness of the research exception. In its Report on the responses to the Public Consultation on the Review of the EU Copyright Rules³²⁸, the Commission reports on the respective position of the different stakeholders. While end users believe that open access is a suitable solution to increase access to research content that should be better supported, institutional users plead for a mandatory and technology-neutral research exception to be adopted at EU level. Publishers on the other hand argue in favour of licensing solutions. The future will tell whether the research exception will be modified as part of the copyright reform.

Database protection

Collections of scientific works, data, or other materials arranged in a systematic or methodical way and individually accessible electronically or by other means may be protected under the European *sui generis* database right (SGDR). Through article 7 of the Database Directive, as implemented in the legislation of the Member States, the maker of a database showing a substantial investment (assessed qualitatively and/or quantitatively) in either the obtaining, verification or presentation of its contents has the exclusive right to prevent the extraction and/or re-utilisation of the whole or of a substantial part, evaluated qualitatively and/or quantitatively, of the contents of that database. Like copyright protection, the *sui generis* database right arises automatically, without any formal requirement, at the time of completion or disclosure of the database to the public.

Where the ‘obtaining, verification or presentation’ of research datasets does involve the substantial investment necessary to qualify for protection, the *sui generis* protection confers two transferable rights on the maker of a database: the right of extraction and the right of re-utilisation of substantial parts of the database, which are respectively defined as follows: ‘(a) ‘extraction’ shall mean the permanent or temporary transfer of all or a substantial part of the contents of a database to another medium by any means or in any form; (b) ‘re-utilisation’ shall mean any form of making available to the public all or a substantial part of the contents of a database by the distribution of copies, by renting, by on-line or other forms of transmission’.

These two concepts have received a broad interpretation from the ECJ³²⁹. In the *Directmedia* case, the Court found that an act of ‘extraction’ occurs when all or part of the contents of the database concerned are transferred to another medium, whether of the same or of a different nature. Such a transfer implies that all or part of the contents of a database can be found in a medium other than the original database³³⁰. In the view of the ECJ, it is immaterial whether the transfer is based on a technical process of copying the contents of a protected database or on a simple manual process; similarly, it is irrelevant that the transfer of the contents of the database may lead to an arrangement of the elements that is different from that in the original one. The Court adds that the transfer of material from a protected database to another database following an on-screen consultation of the first database and an individual assessment of the material contained in that first database is also capable of constituting an extraction.

In the *British Horseracing Board* decision, the Court ruled that the concept of ‘re-utilisation’ must ‘be understood broadly, as extending to any act, not authorised by the maker of the database protected by the *sui generis* right, of distribution to the public of the whole or a part of the contents of the database’. The Court observed that the nature and form of the process are of no relevance in this respect³³¹. Recently, the Court of Justice reaffirmed its broad interpretation of the concept of ‘re-utilisation’ in a case involving the display of information generated as a result of a search in a dedicated meta search engine³³². The technique

employed by a dedicated meta search engine to crawl the targeted databases for specific information, although not identical, is probably comparable to some of the techniques used to text and data mine databases for research purposes: both types of searches make it possible to search the entire contents of that database even if only part of the database is actually consulted and displayed.

The Database Directive does not provide any definition of the terms ‘substantial’/‘insubstantial’ or ‘qualitative’/‘quantitative’. Intuitively, the first method to assess the substantiality of the part extracted and/or re-utilised consists in comparing quantitatively the amount of data taken with the total amount of data contained in the database as a whole. Accordingly, the ECJ decides that a quantitatively substantial part corresponds to ‘the volume of data extracted and/or reutilised and must be assessed in relation to the volume of the contents of the whole of that database’. Furthermore, the Court observes that ‘if a user extracts and/or reutilises a quantitatively significant part of the contents of a database whose creation required the deployment of substantial resources, the investment in the extracted or re-utilised part is, proportionately, equally substantial’³³³. It has been argued that the substantiality of a part can also derive from its economic value, namely from the price that would be paid for such a part, which would rise proportionately with the rate of investments incurred in the obtaining, verification and/or presentation of the part’s contents. But such a direct correlation cannot always be traced, and the Directive recognises a *sui generis* protection on the basis of the investment made in the database’s production rather than in its economic value.

The protection under the *sui generis* right lasts for 15 years from the first of January of the year following the date on which the database was completed. The term of protection for a database may start anew under two conditions, both involving the term ‘substantial’. The first one is a substantial modification of the contents of the database, evaluated either qualitatively or quantitatively, which can consist of additions, deletions or alterations (including rearrangement of the contents). The second is a substantial investment, evaluated qualitatively or quantitatively. This is one of the most controversial and criticised provisions of the Directive since it apparently offers grounds for a perpetual protection of the databases³³⁴.

Finally, it is worth pointing out that, according to article 11 of the Database Directive, only natural persons who are nationals of a Member State or who have their habitual residence in the territory of the EU can benefit from the database right. Furthermore, companies and firms are also entitled to such protection if they are formed according to the law of a Member State and have their registered office, central administration or principal place of business within the EU. Article 11.2 clarifies that in case a company or a firm has a registered office only in the territory of the EU, its operations must be substantially and durably linked with the economy of a Member State. In other words, the protection of the *sui generis* database right is not only unique to Europe in that it is conferred only on EU nationals, whether natural or legal persons, but also because there is no real comparable regime of protection for non-original databases outside the European Union³³⁵.

The Database Directive contains a separate set of exceptions for copyright and the database right. With respect to copyright, Article 6(1) contains a mandatory exception on copyright stating that the lawful user of a database may perform, without prior authorisation, any act covered by Article 5 necessary for the purposes of access to the content of the databases and normal use of the content. Article 6(2) allows Member States to provide for limitations on the copyright owner’s exclusive rights, including the right to make reproduction of a non-electronic database for private purposes and to use it for the sole purpose of illustration for teaching or scientific research, as long as the source is indicated and to the extent justified by the non-commercial purpose to be achieved³³⁶. Since Article 6(2) is optional, Member States have either implemented it in various ways or not at all³³⁷.

With respect to the *sui generis* database right, Article 8(1) states that ‘the maker of a database which is made available to the public in whatever manner may not prevent a lawful user of the database from

extracting and/or re-utilising insubstantial parts of its contents, evaluated qualitatively and/or quantitatively, for any purposes whatsoever'. Article 9 recognises the same optional exceptions on the *sui generis* as in Article 6, but limited to the right of extraction. This means that, where implemented, the substantial extraction of the content of a database is allowed for research purposes, but that no act of re-utilisation can be performed. This restriction actually removes any practical value of the research exception on the database right³³⁸.

The application of Articles 6 and 9 rests on the concept of a lawful user: only a lawful user may benefit from the exceptions of Article 6(1), 8(1) and 9, while the exceptions listed in Article 6(2) extend to anyone. The concept of 'lawful user' is nowhere defined in the Directive. A literal interpretation suggests that once the right holder makes the database available to a user, s/he is deemed to be a lawful user³³⁹. Access may, however, be conditioned by the terms of use or other contractual agreements set by the rights holder. In such a case, "contractual agreement" would need to be interpreted in a broad manner. The use of freely available online databases (websites in many instances), even in the absence of any specific terms of use, on the basis of an implied authorisation, may also qualify as a lawful use, as long as the database is published by (or with the consent of) the rights holder³⁴⁰.

France

In France, copyright is regulated in the *Code de la Propriété Intellectuelle* (Intellectual Property Code, IPC³⁴¹), which deals with literary and artistic property as well as with industrial property. The object of protection is an original work of the mind (*oeuvre de l'esprit*), which enjoys rights of economic exploitation and moral rights³⁴². In particular, original books, as well as other literary, artistic, and scientific writings, conferences and sermons, works of photography, computer programs, and databases are illustratively listed as protected works³⁴³.

Article L122-1 establishes that authors' rights of exploitation include the right of representation (*représentation*) and the right of reproduction. The former covers acts of communication to the public by any means, including by public presentation, public transmission, and any other diffusion throughout a telematics network³⁴⁴. The right of reproduction consists in the material fixation of the work through any process, such as printing, drawing, photography, and any other mechanical or magnetic recording³⁴⁵.

Exceptions and limitations to copyright are regulated under Article L122-5 IPC. Particularly relevant for the purpose of this study is the exemption of short quotations and analysis justified by the educational, informational or scientific character or for criticism, provided that the name of the author and the source are indicated³⁴⁶. Also allowed – under the same general obligation to cite the author and the source – are the representations or reproductions of excerpts of works for the scope of illustration for teaching and research, provided that the public for which such representations or reproductions are intended is mainly composed of pupils, students, teachers or researchers, that there is no commercial exploitation, and under the obligation of a fair compensation for authors. Works conceived for pedagogic purposes are also excluded from the exception.

Regarding the *sui generis* database protection (SGDR), Art. L341-1 requires a financial, material or human substantial investment in the constitution, verification or presentation of the database in order to grant the relative right to the maker. In a number of decisions, French courts have applied a low threshold concerning the requirement for substantial investment, although many of them pre-date the 2004 ECJ saga, which established that investments in the creation of the data do not qualify for SGDR protection³⁴⁷.

Regarding the exceptions and limitation that apply to databases, a few observations can be formulated. Art. L122-5 paragraph 5 of the IPC states that the acts necessary for the purpose of access to the contents of an electronic database and within the limits of the use regulated by contract do not require

the author's authorisation. Two elements of the French provision are not present in Art. 6 of the Directive: the limitation only to electronic databases, and the presence of a contract – the directive employs the term *lawful user*, which includes users who lawfully use the database also in absence of contractual provisions.

The French transposition does not explicitly implement Art. 15 of the Directive, which declares null and void any contractual provision in violation of Article 6(1). Regarding the exceptions for illustration for teaching and scientific research³⁴⁸, the general provision of Article 122-5(e) is applicable to all copyright-protected subject matter and therefore also to databases. The only observation is that the French provision requires fair compensation for such uses, something that the Directive does not mention. As to whether other exceptions and limitations that are traditionally authorised under national law (Article 6(2)(d) of the Directive) are applicable under French law, no specific reference is made in the Code. It seems that an interpretation in line with the general principles of law would admit that pre-existing exceptions that apply to copyright should apply also to databases, as long as those are protected by copyright.

Regarding exceptions to the SGDR, the French legislator took the opportunity to introduce an exception for private purposes for cases of *extraction* of non-electronic databases (Article 342-3(2)), and for extraction and re-utilisation for teaching or scientific research purposes (Article 342-3(4)), although once again requiring fair compensation. The exception for public security or administrative or judicial procedures is to be found in a different part of the Code, which applies generally to all rights covered by that section (copyright, neighbouring rights and the SGDR). The wording of such exceptions is consequently slightly different from that contained in the Directive.

An important observation is the omission by the French legislator in Article L342-3 of provisions dealing with the exceptions and limitations to the SGDR of the lawful user and it should therefore be concluded that such exceptions are available to any user.

Finally, France has introduced a new exception to the SGDR on the basis of the implementation of the Information Society Directive. Such exception applies to legal persons and publicly accessible institutions such as libraries, which are allowed to make extractions and re-use substantial parts of a database for personal consultation by handicapped persons and for non-commercial purposes (Article L342-3(3)).

Germany

The German Copyright Act (UrhG) protects literary, scientific and artistic works, among which it illustratively lists writings, speeches, musical works, choreographies, works of fine art, works of photography, cinematographic works, and illustrations of a scientific or technical nature³⁴⁹. The list is non-exhaustive and is followed by a general purpose clause, defining copyright-protectable works as “personal intellectual creations” of their authors³⁵⁰.

In order to establish that a work is a personal intellectual creation four cumulative criteria are generally considered essential. According to these criteria a work is a “personal creation” (the work has to originate from a person) with an “intellectual content” (an idea that is expressed) in a “perceptible form” (the work needs to be expressed in a form that can be perceived) which has a certain degree of “individuality” (the work has specific characteristics that make it original). This does not mean that works need to be novel or unique, but they need to be creative and individual to the author. Mere ideas are not protected³⁵¹.

Art. 4 of the German Copyright Act grants copyright protection to databases, i.e. to collections of works, data, or other independent elements which, by reason of their selection or arrangement, constitute the author's own intellectual creation³⁵². Protection of course does not extend to the constituent elements of the collections or databases, which can be both protected works or unprotected items. The decisive

consideration is whether the collection itself constitutes an intellectual creation³⁵³. Accordingly, the selection or arrangement of the database elements must constitute an intellectual creation and the work must have reached a certain level of originality. This level of originality has to be found in the structure of the database i.e. in the selection or arrangement of its content³⁵⁴.

According to Article 7 UrhG, the author is the creator of the work. Since a personal creation is required, the author of a work can only be a natural person³⁵⁵. It is worthwhile to note that, in German copyright law, transfers of copyright to others are not contemplated as article 29 I UrhG prohibits this type of alienation. However, it is possible to assign rights of use as stated by Article 31 I UrhG.

Authors' exclusive rights are listed in Arts. 15 et seq. and comprise the exclusive rights of reproduction (Art. 16 UrhG), distribution (Art. 17 UrhG), exhibition (Art. 18 UrhG), and the right of communication to the public (Arts. 19 - 22 UrhG). Extremely relevant in the context of this report is a recent addition to the Act, in Art. 38 UrhG, which deals directly with the issue of licensing scientific publications created as a result of public funding and reads as follows:

The author of a scientific contribution which is the result of a research activity publicly funded by at least fifty percent and which has appeared in a collection which is published periodically at least twice per year has the right, even if he has granted the publisher or editor an exclusive right of use, to make the contribution available to the public in the accepted manuscript version upon expiry of 12 months after first publication, unless this serves a commercial purpose. The source of the first publication shall be indicated. Any deviating agreement to the detriment of the author shall be ineffective³⁵⁶.

This provision is intended to allow the author of a scientific work that is generated in the context of publicly funded – at least 50% – research and published in a periodical collection (at least biannual), to make the accepted version of the manuscript publicly available for non-commercial purposes after an embargo period of twelve months. The right to republish cannot be limited by contractual agreements, which means that even if the author has licensed all her exclusive rights to a publisher, she will still be entitled to the right of republication³⁵⁷.

Exceptions and limitations to copyright are present in Sec. VI (Arts. 44a to 63a). Of particular relevance for this study is the exception of quotation provided by Article 51 UrhG that allows the reproduction, distribution, and communication to the public of a published work for the purpose of quotation (e.g.: inclusion in an independent scientific, literary or musical work) to the extent justified by such a purpose.

The sui generis database right (SGDR) is implemented in Sec. VI of the Act, Arts. 87a-e. The implementation of the sui generis right follows closely the requirements already seen when analysing the EU framework. Of particular interest in the German implementation is the provision regarding the maker of the database, identified as the person who has made the investment, with no explicit reference to the “risk” factor³⁵⁸. However, it has been noted that the *Datenbankhersteller* is in fact the person (including legal person) who bears the risk, as the risk element is implied in the concept of investment³⁵⁹.

Regarding the level of substantial investment, German courts have fully implemented the ECJ rulings that limit investments to acts connected with the obtaining, verification and presentation of the data, and not with the creation of the data themselves³⁶⁰. However, the investment so characterised does not need to be particularly high and even “routine” investments such as the maintenance of the database software, and verification of the entries submitted by third parties qualify³⁶¹.

The German Supreme Court (Bundesgerichtshof, BGH) ruled that the extraction of 75% of a database corresponds to a substantial extraction in quantitative terms and therefore violates the scope of the SGDR³⁶². Similarly, the same Court found that an extraction of one-tenth of a database's content does not constitute a substantial reproduction in quantitative terms and therefore is permitted by the law³⁶³. While this type of numerical pointer can prove particularly helpful in determining the boundaries of the right and consequently what constitute a licit reproduction, it must also be borne in mind that investments that are qualitatively substantial can constitute an infringement. Therefore, an extraction of 10%, when this 10% is qualitatively substantial (because it corresponds to a quantitative or qualitative substantial investment) can be found to infringe the SGDR.

Art. 87c lists the limitations to the SGDR and states that acts of reproduction of substantial parts of the database are permissible in cases of private use (but only in the case of non-electronic databases), personal scientific use (but only to the extent justified by that use and for non-commercial purposes), and illustration for teaching (as long as there is no commercial purpose). A general exception to the rights of reproduction, distribution and communication to the public is also permitted for cases of public order such as court proceedings and other authority mandated uses.

The Netherlands

Article 10 of the Dutch Copyright Act (DCA) contains a non-exhaustive list of categories of works that are literary, scientific and artistic works, and are protected under the DCA provided they are sufficiently original. This includes original adaptations (Article 13 DCA). The criterion of originality is not specified in the Act but has been recognised as a requirement for protection by the courts in several cases³⁶⁴. “Originality” (or “oorsponkelijkheid”) is in Dutch practice typically used as shorthand for a two-pronged test elaborated by the Supreme Court. The work must have an “own, individual character” and “bear the personal stamp of the author”³⁶⁵.

Exclusive rights of the author are governed by either one of two broad concepts: the right to make a reproduction, or “verveelvoudiging”, pursuant to Article 13 DCA and the right to communicate to the public, or “openbaarmaking”, pursuant to Article 12 DCA. The right to make a reproduction encompasses two elements: the right to make copies of a work of authorship, and the right to make adaptations, arrangements, interpretations, translations, or any form of conversion of the work into another work. Actually Article 13 of the Act only refers to the second element of the right of reproduction. Because the issue was so obvious to the Dutch legislator, the first element of the right of reproduction is nowhere explicitly provided for in the Act.

Chapter 6 of the DCA contains numerous limitations and exceptions allowing unauthorised use of protected works for different purposes, by different types of users and under different conditions. The most relevant in the context of this study are the right of quotation (Article 15a), the right to use works communicated by or on behalf of a public authority (Article 15b), and the educational use exception (Article 16). Unique to the DCA and relevant in the context of scientific works produced or disseminated by or on behalf of public authorities is Article 15b, which reads as follows:

The further communication to the public or reproduction of a literary, scientific or artistic work communicated to the public by or on behalf of the public authorities shall not be deemed an infringement of the copyright in such a work, unless the copyright has been explicitly reserved, either in a general manner by law, decree or ordinance, or in a specific case by a notice on the work itself or at the communication to the public. Even if no such reservation has been made, the author shall retain the exclusive right to have works of his that have been communicated to the public, by or on behalf of the public authorities, appear in the form of a collection.

This exception is barely invoked in practice and has led since its introduction in the Copyright Act in 1975 to only three judicial decisions³⁶⁶.

Another exception relevant to the use and dissemination of scientific works is the educational use exception. Long before the implementation of the Directive, the DCA allowed the “taking over of parts of works” for teaching purposes, pursuant to Article 16. Article 16 § 1a gave examples of possible acts falling under the scope of the exception. These include publications, sounds or video recording. Whether these means of reproduction included digital reproduction or online communication was highly uncertain. As a result of the implementation of the Directive, Article 16 of the DCA has been made technology-neutral/independent, so that digital reproductions are also covered, as well as acts of making a work available to the public. Accordingly, all reproductions and communications that comply with the conditions set out in the article are in principle covered. Notably, Article 16 of the DCA contains three additional criteria that do not appear in Article 5(3)(a) of the Directive: 1) the work from which the part is taken must have been published lawfully; 2) the adoption must be in accordance with that which might reasonably be accepted under the rules of social custom; and 3) moral rights have to be observed. In addition, educational use requires that the source of the work used be indicated. Furthermore, and in contrast to the Directive, educational uses have always been and remain permissible under Dutch law provided that an equitable remuneration is paid to the rights owners. Note however, that the Dutch Copyright Act contains no specific research exception.

Pursuant to Article 1a of the Database Act, and in conformity with the Database Directive, a database is protected if it constitutes “a collection of independent works, data or other materials arranged in a systematic or methodical way and which elements are individually accessible by electronic or other means and for which the acquisition, control or presentation of the contents, evaluated qualitatively or quantitatively, bears witness to a substantial investment”. Unconvinced of the need to grant protection for databases beyond that provided by the Dutch copyright regime, Dutch courts have tended, since the very adoption of the Database Act, to interpret the requirement of “substantial investment” rather restrictively. Since the ECJ’s decision, Dutch courts have followed the ECJ’s creation/collection dichotomy and are in general reluctant to conclude that there has been a substantial investment in the absence of clear evidence put forward by the alleged right holder that he incurred substantial costs in the collection, verification or presentation of the content of the database³⁶⁷.

The beneficiary of the protection is not defined in the Dutch Database Act. There is therefore no reference in the Act to any kind of risk of investing. Dutch commentators seem to agree, however, that the *sui generis* right “should only be conferred as a reward for the risks taken by the innovating industry which actually produce the databases, not the commissioning parties³⁶⁸. Other commentators posit that employees, subcontractors or anyone else executing the work without bearing the financial responsibility for the end product cannot be considered as a database right holder³⁶⁹.

The Database Directive is silent on the issue of the source of the funding or the role of public money in the acquisition of rights. The Netherlands is so far the only Member State to have explicitly regulated the exercise of the *sui generis* rights by public sector bodies. Article 8 of the Dutch Database Act denies a public authority the right to exercise its exclusive database rights unless the right is reserved explicitly by a general mention in an act, order or ordinance, or in a specific case by notification on the database itself or while the database is made available to the public. The specific question of the impact of public funding on the nature of the investment made to produce a database was the object of a ruling from the Raad van State (Dutch Council of State) in the *Landmark* case³⁷⁰. In the court of first instance, the District Court of Amsterdam found that the local authorities of Amsterdam could not be qualified as the producer of the database, of which the list of addresses was part, since the production of the database was initiated and financed by the local authorities of Amsterdam, a large part of which was financed by the Ministry of

Housing, Spatial Planning and the Environment. This ruling was confirmed on 29 April 2009 by the Raad van State. The Dutch position is so far unique within the EU.

The Dutch legislator transposed Article 9 of the Directive on permitted exceptions to the database right literally. In addition to these exceptions, the Dutch Database Act contains in Article 8(2) a unique provision, modelled after Article 15b of the Copyright Act. The provision reads:

The right, referred to in Article 2, paragraph 1 shall not apply to databases for which the public authority is the producer, unless the right is expressly reserved either in general by law, order or resolution or in a particular case as evidenced by a notification in the database itself or when the database is made available to the public.

The main rationale behind this provision is that when transposing the Database Directive into Dutch law, the Dutch legislator did not want to put the makers of a database in a better position than authors of copyright-protected works. In other words, since works that are made available by or on behalf of public authorities are covered by the exception of Article 15b of the Copyright Act, there is no reason not to apply a similar exception to databases produced by public authorities.

United Kingdom

In the United Kingdom copyright law is regulated pursuant to the Copyright, Design and Patents Act 1988 (CDPA 1988). The central element of copyright protection is the copyright work characterised as literary, dramatic, musical, and artistic works, films, sound recordings, broadcasts, and published editions (or typographical works). Differently from many other Copyright Acts the UK's lists exhaustively the categories of works that can qualify for copyright protection; that is to say, under UK law a work, in order to enjoy copyright protection, has to fall within one of the eight categories above listed³⁷¹.

Scientific works, although representing a heterogeneous class whose components could fall within more than one of the eight mandatory categories aforementioned, will most likely be included in the first one: literary works. Literary works, in fact, are defined as any work, other than a dramatic or musical work, which is written, spoken or sung, and accordingly includes a) a table or compilation other than a database; b) a computer program; c) preparatory design material for a computer program; and d) a database³⁷². Literary works are not limited to works of literature, but include all works spoken or written, – with the obvious exclusion of dramatic or musical works – as well as symbols and numerals irrespectively of any quality or aesthetic value³⁷³. Literary works include explicitly works that are spoken such as lectures, conference talks and key-notes, yet, as indicated by Sec. 3(2), copyright does not subsist in a literary [...] work unless and until it is recorded, in writing or otherwise.

In accordance with the EU framework, the rights granted by the CDPA are the right of reproduction, distribution, communication to the public, adaptation, rental and lending rights, and the right to perform, show or play the work in public. These rights may be available only for some though not all of the protected subject matter, but for our main purpose – scientific works falling within the literary work category – it is safe to assume that all the aforementioned rights are available to right holders³⁷⁴.

Databases are protected as a distinct class of literary works on the basis of a definition that closely resembles the one provided by the Database Directive. The definition is broad enough to cover most if not all of the material that previously was protected as table and compilations³⁷⁵. The sui generis database right was implemented in the Copyright and Rights in Databases Regulations 1997 amending the parts of the CDPA 1988 regarding the protection offered to databases by copyright (Part II) and regulating in detail the SGDR (Part III) following the framework set forth by the EU Database Directive. In particular the database maker is the person who takes the initiative in obtaining, verifying or presenting the contents of a database

and assumes the risk of investing in that obtaining, verification or presentation³⁷⁶. Fair dealing with substantial parts of a database extracted by the lawful user for the purpose of illustration for teaching or research and not for any commercial purpose and provided that the source is indicated is not an infringement of the database right in a database³⁷⁷.

The UK, as many common law systems, has a partially different approach to exceptions and limitations to copyright, if compared with the national examples – mostly from civil law tradition – thus far analysed. This difference lies in a general provision allowing a certain number of uses so long as they are “fair”. According to Sec. 29 and 30, fair dealing with a [literary, dramatic, musical, or artistic³⁷⁸] work does not infringe any copyright in the work so long as the purpose of the fair dealing is: a) research or private study; b) criticisms or review; or c) news reporting. In order to benefit from this copyright infringement liability exemption it must be shown that the dealing (any use of a work) was fair, which is often a matter of “degree and impression”³⁷⁹. Factors to be considered in order to establish the fairness of someone's use include whether the work was published, how the work was obtained, the amount taken, the use thereof made, the motives and consequences of the dealing, and alternatives to the dealing, i.e. whether the purpose could have been reached by other means³⁸⁰. In addition to the fair dealing provisions, a number of other exceptions and limitations are present in the CDPA. Two that are relevant for our present purpose are contained in Sec. 32 and allow the preparation of copies of a work in the course of, or preparation for, instruction. Another relevant provision, Sec. 50D, implements the EU made rule that the legitimate user of a database does not infringe any copyright when doing any acts which are necessary for the purpose of access and use of the content of the database; this defence cannot be contracted out.

The UK has recently implemented³⁸¹ a number of amendments to the CDPA, which are expected to facilitate the conduct of scientific research and analysis. Of particular interest is the new section 29A, which reads as follows:

“29A Copies for text and data analysis for non-commercial research

(1) The making of a copy of a work by a person who has lawful access to the work does not infringe copyright in the work provided that —

(a) the copy is made in order that a person who has lawful access to the work may carry out a computational analysis of anything recorded in the work for the sole purpose of research for a non-commercial purpose, and

(b) the copy is accompanied by a sufficient acknowledgement (unless this would be impossible for reasons of practicality or otherwise).

(2) Where a copy of a work has been made under this section, copyright in the work is infringed if —

(a) the copy is transferred to any other person, except where the transfer is authorised by the copyright owner, or

(b) the copy is used for any purpose other than that mentioned in subsection (1)(a), except where the use is authorised by the copyright owner.

(3) If a copy made under this section is subsequently dealt with —

(a) it is to be treated as an infringing copy for the purposes of that dealing, and

(b) if that dealing infringes copyright, it is to be treated as an infringing copy for all subsequent purposes.

(4) In subsection (3) “dealt with” means sold or let for hire, or offered or exposed for sale or hire.

(5) To the extent that a term of a contract purports to prevent or restrict the making of a copy which, by virtue of this section, would not infringe copyright, that term is unenforceable.”

The provision basically clarifies that making a copy of a work for the purpose of text and data mining (TDM) is not an infringement of the copyright in the work provided that this is made for the sole purpose of research for non-commercial purposes. The provision has also the upside to make clear that any contractual agreement that has the effect of limiting the possibility of making copies under this provision is unenforceable. The exception does not cover the SGDR, however. It is the opinion of the UK government that the SGDR's fair dealing exception for non-commercial scientific uses offers a parallel defence adequate to the present needs³⁸².

United States of America

The intellectual property clause of the U.S. Constitution authorises Congress to enact copyright legislation with the objective to promote science and useful arts by giving authors the exclusive rights to their writings for a limited time³⁸³. The present Copyright Act, which was last given a comprehensive update in 1976, rests on this constitutional basis³⁸⁴. Sec. 102 of the Copyright Act affords copyright protection to original works of authorship fixed in any tangible medium of expression, now known or later developed, from which they can be perceived, reproduced, or otherwise communicated, either directly or with the aid of a machine or device.

The section proceeds with non-exhaustive lists of examples of what can constitute a work of authorship: literary, musical, dramatic, choreographic, pictorial, audio-visual, architectural works, and sound recordings³⁸⁵. The concept of originality under U.S. copyright law has been defined as requiring a modicum of creativity, therefore excluding from protection the mere investment of time and effort in the creation of works (the so-called “sweat of the brow” doctrine)³⁸⁶. Significantly, the second part of Sec. 102 excludes from protection any idea, procedure, process, system, method of operation, concept, principle, or discovery, regardless of the form in which it is described, explained, illustrated, or embodied in such work. Another major area where copyright cannot apply is in United States government works: Sec. 105 establishes that copyright protection is not available for any work of the United States Government³⁸⁷.

Any work created by an officer or employee of the US government as part of that person's official duties is not subject to copyright within the United States. Consequently, the majority of works, including original databases, produced by US government agencies are in the public domain and can be used and re-used for any purpose without fear of copyright infringement. Two main limitations to this rule need, however, to be noted: a) the rule applies only to works of the US federal government – the works of states and local authorities are protected by copyright; b) the rule operates only for US works and only within US borders – the US government reserves the right to enforce copyright in its own works outside the United States³⁸⁸. US copyright grants authors the exclusive rights to reproduce their work in copies and phonorecords; to prepare derivative works based on the copyrighted work; to distribute copies to the public by sale, rental, lease or lending; and to perform and display the work publicly³⁸⁹.

Among the main innovations that the present Copyright Act introduced was the codification of the so-called “fair use” doctrine. The fair use doctrine was originally developed by US courts and only later codified in § 107 of the Copyright Act 1976. The doctrine is characterised by an open-ended list of purposes for which the use of a work may be regarded as fair, marked by the words ‘such as’, and by four factors to be considered in determining whether or not a particular use is fair.

A recent application of the fair use doctrine can be seen in two cases that involved the Google Books project. The “Google Books Program” actually consists of two components: the “Partner Program” involving the hosting and display of material provided by book publishers or other rights holders, and the “Library Program” involving the digital scanning of books in the collections of several public and university libraries. These programmes entail several activities, including making text available with rightsholder authorisation, and offering the tools for online searching of the content of the books and displaying snippets of the books. *Whether the display of “snippets” is fair use in Authors Guild v. Google is still to be determined, as the case is currently on appeal to the Second Circuit.*

After the rejection of a proposed settlement between The Authors Guild and Google, The Authors Guild continued its lawsuit against Google and at the same time sued HathiTrust, a partnership of major academic research libraries that relies on Google Books Search to create a digital archive of library materials (the HathiTrust Digital Library, or “HDL”). Works within the HDL are used for three purposes: (1) full-text searches; (2) preservation; and (3) facilitating access for print-disabled persons. In both cases, a District Court had to determine whether digitisation of books is a fair use of copyrighted material. The decisions were rendered by different judges (on October 10, 2012³⁹⁰ and November 14, 2013³⁹¹ respectively), both of whom ruled against the Authors Guild and in favour of applying the fair use doctrine.

The Mass Digitisation Project is referred to in the Authors Guild, Inc. et al v. Hathitrust et al case³⁹² and is a project through which Google creates digital copies of works in the Universities' libraries in exchange for which Google provides digital copies to Hathitrust. Considering the different goals of the Mass Digitisation Project under the first fair-use factor³⁹³, the Court stressed that these were to be considered as transformative uses, referring – amongst others – to the new areas and methods of research, such as text mining, that the digital copies enabled. Although one might have expected Google's fair use defence to be weaker than the libraries', Judge Chin in *Authors Guild v. Google* equally ruled that Google's use of the copyrighted works in the context of its book scanning and indexing project constitutes “fair use” under copyright law. The court reasoned that Google's digitisation of books is “highly transformative,” adds value, serves several important educational purposes, and may enhance the sale of books to the benefit of copyright owners.

The fact that Google Books facilitates search, offering an important tool for readers, scholars, researchers, libraries and others to identify and find books, and opens up new fields of research, in particular through text mining, was put forward to demonstrate the transformative character of Google's use of the copyrighted works. In *Authors Guild v. HathiTrust*, the Court refers in a footnote to text mining, noting that “[m]ass digitisation allows new areas of non-expressive computational and statistical research”. This may or may not apply to copying entire digital databases that are already “mass digitised” and created for purposes that include text mining. Admittedly, the Court did not address any intermediate copying activities by TDM researchers themselves. However, considering the outcome of both *Authors Guild v. HathiTrust* and the lower court decision in *Authors Guild v. Google* – concluding that HathiTrust and Google's use of the copyrighted works qualified as fair use – there is precedent indicating that copying acts by TDM researchers for the purpose of extracting non-expressive metadata, within certain parameters, could be considered fair use under US law. However, as is true of all fair use cases in the US, these decisions are highly dependent on the particular facts before the courts. Moreover, these cases are not binding on federal courts outside the Second Circuit.

Canada

The Canadian Copyright Act³⁹⁴ regulates copyright at the federal level since its original enactment in 1921³⁹⁵. Protection is afforded to every original literary, dramatic, musical, and artistic work and grants the exclusive rights to produce or reproduce the work or any substantial part thereof in any material form whatever, to perform the work in public or, if the work is unpublished, to publish the work or any

substantial part thereof³⁹⁶. The Supreme Court of Canada has ruled that an original work is the expression of an idea through an exercise of skill and judgment³⁹⁷.

Contrary to the United States, where works produced by the federal government are not protected by US copyright law, Canada's Copyright Act provides that copyright in works prepared or published by or under the direction or control of Her Majesty and save an agreement to the contrary belongs to the crown, and is exercised in Canada by the government. It is worth noting, however, that the Canadian government has waived the enforcement of copyright in relation to laws and decisions of federal bodies and courts, and many provincial governments have followed this path³⁹⁸. Furthermore, many Departments who administer Crown copyright for material they create implement a policy whereby permission is not required for certain specified uses of Crown copyright material (e.g., for personal or non-commercial purposes), unless otherwise specified in the material.

The scope of protection under the Canadian Copyright Act includes the right to reproduce the work in a material form, to publish the work, to perform the work in public, to communicate the work to the public by telecommunication, and to make an adaptation of the work³⁹⁹. The Act has contained a fair dealing exception since its initial adoption in 1921. To be exempted under the fair dealing exception, the purpose of the dealing must qualify as one of the allowable purposes under the Copyright Act, and the dealing must be fair. The Copyright Act underwent changes in 2012, which included provisions to implement the WCT, as well as new or expanded exceptions (for example, education, parody, and satire were added to the previous allowable fair dealing purposes, namely research, private study, criticism, review or news reporting) and remedies⁴⁰⁰.

A first element in the assessment of whether a dealing is "fair" is to consider whether it falls under one of the allowable purposes of the Copyright Act. In *CCH Canadian Ltd. v Law Society of Upper Canada*, a landmark case⁴⁰¹, the Court was asked to decide upon the application of the fair dealing defence for purposes of research and private study. The Court ruled that 'these allowable purposes should not be given a restrictive interpretation or this could result in the undue restriction of users' rights' (para. 54). The Court in *CCH* also stated that the allowable purposes must be given a "large and liberal interpretation", and that "research" is not limited to non-commercial or private contexts (para. 51)⁴⁰².

Secondly, the dealing must be fair. Whether a dealing is "fair" depends, in addition to the purpose of the use, on a series of criteria set out by the courts and to be assessed on a case-by-case basis, including the character and amount of the dealing, the nature of the work, the existence of any alternatives to the dealing, and the effect of the dealing on the work. A recent line of jurisprudence from the Supreme Court of Canada interpreting the criteria of the "fair dealing" defence has prompted some commentators to see in this development a subtle move towards an American-style "fair use" defence⁴⁰³.

Whether TDM is fair dealing would depend on the specific facts of each individual case to which a court would apply the conditions mentioned above. TDM would need to fall under one of the enumerated purposes and be "fair". Knowing that the Supreme Court has twice reiterated the importance of allowing fair dealing for purposes of research and private study, TDM activities could, under the appropriate factual circumstances, qualify as fair dealing under the new Canadian copyright regime.

Australia

According to the Australian Copyright Act, copyright subsists in an original literary, dramatic, musical, or artistic work published or unpublished⁴⁰⁴. Copyright gives the author of a literary, dramatic, and musical work the exclusive right to reproduce the work in a material form, to publish the work, to perform the work in public, to communicate the work to the public, and to make an adaptation of the work⁴⁰⁵. Reproduction arises when one such work is copied in any material form, where material is to be

interpreted as including any visible or non-visible form of storage, irrespective of whether the work is stored in a form that itself could be reproduced (therefore including e.g. RAM copies)⁴⁰⁶. In particular, Sec. 21(1A) establishes that a work is reproduced if it is converted into or from a digital or other electronic machine-readable form. The right of publication consists in the right of distribution and a work is published only if reproductions of it have been supplied to the public by sale or otherwise⁴⁰⁷. The Copyright Amendment Act of 2000 has introduced in Australia a general right of communication to the public, defined as the right to make available online or electronically transmit a work or subject matter⁴⁰⁸.

Fair dealing for the purpose of research or study does not constitute an infringement of the copyright in the work⁴⁰⁹. The meaning of research and study in relation to fair dealing defence has been analysed in detail by the Court of Appeal⁴¹⁰. Interestingly, Sec. 40(1A) establishes that fair dealing with a literary work (other than lecture notes) does not constitute an infringement of copyright if it is for the purpose of, or associated with, an approved course of study or research by an enrolled external student of an educational institution. Lecture notes are defined as any literary work produced for the purpose of the course of study or research by a person lecturing or teaching in or in connection with the course of study or research⁴¹¹.

In order to determine whether the dealing for the purpose of research or study was fair, Subsection 2 lists, for the purpose of illustration: the purpose and character of the dealing; the nature of the work; the possibility of obtaining the work within a reasonable time at an ordinary commercial price; the effect of the dealing on the potential market for the work; and the amount and substantiality of the part copied in relation to the whole work – in cases of only partial reproductions. A special case is introduced for reproductions, for the purpose of research or study, of all or part of a literary, dramatic or musical work contained in an article in a periodical publication. This case is considered to be fair dealing – i.e. regardless of an evaluation of the elements established by Subsection 2 – for the purpose of research or study unless another article in the publication is also reproduced for the purpose of different research or a different course of study⁴¹².

Another special case where the finding of fairness in the dealing is irrespective of the elements listed in Subsection 2, is provided by Sec. 40(5). It establishes that a reproduction for the purpose of research or study of not more than a reasonable portion (usually a 10% or a chapter) of a published work such as a literary work of at least 10 pages is taken to be a fair dealing with the work for the purpose of research or study. The provision of Subsection 5 does not apply to computer programs or electronic compilations such as databases.

Other cases of fair dealing on top of research and study include criticism or review, parody or satire, and news reporting⁴¹³. The Australian Copyright Act also provides for a number of specific exceptions to copyright (“Acts not constituting infringement of copyright”) in Divisions 3 to 5, which include a variety of cases ranging from copies for private use and copies made by educational establishments, to temporary reproductions made in the course of communications, which however are not particularly relevant for our present purpose.

Unlike the U.S. and Canada – where as seen, and for different reasons, fair use and fair dealing are considered flexible instruments – the Australian fair dealing exception has not received such a broad interpretation from the courts. The Australian Law Reform Commission received Terms of Reference for the Copyright Inquiry on 29 June 2012. The ALRC was asked to consider whether exceptions and statutory licences in the *Copyright Act 1968* are adequate and appropriate in the digital environment and whether further exceptions should be recommended. The Final Report was tabled in Parliament on 13 February 2014⁴¹⁴. The report discusses at length the comparative benefits and drawbacks of introducing a fair use defence or amending the fair dealing defence. The Report contains 30 recommendations for reform. The key recommendation is for the introduction of a fair use exception to Australian copyright law. Only the future will tell whether the Australian Parliament will go that far.

Israel

The Israel Copyright Act of 2007 grants protection to any original literary, dramatic, musical, or artistic work, if the work is fixed in any form⁴¹⁵. Sec. 5 of the Act states that ideas, procedures and methods of operation, mathematical concepts, facts or data, or news of the day are not protected by copyright, which however can subsist in their expression⁴¹⁶. The case law indicates, however, that in addition to originating in an author, a work must show some creativity to attract copyright⁴¹⁷. In particular, the Supreme Court found that a scholar deciphering fragments of old and damaged scrolls and reconstructing missing parts performed a copyright protectable activity even though his intent was to restore the text to its prior state, because such reconstruction required creativity and discretion in choosing among alternatives⁴¹⁸.

According to Sec. 11, an author enjoys the right of reproduction, publication, public performance, broadcast, making available to the public, making a derivative work, and rental. In particular, reproduction is defined as the right to make a copy of the work and includes storage of a work through electronic or any other technological means, making a three-dimensional copy of a two-dimensional work and the other way around, and making a temporary copy of a work⁴¹⁹. The right of publication, that is the right to publish a reasonable quantity of copies of the work, taking into consideration the character of the work, serves as the right of distribution⁴²⁰.

The 2007 Act shifted Israeli copyright law from a British ‘fair dealing’ framework to an American ‘fair use’ framework, accompanied by an additional list of exceptions. Traditionally, the ‘fair dealing’ defence is in principle much narrower than the US inspired ‘fair use’ defence. The main difference lies in the fact that the purposes for which the defence is admissible are enumerated exhaustively in the act⁴²¹. Fair dealing is therefore not an open norm and the interpretation of the purposes listed in article 2(1)(i) of the former Act by the Israeli courts gave rise to some tension in the years preceding the copyright reform. Since the amendment of 2007, the Israeli Copyright Act contains an open-ended fair use defence that can be invoked in a wide variety of cases and situations. Article 19 of the Copyright Act of 2007 is modelled after section 107 of the US Copyright Act of 1976 but contains an interesting feature in paragraph (c), according to which the Minister may make regulations prescribing conditions under which a use shall be deemed a fair use.

The purposes that a court should consider in order to determine whether a dealing was fair for the purpose of private study or research are listed in Sec. 19(b) and *include* the purpose and character of the use, the character of the work, the scope of the use, in both quantitative and qualitative terms, and the impact of the use on the potential market of the work. Importantly, the list is non exhaustive and courts have considered the attribution to the author or copyright owner as an important additional factor⁴²².

The amendments of 2007 were limited not only to the implementation of the fair use defence. An extensive number of additional exceptions were introduced in the Israeli Copyright Act covering a number of different uses of works, none of which are directly applicable to TDM activities. The new Israeli fair use provision has yet to be tested in a TDM case. At this time, it is impossible to predict how a judge would rule on the issue, but it is fair to say that in rendering judgment in new situations Israeli courts tend to look at American case law for inspiration.

Japan

The Japanese Copyright Act 1970⁴²³ defines its overall purpose as that of securing protection of [copyright] paying due regard to fair exploitation of cultural products and in order to contribute to the development of culture⁴²⁴. Works are defined as productions in which thoughts or sentiments/emotions are expressed in a creative way and which fall within the literary, scientific, artistic or musical domain, such as theses and lectures, figurative works of a scientific nature, and computer programs⁴²⁵. Collections of works

and databases are explicitly protected under the conditions seen above in the Section dedicated to international protection⁴²⁶.

Arts. 21 to 27 regulate the scope of protection of copyright, which include the rights of reproduction, distribution, adaptations and translation, rental and lending rights and public performance. The right of reproduction is defined by Art. 21 as the right to reproduce a work in a tangible form by means of printing, photography, copying, sound recording, visual recording or other methods. Temporary reproductions and reproductions for private use are exempted⁴²⁷. Art. 26-2 regulates the right of distribution, i.e. the right to control the exploitation of copies of works by making them available to the public by way of the assignment of the original copies.

Arts. 30 to 50 constitute Subsection 5, which is dedicated to exceptions and limitations to copyright. Of particular interest for our purpose is certainly Art. 30(4), introduced in 2012, which allows a publicly disclosed work to be used as needed for the development of technology and in experiments to test audio or visual recording devices⁴²⁸.

Another amendment, of 2009⁴²⁹, introduced, alongside other limitations an exception aimed at boosting Japan's internet economy,⁴³⁰ an exception specifically designed to permit TDM. Article 47^{septies} of the Copyright Act⁴³¹ contains an explicit provision to allow text mining, which reads:

For the purpose of information analysis ('information analysis' means to extract information, concerned with languages, sounds, images or other elements constituting such information, from many works or other much information, and to make a comparison, a classification or other statistical analysis of such information; the same shall apply hereinafter in this Article) by using a computer, it shall be permissible to make recording on a memory, or to make adaptation (including a recording of a derivative work created by such adaptation), of a work, to the extent deemed necessary. However, an exception is made of database works which are made for the use by a person who makes an information analysis.

A report issued by the subdivision on Copyright of the Council for Cultural Affairs in January 2009 presents the following examples of information analysis: (1) website information analysis and language analysis in which the use of a specific language or character string is analysed and statistically processed and (2) sound analysis and video/image analysis in which the meaning of the sound wave, video, character string, etc., comprising a certain sound, video, image, etc., is analysed. Although the types of works subject to this provision are not limited, the reverse engineering of computer programming falls outside the scope of this exception: reverse engineering cannot be regarded as "information analysis" because no statistical analysis is conducted.

The wording of the last sentence of the provision may be due to difficulties in translation. According to the AIPPI report of the Japanese Group, when the results of information analysis are presented, it is prohibited to exploit the works subject to the information analysis. The results may be presented or provided only if the results are presented or provided in the form of statistical data, etc., in which the works subject to the analysis are not exploited.

Recently, Japan has seen the introduction of new services that enable users to search and analyse users' comments on the Internet including blogs, review sites and social media. The establishment of said Article is one of the factors that have promoted the emergence of those new services⁴³².

Database related provisions were introduced in Japanese Copyright Law for the first time in 1986. The Japanese legislator considered that separate protection from that afforded to compilations under copyright law should be afforded to (electronic) databases, and decided to introduce provisions specifically

drafted for electronic databases into the Copyright Law. Based on this distinction between compilations and databases, it was thought that databases that should be protected under the new provisions were computer-searchable databases, and at the same time, because creativity of their data arrangement does not need to be protected, electronic databases are excluded from the definition of “compilations.” Art.2(1)(xter) of the Law defines the term “database” as “an aggregate of information such as articles, numerals or diagrams, which is systematically constructed so that such information can be searched for with the aid of a computer.”

The Japanese Copyright Law does not give any indication about whether the reproduction of a “portion of” a database constitutes an infringement of the database. However, it seems that because a database or compilation can enjoy copyright protection for the very reason of “the selection or arrangement of their contents” (compilations) or “the selection or systematic construction of information”(databases), whether a reproduction of a “portion” of a compilation or database is allowed or not should be judged by answering the question whether the selection or systematic construction of information was reproduced or not (for an electronic database) or whether the selection or arrangement of their contents was reproduced or not (for a non-electronic database)⁴³³.

South Korea

The Copyright Act of the Republic of Korea protects the creations of human ideas or emotions, as long as the ideas or emotions are expressed that can be perceived and the creations are original⁴³⁴. Copyright subsists in a variety of works of authorship ranging from literary works to computer program works⁴³⁵.

The Copyright Act provides for authors, in addition to moral rights, exclusive rights of reproduction, public performance, public transmission, exhibition, rental and adaption⁴³⁶. Reproduction comprises not only the act of fixation but also the act of making copies, whether permanent or temporary. The broader concept of public transmission includes, in particular, broadcasting, interactive transmission and digital audio transmission. Distribution is defined as the act of transfer or rental of originals or copies in tangible format to the general public⁴³⁷.

The Copyright Act also provides limitations and/or exceptions to authors' exclusive rights in Article 23 through 36. Relevant for this report are Article 25 (use for school education), Article 29 (quotation), Article 31 (use by libraries or archives) and Article 35-3. Article 25 permits to publish works made public in school textbooks for primary and secondary schools and to use a portion of works made public by designated educational institutions for teaching purposes under certain conditions. According to Article 28, anyone can quote works made public for news reporting, criticism, education and research, etc., provided that the quotation is within a reasonable limit and compatible with fair practices. Libraries and archives are exempted from liability for use of the books held by them under certain conditions and circumstances as referred to in Article 31. The Republic of Korea introduced a U.S. type fair use clause, i.e., Article 35-3 in 2011 when the Copyright Act was amended to comply with Korea-U.S. FTA obligations. According to the Article, it is permissible to use works for news reporting, criticism, education and research, etc., when such use does not conflict with the normal exploitation of works and does not unreasonably prejudice the legitimate interests of the author. Factors in determining the use as fair to be considered: the purpose and character of the use, including whether such use is for profit or non-profit purposes; the type and purpose of the work; the amount and substantiality of the portion used in relation to the work as a whole; and the effect of the use on the present or potential market for or value of the work.

Compilations are protected as original works of authorship, if they are creative in selection, arrangement or composition of their contents. Databases are defined as compilations of which their contents are arranged or composed in a way that anyone can individually access or search such contents.⁴³⁸

Thus original databases are protected by copyright, as they meet such conditions. Non-original databases fall under statutory subject matter of protection in accordance with the new Chapter IV introduced in 2003. The database producer who makes a considerable investment in human or material resources for the production of a database, or renewal, verification, or supplementation of their contents has the rights of reproduction, distribution, broadcasting or interactive transmission⁴³⁹. A foreign national can be a beneficiary of sui generis protection on the condition of reciprocity, if he or she is protected in accordance with a treaty to which the Republic of Korea has acceded⁴⁴⁰. The rights of the database producer are limited more broadly than copyrights, as the limitations and exceptions to copyright are applicable mutatis mutandis to the rights of the database producer on the one hand, and the use of the whole or substantial portion of the database is permissible for educational, academic or research purposes, or for reporting current events on the other hand⁴⁴¹. The term of protection is renewable 5 years⁴⁴².

Conclusion

This brief overview of the copyright and, where applicable, database legislation in force in several jurisdictions shows that the scope of protection granted to research output varies significantly from one country to another. Although the implementation of OA principles is based on contractual arrangements between authors, publishers, and universities, the framework set by the copyright regime is a determinant factor in how those arrangements take form. The ways in which copyright law defines the scope of rights and limitations and exceptions to them serve as the ground rules for licensing agreements. If they are to support OA, whether access is granted through the ‘Golden’ or the ‘Green Road’ (see section below), copyright regimes should create a favourable environment for the dissemination and re-use of publicly funded scholarly publications.

Open Access Policies

In this section we pay closer attention to OA Principles and their implementation in practice. The first subsection discusses what Open Access means, with a succinct review of the main principles and implementation options. The second subsection takes a brief look at the OA policies in two regions in the world, Europe and the United States.

Defining Open Access

Publicly funded scientific research constitutes an essential building block for further progress and innovation, one that is often seen as a collective good. For this reason, a common assumption is that, for the greater good of science and the public interest, publicly funded research should be made accessible without restriction. This principle of unfettered access also entails the freedom to use and re-use publicly funded scientific research.

The term ‘open access’ was first formally defined at a meeting in Budapest in early December 2001. Out of that meeting came the so-called Budapest Open Access Initiative⁴⁴³ and ‘open access’ was defined as the ‘free availability of scientific literature on the public internet, permitting any users to read, download, copy, distribute, print, search, or link to the full texts of these articles, crawl them for indexing, pass them as data to software, or use them for any other lawful purpose, without financial, legal, or technical barriers other than those inseparable from gaining access to the internet itself. The only constraint on reproduction and distribution, and the only role for copyright in this domain, should be to give authors control over the integrity of their work and the right to be properly acknowledged and cited.’ The Budapest Open Access Initiative was followed up some 15 months later by the Bethesda Statement⁴⁴⁴, which came out of a one-day meeting of scientists, funding agencies, librarians, scientific societies and publishers, held in April 2003. In October of the same year, the Max Planck Society in Germany convened a meeting on ‘open access to Knowledge in the Sciences and Humanities.’ This meeting widened the discussion to

include the humanities and produced the ‘Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities (October 2003)’⁴⁴⁵.

OA material can potentially include original scientific research results, like articles and monographs, as well as raw data and metadata, source materials, digital representations of pictorial and graphical materials and scholarly multimedia material. On the basis of the Berlin and Budapest statements and initiatives, the three following essential characteristics of OA emerge: free accessibility, further distribution, and proper archiving⁴⁴⁶. The Berlin Declaration gives a definition of a contribution that qualifies as open access:

1. The author(s) and right holder(s) of such contributions grant(s) to all users a free, irrevocable, worldwide, right of access to, and a license to copy, use, distribute, transmit and display the work publicly and to make and distribute derivative works, in any digital medium for any responsible purpose, subject to proper attribution of authorship (community standards, will continue to provide the mechanism for enforcement of proper attribution and responsible use of the published work, as they do now), as well as the right to make small numbers of printed copies for their personal use.

2. A complete version of the work and all supplemental materials, including a copy of the permission as stated above, in an appropriate standard electronic format is deposited (and thus published) in at least one online repository using suitable technical standards (such as the Open Archive definitions) that is supported and maintained by an academic institution, scholarly society, government agency, or other well-established organisation that seeks to enable open access, unrestricted distribution, interoperability, and long-term archiving.

Since the lack of access to complete literature can impede advances in knowledge, the main thrust of the Declaration lies in the creation of a new ‘open access’ business model for scientific publishing or, in the absence of this, of institutional repositories where all scientific and scholarly publications are to remain freely accessible. According to the Berlin Declaration, the only main constraint on reproduction and distribution of articles should be to give authors control over the integrity of their work and the right to be properly acknowledged and cited. The principles set out in the Berlin Declaration are primarily aimed at governments, universities, research institutions, funding agencies, foundations, libraries, museums, archives, learned societies and professional associations. Remarkably, neither publishers nor authors are listed in this enumeration⁴⁴⁷.

These principles are transposed into reality following two main roads: the ‘Golden road’ and the ‘Green road’ of open access publishing. The two roads to open access should not be confused, for they are complementary to each other. Under the ‘Golden road’, journals directly provide free open access to their journals and articles. As the Berlin Recommendation of 2003 implies, the ‘Golden road’ is the preferred way for the full deployment of the open access principles. This Recommendation states that “in order to implement the Berlin Declaration institutions should: implement a policy to require their researchers to deposit a copy of all their published articles in an open access repository; and encourage their researchers to publish their research articles in open access journals where a suitable journal exists and provide the support to enable that to happen”⁴⁴⁸. Open Access journals are usually licensed under one of the six core Creative Commons (CC) licences, preferably under the most liberal type, e.g. a CC- Attribution Licence 4.0⁴⁴⁹. Among the more successful open access journal and databases are the *Public Library of Science* (PLOS)⁴⁵⁰, Biomed Central⁴⁵¹ and the open access alternative offered by Springer Open Choice Publishing⁴⁵².

The ‘Green road’, by contrast, centres on self-archiving of articles published through traditional channels, where authors provide access to their own published articles by making their own e-prints free

for all. Open access self-archiving is *not* self-publishing; nor is it about online publishing without quality control (peer review); nor is it intended for writings for which the author wishes to be paid, such as books or magazine/newspaper articles. While the Green Road does meet the main OA requirements, namely free access, possibility to copy, use, distribute the work and permanent archiving, the fact that the publications are first published through traditional channels means that authors only retain limited rights on their publication/data. Self-archived articles are usually accompanied by the text of a licence telling users what they can and cannot do with the article. As a result, it is most often unclear whether self-archived articles can be distributed, re-used or modified. By contrast, releasing research results along the Golden Road ensures a better access, clearer re-use possibilities, visibility and ‘fundability’ of research output on the internet⁴⁵³.

European Open Access Policy

On July 17, 2012, the European Commission published its Communication to the European Parliament and the Council entitled ‘Towards better access to scientific information: Boosting the benefits of public investments in research’⁴⁵⁴. As the Commission observes, ‘discussions of the scientific dissemination system have traditionally focused on access to scientific publications — journals and monographs. However, it is becoming increasingly important to improve access to research data (experimental results, observations and computer-generated information), which form the basis for the quantitative analysis underpinning many scientific publications’⁴⁵⁵. The Commission emphasises that through a more complete and wider access to scientific publications and data, innovation will accelerate and researchers will collaborate more so that duplication of effort will be avoided. Moreover, open research data will better enable subsequent researchers to build on previous research results.

The Communication marks an official new step on the road to open access to publicly funded research results in science and the humanities in Europe. Scientific publications are no longer the only elements of an open access policy: research results upon which publications are based must now also be made available to the public. To implement this policy, the European Commission set up a pilot initiative on Open Access (OA) to peer reviewed research articles in its Seventh Research Framework Programme (FP7), otherwise known as the OpenAire project, to ensure that the results of the research it funds are disseminated as widely and effectively as possible to guarantee maximum exploitation and impact in the world of researchers and beyond⁴⁵⁶. OpenAire pursues three objectives: to build support structures; to establish and operate an electronic infrastructure; and to manipulate research datasets. All of this under open access conditions.

Following on from this, the Commission agreed that open access to scientific publications should be a general principle of the current Horizon 2020 research framework programme. In the model grant agreement for Horizon 2020, the Commission states that, in addition to scientific publications, the beneficiaries must deposit the data and associated metadata that are needed to validate the results presented in scientific publications in a research data repository as soon as possible. The beneficiaries must also take measures to make it possible for third parties to access, mine, exploit, reproduce and disseminate the data, free of charge for any user. In all cases, the Commission also encourages authors to retain their copyright and grant adequate licences to publishers.

The European Open Access Policy is not binding on the Member States who are free to adopt the policy that best suits the needs of their own scientific community. This leads to a mosaic of open access policies across Europe, ranging from the mandatory ‘Golden Road’ for publications and data put in place by the Research Councils of the United Kingdom (RCUK), to the ‘Green Road’ for publications and data of the Netherlands, to the ‘Green Road’ for publications in Germany and to other more equivocal policies in a number of Member States. In the recent years, the national research councils of the UK and the Netherlands have issued policy statements according to which research grants will be awarded only

provided that the applicants commit to publishing their results, both publications and data, under OA conditions.

Open Access in the United States

Since 2008, all publications that arise from National Institutes of Health (NIH) funds have to comply with the NIH Public Access Policy. The policy requires the final peer-reviewed paper to be deposited in PubMed Central, NIH's digital full-text archive, upon acceptance for publication, with an indication of when, within a period of 12 months (the so-called embargo period), the paper will become accessible to the general public⁴⁵⁷. More recently, thanks to a US government directive issued by the Office of Science and Technology Policy [Public Access Directive], all federal agencies with more than USD 100 million in research and development expenditure are required to develop plans to make the published results of federally funded research freely available to the public within one year of publication⁴⁵⁸.

However, the NIH Public Access Policy does not specify under which conditions users may reuse the publications and data made available on the website. This means that PubMed Central users can only reuse the works in accordance with the copyright law of their own jurisdiction⁴⁵⁹. In fact, while it may be possible to reproduce the paper for private or research purposes (make copies), it may not be possible to redistribute the paper (post it on one's own website) nor to modify the paper, outside what is allowed by fair use or other exceptions or limitations to copyright law. All these rights remain within the author's domain or, more often, within the publisher's. The Directive specifically calls for agencies to implement measures to prevent the unauthorised mass redistribution of scholarly publications⁴⁶⁰. Consequently, users enjoy public access, but not open access⁴⁶¹.

In some cases, data mining or bulk downloading is prohibited and enforced purely on the basis of contractual terms. For example, the NIH PubMed Central Public Access Policy prohibits the use of crawlers or systematically downloading articles that are individually available for public access on their repositories, referring to copyright restrictions⁴⁶². However, in May of 2013, President Obama issued a Policy Memorandum, "Open Data Policy – Managing Information as an Asset". The Policy Memorandum requires U.S. federal agencies to manage information created, collected, processed, disseminated, or disposed of, by or for the United States Federal Government as an asset throughout its life cycle to promote openness and interoperability, and properly safeguard systems and information. Under the Policy Memorandum, federal agencies are required to make data publicly available and fully reusable to the widest range of users for the widest range of purposes permitted by law – subject only to privacy, confidentiality, security, or other valid restrictions. Data is to be provided in convenient, modifiable, machine-readable, non-proprietary, publicly available, open formats that permit retrieval, downloading, indexing, searching and reuse. Data must be described fully so that consumers have sufficient information to understand and make effective use of federal data. Primary data is to be published in its complete form in a timely manner, and agencies must designate points of contact to assist with data use and to respond to complaints about adherence to federal open data requirements⁴⁶³.

Unresolved Issues

OA principles entail more than just granting access to research data with limited or no restrictions. The core of OA principles aspire to make research data available for any type of re-use by any user. What are the implications of applying OA principles to research results that are not funded entirely by public money? How does the funding of research by public/private partnerships affect ownership of the data and the licensing conditions applied? To what extent is copyright law implicated by 'Text and Data Mining' (TDM) and what is the role of OA principles?

Public/Private Partnerships

Ownership issues may arise in cases of research outputs from public/private partnerships. The funding of a research project through external sources, whether public or private, usually leads to the application of different rules of ownership. For copyrightable research outputs, the parties may elect to contract around copyright law. A contract typically involves at least three parties: the author, the university or other research institutions and the sponsoring or commissioning party. Depending on the law, the internal policy of the institution or on the bargaining position of the respective parties, the copyright ownership may be transferred either to the university or to the external entity.

The issue of rights ownership is very important in the context of public/private partnerships because it can greatly influence the manner in which research output will be published: either through traditional channels with a full reservation of copyright, or through 'Green' or 'Golden Road' OA.

Although the possibility exists under the Creative Commons licensing system to restrict use for commercial purposes, the distinction between commercial and non-commercial use in the Creative Commons licences raises pressing questions not only in the scientific publishing sector, but also in several other sectors of the copyright industry, because it leaves much room for interpretation. For example, for the purposes of an OA contribution, would the 'responsible purpose' referred to in the Berlin Declaration include a commercial use? Would a pharmaceutical company's distribution among thousands of physicians of an OA scientific article promoting its product fall under such a 'responsible purpose'?

Text and data mining

TDM is an important technique used in science and other disciplines for analysing and extracting new insights and knowledge from the exponentially increasing store of digital data ('Big Data'), which is certain to become more important as researchers acquire the skills and the technology to address and investigate datasets of increasing size, complexity and diversity in all media: text, numbers, images, audio files and in any other form.

The uncertainty in current legal frameworks regarding the scope of protection of works and databases is bound to create obstacles to TDM activities. A system resting solely on licensing agreements might be insufficient to allow TDM to take place in all instances where it would be socially desirable. Firstly, because only a portion of the databases that are interesting for TDM research may be offered as part of publishers' subscription agreements and an even smaller portion could be available under OA conditions. If TDM were allowed to take place without restriction, transaction costs would be lower. Otherwise, some databases may remain out of reach of researchers. Moreover, transaction costs would rise if researchers had to reconcile the terms and conditions of non-standard or non-interoperable licenses.

Private actors are not subject to any obligation to open up or share their data with third parties. Even in situations where such data does not enjoy any special copyright or database protection, restrictions on the (re-)use may flow from contractual requirements (in terms and conditions) set by the holder of the data or from the application of technological protection measures.

It is against this background that the United Kingdom has recently introduced a specific exception in the Copyright, Designs and Patent Act to allow TDM activities for non-commercial research to take place without the rights holder's prior authorisation under the conditions stated in the law. Together with the Japanese provision, the UK provision is a step towards facilitating scientific research that may in time be followed by other jurisdictions.

NOTES

284 This chapter is based on a background paper prepared by L. Guibault and T. Margoni, DSTI/STP(2014)37. The background paper, in turn, was based in part on L. Guibault and A. Wiebe (ed.), *Safe to be open : Study on the protection of research data and recommendations for access and usage*, Göttingen: Universitätsverlag Göttingen, 2013.

285 See for example the TIP interim report draft of the OECD Open Science project: DSTI-STP-TIP(2014)9-ENG, in particular the section "The impacts of Open Science: an Overview".

286 OECD Principles And Guidelines For Access To Research Data From Public Funding, Paris, OECD, 2007; Communication of the European Commission on scientific information in the digital age: access, dissemination and preservation, COM(2007)56, Brussels, 14.2.2007.

287 Id.

288 See L. Guibault, 'Owning the right to open up access to scientific publication', in L. Guibault, C. Angelopoulos, *Open Content Licensing: From Theory to Practice*, Amsterdam University Press, 2011, p. 137-167.

289 See Art. 2(1) of the Berne Convention for the Protection of Literary and Artistic Works, signed in Berne, Switzerland on 9 September 1886. As of 2014 the Berne Convention has been ratified by 167 countries, including all OECD members.

290 See Art. 2 "Scope of copyright protection" of the World Intellectual Property Organization (WIPO) Copyright Treaty, adopted in Geneva on December 20, 1996; and Art. 9(2) "Relation to the Berne Convention" of the Agreement on Trade-Related Aspects of Intellectual Property Rights (TRIPS), signed in Marrakesh on 15 April 1994, as Annex 1C of the Marrakesh Agreement Establishing the World Trade Organization. The two provisions are virtually identical, with the only difference being the presence of the auxiliary "shall" in the TRIPS definition.

291 See e.g. the U.S Supreme Court decision in *Harper & Row v. Nation Enterprises*, 471 U.S. 539 (1985) "copyright's idea/ expression dichotomy strikes a definitional balance between the First Amendment and the Copyright Act by permitting free communication of facts while still protecting an author's expression". A similar statement is present in art. 1(2) of the European Software Directive "protection ... shall apply to the expression ... of a computer program. Ideas and principles ... are not protected".

292 See Arts. 8 and seq. Berne, Arts. 6 and seq. WCT, and Arts. 11 and seq. TRIPS.

293 See Art. 9(1) Berne. See also P. Goldstein, B. Hugenholtz, *International Copyright – Principles, law, and practice*, 3rd ed., Oxford University Press, 2013, at 307.

294 See the WIPO Guide to the Berne Convention, 1978, at 54.

295 See S. Ricketson and J. Ginsburg, *International copyright and neighbouring rights*, Oxford University Press, 2006, at 11.26.

296 Although a right to distribute copies can be identified with regard to specific cases such as cinematographic adaptations and reproductions of literary works (Art. 14(1), *droit de suite* of original works of art and written manuscripts (Art. 14ter), and the seizure of infringing copies of works (Art. 16); See Ricketson, cit., at 11.45.

297 A similar provision is contained in Arts 8 and 12 of the WIPO Performance and Phonograms Treaty.

298 See Goldstein, cit., at 311.

299 See De Wolf Study, at 61.

- 300 Among the jurisdictions already offering a general right to communicate a work to the public are Germany and France, while recognising specific rights Canada, the U.K., and Japan can be listed; See Goldstein, cit., at 324.
- 301 See Goldstein, cit., at 325.
- 302 See German Supreme Court (BGH), 17 July 2003, Case I ZR 259/00 (Paperboy); ECJ Case C-466/12 of 13 February 2014, *Svensson et al. v. Retriever Sverige* [Svensson]; M. Senftleben, in Dreier & Hugenholtz, *Concise European Copyright Law*, 2006, WCT note 6(a); Goldstein, cit., 336.
- 303 Art. 2(2) Berne.
- 304 See Art. 2(4) Berne.
- 305 See Art. 2(7) Berne.
- 306 See Art. 3(1) Berne.
- 307 See Art. 2(8) Berne.
- 308 See Guibault L., *Copyright limitations and contracts*, London 2002, at 28.
- 309 See Art. 10(1) Berne.
- 310 See Art. 10(2) Berne.
- 311 See Art. 10bis(1) Berne.
- 312 See Art. 10bis(2) Berne.
- 313 See Art. 9(2) Berne Convention. This article is the first appearance in an international agreement of the so called Three-step Test. Art. 9(2) was introduced in the 1967 Stockholm revision of the Convention and in its Berne formulation only applies to the right of reproduction. Successively, it was introduced in all the other major international agreements. See in general Senftleben M., *Copyright, limitations and the Three-step test*, London, 2004; Gervais D., *Towards a new core international copyright norm: the reverse Three-step test*, *Marquette Intellectual Property Law Review*, Vol. 9, 2005, p.1; Ricketson S., *WIPO study on limitations and exceptions of copyright and related rights in the digital environment*, WIPO, 2003.
- 314 See Art. 2(5) Berne Convention.
- 315 See Art. 10(2) TRIPS Agreements.
- 316 Directive 2009/24/EC of the European Parliament and of the Council of 23 April 2009 on the legal protection of computer programs (Codified version) (Text with EEA relevance) *OJ L 111*, 5.5.2009, p. 16–22, art. 1(3).
- 317 Directive 96/9 of the European Parliament and of the Council of 11 March 1996 on the legal protection of databases, *OJ L 77*, 27.3.1996, p. 20–28, art. 3(1).
- 318 Directive 2006/116/EC of the European Parliament and of the Council of 12 December 2006 on the term of protection of copyright and certain related rights (codified version), *OJ L 372*, 27.12.2006, p. 12–18, art. 6.
- 319 M. van Eechoud, *Along the Road to Uniformity - Diverse Readings of the Court of Justice Judgments on Copyright Work*, *JIPITEC : Journal of Intellectual Property, Information Technology and E-Commerce Law*, 2012-1, p. 60-80; *Infopaq International A/S v Danske Dagblades Forening*, Case C-5/08, Judgment of the Court, 16 July 2009; *Bezpečnostní softwarová asociace v. Ministerstvo kultury*, C-393/09, Judgment of the Court (Third Chamber) of 22 December 2010;

- 320 See Football Dataco. See also T.-E. Synodinou, The Foundations of the Concept of Work in European Copyright Law, in: T.-E. Synodinou (ed.), *Codification of European Copyright Law – Challenges and Perspectives*, The Hague, Kluwer Law International, 2012, pp. 93-113, p. 101.
- 321 See Art. 2 Information Society Directive; See also L. Guibault, G. Westkamp et al., *Study on the implementation and effect in member state's laws of directive 2001/29/EC on the harmonisation of certain aspects of copyright and related rights in the information society*, Institute for Information law, University of Amsterdam, 2007, at 35-36.
- 322 See Art. 5(1) InfoSoc Directive; J. Triaille (ed.), *Study on the application of Directive 2001/29/EC on copyright and related rights in the information society*, De Wolf & Partners, 2013 [De Wolf Study], at 111.
- 323 L. Guibault, Why Cherry Picking Never Leads to Harmonisation: The Case of the Limitations on Copyright under Directive 2001/29/EC, *JIPITEC*, 2010-2, p. 55-66.
- 324 De Wolf Study, p. 403.
- 325 European Commission, Green Paper on Copyright in the Knowledge Economy, Brussels, 16 July 2008, COM(2008) 466/3, p. 6.
- 326 European Commission, Green Paper on Copyright in the Knowledge Economy, Brussels, 16 July 2008, COM(2008) 466/3, p. 6.
- 327 Brussels, 19.10.2009, COM(2009) 532 final.
- 328 European Commission, Report on the responses to the Public Consultation on the Review of the EU Copyright Rules, DG Internal Market and Services, Brussels, July 2014, p. 59-60.
- 329 Directmedia Publishing GmbH v Albert-Ludwigs Universität Freiburg, C-304/07, [2009] 1 C.M.L.R. 7.; Apis – Hristovich EOOD v Lakorda AD, C-545/07 [2009] ECRI-1627.
- 330 Directmedia decision, para 36.
- 331 BHB decision, para 61.
- 332 Innoweb B.V. v Wegener ICT Media B.V. and Wegener Mediaventions B.V., C-202/12, Decision of the Court of Justice, 19 December 2013.
- 333 BHB decision, para. 70.
- 334 J.H. Reichman and R. L. Okediji, When Copyright Law and Science Collide: Empowering Digitally Integrated Research Methods on a Global Scale, 2012, *Minn. L. Rev.* Vol. 96, p. 1362, 1451.
- 335 Among the countries outside the European Union that recognise some protection on non-original databases are South-Korea and Mexico
- 336 L. Guibault and A. Wiebe (eds.), *Safe to be open - Study on the protection of research data and recommendations for access and usage*, Göttingen University Press, Göttingen, 2013, p. 33-34.
- 337 See Nauta Dutilh, *The implementation and application of Directive 96/9/EC on the legal protection of databases*, Brussels, 2001, Contract ETD/2001/B5-3001/E/72, available at: http://ec.europa.eu/internal_market/copyright/prot-databases/index_en.htm
- 338 De Wolf Study, p. 365. See also: A. Beunen, *Protection for Databases – The European Database Directive and its effects in the Netherlands, France and the United Kingdom*, Nijmegen, Wolf Legal Publishers, 2007, p. 219.

- 339 See Recital 34 offers some guidance: ‘Whereas, nevertheless, once the rightholder has chosen to make available a copy of the database to a user, whether by an online service or by other means of distribution, that lawful user must be able to access and use the database for the purposes and in the way set out in the agreement with the rightholder, even if such access and use necessitate performance of otherwise restricted acts’.
- 340 See M.M.M. van Eechoud et al., *Harmonising European Copyright Law – The Challenges of Better Law Making*, Alphen aan den Rijn, Kluwer Law International, 2009, p.
- 341 Code de la propriété intellectuelle created by Loi n° 92-597 du 1er Juillet 1992, as amended.
- 342 See Art. L111-1 IPC.
- 343 See Art. L112-2 and L1112-3 IPC.
- 344 See Article L122-2 IPC; A. Lucas, H. Lucas, et al, *Traité de la propriété littéraire et artistique*, 4th Ed., Paris, 2012, at 247.
- 345 See Art. L122-3; Lucas et al., *op. cit.*, at 253.
- 346 See Article L122-5(3) IPC.
- 347 See Beunen, *op. cit.*, p. 144. See also TGI Paris 25 April 2003; TGI Strasbourg 22 July 2003.
- 348 See article 10(2) of the Berne Convention : “It shall be a matter for legislation in the countries of the Union, and for special agreements existing or to be concluded between them, to permit the utilisation, to the extent justified by the purpose, of literary or artistic works by way of illustration in publications, broadcasts or sound or visual recordings for teaching, provided such utilisation is compatible with fair practice” and article 5(3)a) of the Directive 2001/29/EC “(a) use for the sole purpose of illustration for teaching or scientific research, as long as the source, including the author’s name, is indicated, unless this turns out to be impossible and to the extent justified by the non-commercial purpose to be achieved”.
- 349 See Arts. 1 and 2 German Copyright Act (Act of 9 September 1965 “Gesetz über Urheberrecht und verwandte Schutzrechte, UrhG).
- 350 Cf. A. Nordemann, in Fromm/Nordemann (eds), *Urheberrecht*, 10th edition, Stuttgart, Kohlhammer, 2008, § 2, para. 11.
- 351 See G. Schulze, in Dreier/Schulze (eds), *UrhG*, 3rd edition, Munich, Beck, 2008, § 2, para. 37.
- 352 See Art. 4 German Copyright Act.
- 353 See Cf. G. Schulze, in Dreier/Schulze (eds), *UrhG*, 3rd edition, Munich, Beck, 2008, § 4, para. 4.
- 354 See BGH, MMR 2007, 589 – Gedichttitelliste I.
- 355 See G. Schulze, in Dreier/Schulze (eds), *UrhG*, 3rd edition, Munich, Beck, 2008, § 7, para. 2.
- 356 See Art. 38(4) German Copyright Act.
- 357 See R. Hilty et al, *Zum Referentenentwurf eines Gesetzes zur Einführung einer Regelung zur Nutzung verwaister Werke und weiterer Änderungen des Urheberrechtsgesetzes sowie des Urheberrechts-wahrnehmungsgesetz, Stellungnahme des Max-Planck-Instituts für Immaterialgüter- und Wettbewerbsrecht zur Anfrage des Bundesministeriums der Justiz vom 20. Februar 2013*, available at http://www.ip.mpg.de/files/pdf2/Stellungnahme-BMJ-UrhG_2013-3-15-def1.pdf; V. Moscon, *Open Access to Scientific Articles: Comparing Italian with German law*, available at <http://kluwercopyrightblog.com/2013/12/03/open-access-to-scientific-articles-comparing-italian-with-german-law>.

- 358 See T. Dreier, in Dreier/Schulze (eds), *UrhG*, 3rd edition, Munich, Beck, 2008., §§ 87a et seq., para. 8.
- 359 See G. Westkamp, *Der Schutz von Datenbanken und Informationssammlungen im britischen und deutschen Recht*, Munich 2003, p. 406.
- 360 See BGH, GRUR 2005, 857, 858 – HIT BILANZ.
- 361 See BGH, MMR 2011, 676 – Zweite Zahnarztmeinung II, at 20 et seq.
- 362 See BGH, MMR 2010, 41 – Gedichttitelliste III.
- 363 See BGH, MMR 2011, 676 – Zweite Zahnarztmeinung II, at 14.
- 364 Dutch Supreme Court, January 1991, NJ 1991, 608, m.nt. DWFV, AA 1992, 31 m.nt. HCJ, IER 1991, 96 m.nt. FWG, AMI 1991, 177 m.nt. JHS, CR 1991, 84 m.nt. Hugenholtz (Van Dale/Romme I).
- 365 See M. van Eechoud, P. Bernt Hugenholtz, S. van Gompel, L. Guibault, and B. van der Sloot, ‘Dutch report’, ALAI Study Days, Dublin, June 2011, p. 1.
- 366 Supreme Court, 14 June 1968, NJ 1968, 276 (*DNB v C&A, Telegraaf – ‘Bankbiljet’*); Supreme Court, 29 May 1987, NJ 1987, 1003 (*Struycken & Unger v Riet – ‘Beatrix Zegel’*); Court of Appeal Arnhem, 24 June 2008, Computerrecht 2008, 138 (*Openbareverkoop.nl v Internet notarissen*) with comment O. Volgenant.
- 367 Case where substantial investment was found to exist: Court of Appeal of Arnhem, (*Openbareverkoop v Internetnotarissen*), IEPT 2008-06-24; cases where no substantial investment was found to exist: Court of Appeal of Arnhem, 4 July 2006, (*NVM v Zoekalhuizen*), IEPT 2006-06-04; Court of Appeal of The Hague, (*KNMP v ID/Farma*) IEPT 2008-07-10; Court of Appeal of Amsterdam, (*PR Aviation v Ryanair*), IEPT 2012-03-13.
- 368 A. Beunen, op.cit., p. 150.
- 369 J.H. Spoor, D.W.F. Verkade and D.J.G. Visser, *Auteursrecht*, 3rd edition, Deventer, Kluwer, 2005, p. 625.
- 370 ABRvS 29 April 2009, n 07/786, AMI 2009-6 (College B&W Amsterdam/Landmark; with annotation from M. Van Eechoud).
- 371 See L. Bently and B. Sherman, *Intellectual Property Law*, 3rd ed., 2009, Oxford, at 58 and seq.
- 372 See CDPA 1988 Sec. 3.
- 373 See *University of London Press v. University Tutorial Press* [1916] 2 ch 601; L. Bently, cit. at 61.
- 374 See L. Bently, cit. At 136.
- 375 See L. Bently, cit., at 66.
- 376 See Sec. 14 Copyright and Rights in Databases Regulations 1997.
- 377 See Sec. 20 Copyright and Rights in Databases Regulations 1997.
- 378 The expression “literary, dramatic, musical or artistic” has been deleted by the Copyright and Rights in Performances (Research, Education, Libraries and Archives) Regulations 2014, in force since 1st June 2014.
- 379 See *Ashdown v. Telegraph Group* [2002] ch 149, 172, at 71.
- 380 See L. Bently, cit., at 206 et seq.

- 381 These yet to be implemented provisions should refer to Personal copies for private use, Quotation and Parody, scheduled entry into force on October 1st 2014.
- 382 See the official opinion of the UK Government in the document titled “Technical Review of Draft Legislation on Copyright Exceptions: Government Response”, at 13 (“The Government’s view is that this existing exception will permit the extraction of whole works if required for text and data mining through the provision for “fair dealing with a substantial part”).
- 383 See Art. I, Sec. 8, Clause 8 of the United States Constitution.
- 384 See Copyright Act 1976 Pub. L. 94-553 (Oct. 19, 1976), in force since 1978.
- 385 See Sec. 102 Copyright Act 1976.
- 386 See Feist Publications, Inc., v. Rural Telephone Service Co., 499 U.S. 340 (1991).
- 387 “... but the United States Government is not precluded from receiving and holding copyrights transferred to it by assignment, bequest, or otherwise”.
- 388 See e.g. “Copyright and Other Rights Pertaining to U.S. Government Works”, available at <http://www.usa.gov/copyright.shtml>
- 389 See Sec. 106 Copyright Act 1976. See M. Nimmer and D. Nimmer, Nimmer on Copyright, 2013 Rev. Ed., at ch. 8 Vol. 2.
- 390 Text of the decision available from: <http://docs.justia.com/cases/federal/district-courts/new-york/nysdce/1:2011cv06351/384619/156>. It should be noted that The Authors Guild is appealing the decision in Authors Guild v. HathiTrust to the Second Circuit and announced it would also appeal the ruling in Authors Guild v. Google.
- 391 Text of the decision available from: <http://www.nysd.uscourts.gov/cases/show.php?db=special&id=115>
- 392 See http://scholar.google.com/scholar_case?case=7845449963545508939&hl=en&as_sdt=6&as_vis=1&oi=scholar
- 393 The first fair use factor considers the "purpose and character of the use, including whether such use is of a commercial nature or is for nonprofit educational purposes.”
- 394 See Canadian Copyright Act (R.S.C., 1985, c. C-42).
- 395 See Canadian Copyright Act (S.C., 1921, C-24). Before this time the Copyright Act of the United Kingdom applied.
- 396 See Sec. 3.1 Canadian Copyright Act.
- 397 CCH Canadian Ltd. v. Law Society of Upper Canada, 2004 SCC 13, [2004] 1 S.C.R. 339, at para. 16. See D. Gervais & E. Judge, Intellectual Property: The law in Canada, Toronto, 2005, at 21; see also M. Perry, T. Margoni, From music tracks to Google maps: Who own computer-generated works?, in CLSR 26 (2010), 621 – 629.
- 398 See Reproduction of Federal Law Order (SI/97-5); Gervais, cit., at 44.
- 399 Canadian Copyright Act, Statute of Canada, ch. C-42, art. 3.
- 400 “Education, parody or satire” added by the Canadian Modernization Act of 2012 (formerly Bill C-11).
- 401 CCH Canadian Ltd. v Law Society of Upper Canada, 2004 SCC 13 at para 48, [2004] 1 SCR 339 [CCH] <http://scc.lexum.org/decisia-scc-csc/scc-csc/scc-csc/en/item/2125/index.do>.

- 402 See CCH, cit., at 84; Gervais, cit., at 81. See also *Alberta (Education) v. Canadian Copyright Licensing Agency (Access Copyright)*, 2012 SCC 37.
- 403 Michael Geist, *Fairness Found: How Canada Quietly Shifted from Fair Dealing to Fair Use*, in M. Geist (ed.), *The Copyright Pentology*, Ottawa, University of Ottawa Press, 2013, pp. 157-186.
- 404 See Copyright Act 1968, as amended.
- 405 See Secs. 31 and 32 Copyright Act 1968.
- 406 See B. Sherman, *Australia*, in P. Geller, L. Bently, *International copyright law and practice*, Lexis Nexis, 2013, at 8[1][b].
- 407 See Sec. 29(1); See also Sherman, cit., at 8[1][b][i][B].
- 408 See Sec. 10(1) as amended by the Copyright Amendment (Digital Agenda) Act of 2000.
- 409 See Sec. 40(1).
- 410 See *De Grais v. Neville Jeffress Pidler* [1990] FCA 352, at 22 et seq. (excluding research or study in a case involving a method of retrieval of material published in newspapers).
- 411 See Sec. 40(1B).
- 412 See Sec. 40(3) and (4).
- 413 See Secs. 41, 41(A), and 42. Sec. 41(A) regarding “parody or satire” was introduced by the Copyright Amendment Act of 2006.
- 414 See: Australian Law Reform Commission, *Copyright and the Digital Economy*, 2014.
- 415 See Copyright Act 2007, Sec. 1.
- 416 See Copyright Act 2007, Sec. 1; See also M. Birnhack, N. Wilkof, J. Weisman, in Geller & Bently, cit., at 2[1][b].
- 417 See Birnhack et al, in Geller & Bently, cit., at 2[1][b][ii].
- 418 See *Eisenman v. Qimron*, C.A. 2790/93, 2811/93, 54(3) P.D. 817 [2000], cited i.e. by D. Nimmer, *Copyright in the Dead Sea Scrolls – Authorship and Originality*, in 38 Hous. L. Rev. 1 (2001); See also N. Elkin-Koren, *Of Scientific claims and proprietary rights: Lessons from the Dead Sea Scrolls case*, in 38 Hous. L. Rev. 445 (2001).
- 419 See Sec. 11(1), Copyright Act.
- 420 See Birnhack et al, in Geller & Bently, cit., at 8[1][b][iii].
- 421 Meera Nair, ‘Canada and Israel – Cultivating Fairness of Use’, PIJIP Research Paper, No. 2012-04 American University, Washington College of Law.
- 422 See Birnhack et al, in Geller & Bently, cit., at 8[2][a].
- 423 See Copyright Act Law N. 48 of 1970.
- 424 See Art. 1 Copyright Act.
- 425 See Art. 2(1)(i) and 10(1) Copyright Act.

- 426 See Art. 12 Copyright Act; See also T. Doi, in Geller & Bently, cit., at 2[3][b].
- 427 See T. Doi, in Geller & Bently, cit., at 8[1][b][1].
- 428 See Art. 30(4) introduced by Law No. 43 2012 cited by T. Doi, in Geller & Bently, cit., at 1[2][c].
- 429 See Law No. 53 of 2009.
- 430 Yoshiyuki Tamura, Rethinking Copyright Institution for the Digital Age, 1 W.I.P.O.J. 63-74 (2009)
- 431 Japan Copyright Act: <http://www.cric.or.jp/english/clj/cl2.html>
- 432 Kei Iida, Sayuri Imako, Yasutaka Iwamoto, Ong Poh Chuan, Hirohito Katsunuma, Kei Konishi, Junko Kobayashi, Yasuhiko Takada, Takashi Nakazaki, Question Q216B Exceptions to Copyright protection and the permitted Uses of Copyright works in the hi-tech and digital sectors AIPPI National Group: Japanese Group, p. 9.
- 433 Japanese Group, Database protection at national and international level, AIPPI, Report Q182, Geneva, 2004.
- 434 See Copyright Act of the Republic of Korea as substantially revised by Law No. 8101, December 28, 2006, and subsequently amended. See Art. 2 Item 1.
- 435 See Art. 4.
- 436 See Arts. 16 to 22 Copyright Act.
- 437 See J. Yang, C. Shin, in Geller & Bently, cit., at 8[1][b].
- 438 See Art. 2 Items 18 and 19.
- 439 See Art. 2 Item 20 and Art. 93.
- 440 See Art. 91.
- 441 See Art. 94.
- 442 See Art. 95.
- 443 Available at: <http://www.opensocietyfoundations.org/openaccess/>.
- 444 Available at: <http://www.earlham.edu/~peters/fos/bethesda.htm/>.
- 445 Available at: <http://oa.mpg.de/lang/en-uk/berlin-prozess/berliner-erklarung/>.
- 446 Open Society Institute, *Open Access Publishing and Scholarly Societies – A Guide*, New York, OSI, 2005, p. 6; see also: P. Suber, *Open Access*, MIT Press, 2012.
- 447 <http://openaccess.mpg.de/Berlin-Declaration>
- 448 RECOMMENDATION IN ORDER TO MOVE FORWARD, adopted by the delegates of the "Berlin 3 open access" conference (Feb 28th - Mar 1st, 2005, University of Southampton, UK).
- 449 See: <http://creativecommons.org>
- 450 See: <http://www.plos.org/about/openaccess.html>.
- 451 See: <http://www.biomedcentral.com/info/about/copyright>.

452 See: <http://www.springer.com/sgw/cda/frontpage/0,11855,5-40359-12-161193-0,00.html>

453 L. Guibault, Owing the Right to Open Up Access to Scientific Publications, in: L. Guibault and C. Angelopoulos (ed.), *Open Content Licences: From Theory to Practice*, Amsterdam, Amsterdam University Press, 2011, pp. 137-167, 157.

454 Brussels, 17.7.2012 COM(2012) 401 final.

455 Id., p. 3.s

456 <https://www.openaire.eu/en/support/faq>

457 “The Director of the National Institutes of Health shall require that all investigators funded by the NIH submit or have submitted for them to the National Library of Medicine’s PubMed Central an electronic version of their final, peer-reviewed manuscripts upon acceptance for publication, to be made publicly available no later than 12 months after the official date of publication: *Provided*, That the NIH shall implement the public access policy in a manner consistent with copyright law”, see Division G, Title II, Section 218 of PL 110-161 (Consolidated Appropriations Act, 2008), as confirmed by Division F, Section 217 of PL 111-8 (Omnibus Appropriations Act, 2009); for references see <http://publicaccess.nih.gov/policy.htm> (last accessed 06/2013).

458 See <http://www.whitehouse.gov/blog/2013/02/22/expanding-public-access-results-federally-funded-research> with direct links to the Directive.

459 Interestingly the PubMed Central copyright notice prohibits bulk downloading of papers for copyright reasons: “Bulk downloading of articles from the main PMC web site, in any way, is prohibited because of copyright restrictions”, available at <http://www.ncbi.nlm.nih.gov/pmc/about/copyright>

460 See Public Access Directive, sec. 3.

461 Indeed, PubMed Central offers a specific Open Access subset: <http://www.ncbi.nlm.nih.gov/pmc/tools/openftlist>

462 “Crawlers and other automated processes may NOT be used to systematically retrieve batches of articles from the PMC web site. Bulk downloading of articles from the main PMC web site, in any way, is prohibited because of copyright restrictions. PMC has two auxiliary services that may be used for automated retrieval and downloading of a special subset of articles from the PMC archive. These two services, the PMC OAI service and the PMC FTP service, are the only services that may be used for automated downloading of articles in PMC. See the PMC Open Access Subset for information about which articles are included in this special subset, and for links to the PMC OAI and FTP services. Do not use any other automated processes for bulk downloading, even if you are only retrieving articles from the PMC Open Access Subset. Articles that are available through the PMC OAI and FTP services are still protected by copyright but are distributed under a Creative Commons or similar licence that generally allows more liberal use than a traditional copyrighted work. Please refer to the licence statement in each article for specific terms of use. The licence terms are not identical for all the articles”, <http://www.ncbi.nlm.nih.gov/pmc/about/copyright>.

463 See: <https://www.whitehouse.gov/sites/default/files/omb/memoranda/2013/m-13-13.pdf>

REFERENCES

- Bargheer, M., S. Bellem and B. Schmidt, 'Kapitel 1: open access und Institutional Repositories – Rechtliche Rahmenbedingungen', in G. Spindler (ed.), *Rechtliche Rahmenbedingungen von open access-Publikationen*, Göttingen, Universitätsverlag Göttingen, 2006, Göttinger Schriften zur Internetforschung.
- Bently, L. and B. Sherman, *Intellectual Property Law*, 3rd ed., Oxford, 2009.
- Beunen, A., *Protection for databases*, Nijmegen: Wolf 2007.
- Geist, M., 'Fairness Found: How Canada Quietly Shifted from Fair Dealing to Fair Use', in M. Geist (ed.), *The Copyright Pentalogy*, Ottawa, University of Ottawa Press, 2013.
- Gervais, D., 'Towards a new core international copyright norm: the reverse Three-step test', 9 *Marquette Intellectual Property Law Review* (2005).
- Gervais, D. and E. Judge, *Intellectual Property: The law in Canada*, Scarborough: Thomson Carswell, 2005.
- Dinwoodie, G. and R. Dreyfuss, 'Patenting science: protecting the domain of accessible knowledge', in L. Guibault and P.B. Hugenholtz, *The Future of the Public Domain – Identifying the Commons in Information Law*, The Hague: Kluwer Law International, 2006.
- Dreier, T. and G. Schulze (eds), *UrhG*, 3rd edition, Munich, Beck, 2008.
- van Echoud, M., P. Bernt Hugenholtz, S. van Gompel, L. Guibault and N. Helberger, 'Harmonizing European Copyright Law – The Challenges of Better Law Making', Alphen aan den Rijn: Kluwer Law International, 2009.
- van Echoud, M., S. van Gompel, L. Guibault, B. van der Sloot and B. Hugenholtz, 'Dutch report', ALAI Study Days, Dublin, June 2011.
- Elkin-Koren, N., 'Of Scientific Claims and Proprietary Rights: Lessons from the Dead Sea Scrolls Case', in 38 *Houston Law Review* 445 (2001).
- European Commission, Green Paper on Copyright in the Knowledge Economy, Brussels, 16 July 2008, COM(2008) 466/3.
- European Commission, Report on the responses to the Public Consultation on the Review of the EU Copyright Rules, DG Internal Market and Services, Brussels, July 2014.
- Goldstein, P., and B. Hugenholtz, *International Copyright – Principles, law, and practice*, Oxford: Oxford University Press, 2013.
- Guibault, L. and T. Marogni, "The Legal Aspects of Open Science", background paper for the OECD, 2014.

- Guibault, L., 'Licensing Research Data under Open Access Conditions under European Law', in: D. Beldiman (ed.), *Access to Information and Knowledge: 21st Century Challenges in Intellectual Property and Knowledge Governance*, Cheltenham: Edward Elgar, 2013.
- Guibault, L. 'Owning the Right to Open Up Access to Scientific Publications', in L. Guibault & C.J. Angelopoulos (ed.), *Open Content Licensing: From Theory to Practice*, Amsterdam: Amsterdam University Press, 2011.
- Guibault, L. and G. Westkamp et al., Study on the implementation and effect in member state's laws of directive 2001/29/EC on the harmonisation of certain aspects of copyright and related rights in the information society, Institute for Information law, University of Amsterdam, 2007.
- Guibault, L., 'Why Cherry Picking Never Leads to Harmonisation: The Case of the Limitations on Copyright under Directive 2001/29/EC', *2 Journal of Intellectual Property, Information Technology and E-Commerce Law* (2010).
- Guibault, L. and A. Wiebe (eds.), *Safe to be open - Study on the protection of research data and recommendations for access and usage*, Göttingen: Göttingen University Press, 2013.
- Guibault, L., 'Owning the Right to Open Up Access to Scientific Publications', in: L. Guibault and C. Angelopoulos (ed.), *Open Content Licences : From Theory to Practice*, Amsterdam: Amsterdam University Press, 2011.
- Hilty, R. M., K. Köklü and S. Nérison, 'Zum Referentenentwurf eines Gesetzes zur Einführung einer Regelung zur Nutzung verwaister Werke und weiterer Änderungen des Urheberrechtsgesetzes sowie des Urheberrechts-wahrnehmungsgesetz, Stellungnahme des Max-Planck-Instituts für Immaterialgüter- und Wettbewerbsrecht zur Anfrage des Bundesministeriums der Justiz vom 20. February 2013, available at http://www.ip.mpg.de/files/pdf2/Stellungnahme-BMJ-UrhG_2013-3-15-def1.pdf
- Iida, K., S. Imako, Y. Iwamoto, O. P. Chuan, H. Katsunuma, K. Konishi, J. Kobayashi, Y. Takada, T. Nakazaki, Question Q216B Exceptions to Copyright protection and the permitted Uses of Copyright works in the hi-tech and digital sectors AIPPI National Group: Japanese Group.
- Japanese Group, Database protection at national and international level, AIPPI, Report Q182, Geneva, 2004.
- Lemley, M., 'Are universities patent trolls?', *18 Fordham Intellectual Property, Media & Entertainment Law Journal* (2008).
- Love, B., 'Do University Patents Pay Off? Evidence from a Survey of University Inventors in Computer Science and Electrical Engineering', *16 Yale Journal of Law and Technology* (2014). (forthcoming)
- Moscon, V., Open Access to Scientific Articles: Comparing Italian with German law, available at <http://kluwercopyrightblog.com/2013/12/03/open-access-to-scientific-articles-comparing-italian-with-german-law>.
- Nair, M. 'Canada and Israel – Cultivating Fairness of Use', PIJIP Research Paper, No. 2012-04 American University, Washington College Of Law.

- Nauta Dutilh, *The implementation and application of Directive 96/9/EC on the legal protection of databases*, Brussels, 2001, Contract ETD/2001/B5-3001/E/72, available at: http://ec.europa.eu/internal_market/copyright/protdatabases/index_en.htm
- Nimmer, D., 'Copyright in the Dead Sea Scrolls – Authorship and Originality', 38 *Houston Law Review* (2001).
- Nordemann, A. in J. Fromm and A. Nordemann (eds), *Urheberrecht*, 10th edition, Stuttgart: Kohlhammer, 2008.
- OECD, *Principles And Guidelines For Access To Research Data From Public Funding*, Paris, OECD, 2007
- Communication of the European Commission on scientific information in the digital age: access, dissemination and preservation, COM(2007)56, Brussels, 14.2.2007
- Open Society Institute, *Open Access Publishing and Scholarly Societies – A Guide*, New York: OSI, 2005.
- Perry, M. and T. Margoni, 'From music tracks to Google maps: Who own computer-generated works?', 26 *Computer Law & Security Review* (2010).
- RECOMMENDATION IN ORDER TO MOVE FORWARD, adopted by the delegates of the "Berlin 3 open access" conference (Feb 28th - Mar 1st, 2005, University of Southampton, UK).
- Reichman, J.H. and R.L. Okediji, 'When Copyright Law and Science Collide: Empowering Digitally Integrated Research Methods on a Global Scale' 96 *Minnesota Law Review* (2012).
- Ricketson, S. and J. Ginsburg, *International copyright and neighbouring rights*, Oxford: Oxford University Press, 2006.
- Ricketson, S. WIPO study on limitations and exceptions of copyright and related rights in the digital environment, WIPO, 2003.
- Senftleben, M., *Copyright, limitations and the Three-step test*, The Hague, London, Kluwer Law International, 2004.
- Sherman, B. 'Australia', in P. Geller & L. Bently, *International copyright law and practice*, Lexis Nexis, 2013.
- Spoor, J.H., D. W. F. Verkade, D. J. G. Visser, *Auteursrecht*, Deventer: Kluwer, 2005.
- Synodinou, T.E., 'The Foundations of the Concept of Work in European Copyright Law', in: T.E. Synodinou (ed.), *Codification of European Copyright Law – Challenges and Perspectives*, The Hague: Kluwer Law International, 2012.
- Triaille, J., (ed.), *Study on the application of Directive 2001/29/EC on copyright and related rights in the information society*, De Wolf & Partners, 2013 [De Wolf Study].