



Perspectives from Intellectual Property Law

Lisa van Dongen

Contents

- 18.1 Introduction – 442**
- 18.2 Meeting the Criteria – 443**
 - 18.2.1 The Formal Requirements of Copyright – 443
 - 18.2.2 *Sui Generis* Database Right – 445
 - 18.2.3 Trade Secret Right – 446
 - 18.2.4 Summary – 447
- 18.3 The Scope of Protection – 448**
 - 18.3.1 Copyright: Protected Subject Matter – 448
 - 18.3.2 *Sui Generis* Database Protection – 449
 - 18.3.3 Trade Secret Right – 451
 - 18.3.4 Summary – 451
- 18.4 Exceptions and Limitations – 452**
 - 18.4.1 Limitations of the Rights – 452
 - 18.4.2 Exceptions: Common Ground – 454
 - 18.4.3 Exceptions Specific to the Right – 454
- 18.5 Alternative Sources – 455**
 - Further Reading – 458**

Learning Objectives

- What copyright, sui generis database rights, and trade secrets entail and how to determine their beneficiaries.
- When and how the use of third-party datasets is restricted by these rights and when not.
- The potential and limitations of alternative sources to complement or substitute third-party datasets, such as data portability rights and public sector information.

18.1 Introduction

More and more information is collected via the use of smart devices (e.g., smart thermostat, smart phone), internet services (e.g., Google and Facebook), sensors (e.g., in cars, smart homes, and cities), and cameras. The resulting datasets contain a lot of information about individuals, but also about society at large. These datasets allow their observers to spot problems and explore ways to address them, but also to spot opportunities and explore how to exploit them. For example, by studying information from the sensors of cars, the sensors and cameras pointed at the roads, and traffic light systems, it is possible to identify the causes of car accidents and propose solutions to decrease the number of accidents on a certain block. However, the access to such datasets generated by others is often restricted. There is a big group of actors who do not want others to use “their” data. A very important factor that helps such actors restrict access to their datasets is intellectual property law. The holder of intellectual property rights on a dataset has the ability to restrict access of everyone else to (parts of) his/her dataset, as well as impose limits on its use. To understand how to navigate this field of law, it is important to first understand what purpose intellectual property rights serve.

As articulated in the Enforcement Directive, the main underlying reason in current intellectual property law systems is incentivizing (investment in) innovation. Intellectual property rights have been created as artificial property rights to correct certain market failures. Think of the market as a field filled with fruits. If everyone is free to use the field and its fruits without any restrictions, it is likely that many will do so. What is unlikely, however, is that everyone using the field will also individually invest in it. This is due to the uncertainty that it will yield them any results or even allow them to recoup their investment; after all, everyone is free to use the field without restrictions. When the projected proceeds are smaller, less people will be willing to invest. Moreover, any investments that are made are likely to be smaller. This is where intellectual property rights come in. They are tools to correct this “market failure” by rewarding those who invest in innovation with a set of exclusive rights for a limited time. These rights are tools for the right holder to legally restrict the access and use of his/her intellectual property. This allows the right holder to charge higher prices to recoup investments and make a profit.

This chapter aims to provide an introduction to the basics of intellectual property rights in the EU. It uses simplifications and does not always provide the entire picture to maximize understanding of the material. Such simplifications are generally pointed out, and sources on the topic have been included in the references for those wishing to gain a deeper understanding of such an underdeveloped concept. It is, therefore, not to be used as a substitute for legal advice or as a basis for academic debates. Furthermore, while there are many different types of intellectual property rights, only **copyrights**, ***sui generis* database rights**, and **trade secrets** will be discussed here. Under the legal framework of the EU, the subject matter and conditions of these intellectual property rights are closely related to data and software, as will become apparent further in the chapter. Other rights such as patents currently play a more complicated role in the EU in data and software *inter alia* due to limitations in patentability of subject matter such as mathematical methods and computer programs as such. Such limitations have also started playing more of a role in, for instance, the United States, as can be inferred from the case law of their Supreme Court (i.e., on the “abstract idea” concept) between 2010 and 2014. This subject thus requires more attention than it could receive in this limited contribution.

This chapter thus focuses only on these particular intellectual property rights from an EU perspective. The questions explored in the following sections will focus on establishing for each of these intellectual property rights when it would be applicable (► Sect. 18.2), followed by what this means for the data’s usage by a third party (► Sect. 18.3), as well as limitations and exceptions (► Sect. 18.4). This chapter concludes by discussion of ways to gain lawful access to datasets covered by one or more of these intellectual property rights and alternative sources.

18.2 Meeting the Criteria

18.2.1 The Formal Requirements of Copyright

Something might be protected by copyright if it meets the three cumulative criteria for copyright protection. Following the Berne Convention, these criteria require that it is (1) an expression (2) that is original (3) in the area of literature and art. In the EU’s copyright regime, factors such as labor or investment are not relevant. There are three elements of a dataset that are capable of meeting these requirements:

- The contents of the dataset
- The selection of the data
- The arrangement of the data

If one or more of these elements meet the formal requirements, there might be copyright protection on those elements of the dataset. In that case, there would be legal restrictions on its use. It is thus important to understand these criteria to be able to determine the likelihood of copyright protection on a dataset to ensure lawful use.

The underlying premise of the criterion **expression** is that facts and ideas are not created but discovered. This is also confirmed in *Feist Publications, Inc., v. Rural Telephone Service Co.*, which shows that the United States and the European Union approach this criterion in a similar fashion. What this means for copyright is that it does not protect *what* is said, but *how* it is said. A good rule of thumb is looking at it as a spectrum in which facts and ideas are on one side and expressions on the other based on specificity. In principle, the more specific a fact or idea becomes, the closer the needle generally moves towards expression. The reasoning behind this is that an author can convey a fact or idea choosing his/her own words, thereby creating something both beyond and separate from the fact or idea. To illustrate, look at the difference in detail in the following sentences in **Table 18.1**.

This requirement is a possible hurdle for copyright protection on a dataset. For example, data in such datasets together may create a very specific picture, but if the data is merely displayed as variables in a table, the data lack expression.

In *Football DataCo Ltd.*, the second element—**originality**—was understood as a margin of discretion to make free and creative choices that is utilized. In simpler terms, it requires that the creator has put his/her personal stamp on it. However, this of course should not be taken literally. For instance, putting your logo on something does not make it original. The bar for meeting this criterion is not very high in practice. Such creative choices can be as simple as selecting lighting, a background and an angle for making a picture, or word choice in a text or code. It is, however, important to emphasize that there should be room to make such choices by the creator. For instance, a passport photo has to meet a

Table 18.1 Expression

Example sentence	Level of detail	Fact/idea/expression
This house is green	Very little detail and very general	Fact/idea
This three-story house is three different shades of green	More detail, but still quite general	Fact/idea, but already more towards expression
This three-story living accommodation is a mix of shades of green, amongst which olive, moss, and even some hints of metallic green around the corners of its windows and doors	A lot of detail and very specific	Expression

Note to table: Author's own table

number of strict requirements. Such predetermined settings affect the room the photographer has to make his/her own creative decisions. For a passport photo, it is thus highly unlikely that the photographer would be able to meet the originality requirement. Another example is functionality requirements. Software code is capable of attracting copyright protection since the Software Directive came about, but, as confirmed in *Bezpečnostní softwarová asociace*, the expression in the code cannot amount to originality if “dictated by their technical function.”

The circumstance that the author has room to make creative choices is thus vital for meeting the originality requirement. Moreover, in the absence of such requirements, there is still the matter of whether creative choices are actually made. The selection and/or arrangement of data in a dataset can, for example, meet the minimum threshold of creativity, but these choices are generally made based on utility in practice; the choices made in selecting data are often determined by a company’s primary business, and the data are arranged for practical reasons such as by alphabetical order or by date.

The last criterion requires that it is a work in the area of **literature and art**. What constitutes art or literature is understood very broadly in the copyright regime. For instance, literature for the purpose of copyright protection can include essentially anything involving the written word. As mentioned above, it can even cover the code in software. This means that data—whether numeric or text—also falls within this broad category. Some other examples of works that may be protected are books, paintings, sketches, maps, architecture, preparatory design material for software code, films, musical compositions, lyrics, topography, choreographic works, and so forth: Article 2(1) of the Berne Convention contains well over 20 examples of types of works falling within the ambit of literature and art.

18.2.2 *Sui Generis* Database Right

If materials such as datasets and preparatory design material for software code are part of a database, their use may be restricted by *sui generis* database protection. Due to the limited protection provided by the copyright regime in databases, the Database Directive was adopted in 1996 to further strengthen the information economy in the EU. To this day, this regime is still very much a European creation (there is, for instance, no equivalent in the USA). The *sui generis* database right thus protects databases without originality. However, that does not mean that if there is copyright protection on the contents, selection, and/or arrangement of the database, there cannot also be *sui generis* database protection. The two rights can coexist on a single database. A dataset is likely to be covered by this right if it is (1) a database for which a (2) relevant investment was made (3) that is substantial.

For a dataset to satisfy the first condition—that it is a **database**—it is first required that the dataset is a collection or compilation of materials. Such materials include copyrighted works, numbers, facts, and data, but are not limited to those categories. Next, such materials must then be organized, stored, and accessible via electronic or nonelectronic means. This means that a written document meeting all the other requirements could also qualify as a database. However, for a physical database, it is not necessary that the materials are physically stored in an organized manner.

The second criterion requires that a **relevant investment** is made. This means that the investment must be made in the collection, verification, and/or presentation of data for the database. As clarified in *BHB v William Hill*, investment in other categories such as in the creation of data is not relevant for meeting this criterion. Such an investment can be made by way of financial resources, human resources, and material resources. Investment via human resources can, for example, be made in effort or time. For material resources, the investment is made in equipment to build the database such as hardware and software. Of course, such type of investments also cost money. Moreover, human input is generally required in operating equipment to make a database. In reality, the connection between these three types of investment thus often makes for a combination of the three with the emphasis on financial resources. Moreover, such investments should not have been made for other purposes. For instance, computers used to create the database are often not solely bought for that purpose. In that case, the investment generally does not count towards the coming into being of *sui generis* database protection.

The last criterion—that the investment must be **substantial**—is a bit more ambiguous. The Database Directive does not provide conclusive guidance on what this criterion means or how it should be applied. Case law so far has mostly dealt with high sums of financial investment, so these cases do not provide much guidance on the substantial threshold either. Unfortunately, the exact ceiling and floor of this criterion are also still subject to heavy academic debate, but it would be beyond the purposes of this chapter to include these. This threshold should, different than the word substantial might suggest, not be interpreted as “high.” Instead, this criterion is best understood as requiring an investment that is not too unsubstantial. These perimeters in the main text—not high, just not too unsubstantial—are generally accepted in EU member states such as Germany. A clear example of such an insubstantial investment would be a single employee of a big company devoting only a few hours to making the database. An example of something that would qualify would be investments in verifying a great quantity of data with another dataset.

18.2.3 Trade Secret Right

Following the Trade Secret Directive, if the dataset consists of (1) information not known in the relevant circles, (2) is of commercial value, and (3) is kept secret by the company in question, the dataset may be protected as a trade

secret. The first criterion requires that the information in question is **not readily accessible or known in the relevant circles**. The relevant circles refers to people generally dealing with this type of information, which means that the relevant circle may differ per type of information if the protected subject matter consists of different types of information. It can, therefore, not cover insignificant information or the kind attained through normal employment experience. Information that can be covered by a trade secret right at least includes know-how, business information, or technological information, but may be defined broader in domestic law.

Second, the information should be of **commercial value**. It does not matter whether it does so actually or potentially. What is important is that the interests of the right holder of the trade secret—whether scientific, technical, business, or financial in nature—would be harmed if the trade secret would be compromised. It should thus have commercial value because it is secret. If the value would not be affected negatively if it would be misappropriated, satisfaction of the second criterion is questionable.

Finally, the holder of the trade secret right should make reasonable efforts in **keeping the information secret**. Of course, this is subject to the circumstances of the case. In some cases, it might be more difficult to keep the information a secret or the circumstances may require different measures than in others. The fact that many people know does not necessarily mean that the company has failed in its effort to satisfy this criterion. For instance, many employees might require knowledge of (parts of) the trade secret in order to be able to make a product. As long as they are under contractual obligations to secrecy, it does not matter how many know. The same is true for distributors who have received certain information under a nondisclosure agreement to be able to do their job.

18.2.4 Summary

The formal requirements of each of the intellectual property rights can be broken down into three basic components. Put next to each other in a table, that creates the following picture (■ Table 18.2).

■ Table 18.2 Formal requirements

Copyright	<i>Sui generis</i> database right	Trade secret right
<ol style="list-style-type: none"> 1. Expression 2. Originality 3. Literature and art 	<ol style="list-style-type: none"> 1. Database 2. Relevant investment 3. Substantial 	<ol style="list-style-type: none"> 1. Not readily accessible or known in relevant circles 2. Commercial value 3. Kept secret

Note to table: Author's own table

18.3 The Scope of Protection

18.3.1 Copyright: Protected Subject Matter

If a dataset would be protected via one or more of these routes, there is still the limitation of what these rights actually protect and against what. When a dataset or software code meets the requirements for copyright protection, this protection is limited to the **original expression** only. This means that the protection can never extend to, amongst other things, factual content or ideas. Additionally, if only the selection and/or arrangement of a dataset are protected by copyright and not the data itself, the expression exists only in the selection and/or arrangement. For software code, this means that copyright can only rest on code *not* dictated by technical functions. A third party would thus be able to use the contents of the dataset or such unprotected parts of the software's code.

Moreover, copyright only protects the original expression against certain types of use by others. In other words, the copyright holder has certain rights to *exclude*. Different from what the term “copyright” suggests, it constitutes not one right but a bundle of rights. The bundle of rights contains exploitation rights, otherwise known as economic rights. There are several economic rights included in the InfoSoc Directive, but only the right to reproduction and the right to make public are of particular relevance for data usage and software. The **right to reproduction** entails that, in principle, only the copyright holder has the right to make copies of his/her work. Furthermore, it is important to note that a reproduction does not have to be exact. Making a photo of a painting is also reproducing the work. The means used to make a copy do not matter for this right. Furthermore, it does not have to be a copy of the entire work. What is important is that enough should be copied to display the intellectual and creative work of the artist. A sample as little as 11 words from newspaper articles has been found capable of doing that in *Infopaq v Danske Dagblades Forening*. Consequently, it is arguable that a small part of the dataset or code could also convey the creative choices of the author. If so, in the absence of an applicable exception, even the use of such small excerpts requires authorization. Second, there is the **right to make public**. Think, for instance, of putting a protected content on a website or using hyperlinks to protected content. Take care that this is somewhat oversimplified. What should be understood as making available to the public and who should be understood to be conducting this act are still evolving due to certain recent legislative and judicial developments at the EU level. In most cases, a reproduction of some sorts is necessary to be able to make it public. Notable exceptions here are the use of hyperlinks or displaying the original (i.e., a painting in a museum).

What this means for third parties is that they cannot lawfully engage in these uses of the original expression without authorization. The copyright holder can, for instance, grant others permission to reproduce his/her work via a license. Since 11 words could already convey creative choices from the author, the requirement of having to obtain authorization kicks in fast. In principle, such authorization can

only be obtained from the copyright holder. The **right holder** is generally a natural person—the author or creator. When a work has been created in assignment, the allocation of the copyright depends on who has made the creative choices. In some cases, the creative choices may have been made by several actors, which generally leads to shared rights to a work. However, this is different in case of creation under employment. For instance, the exploitation rights on a work are located with the employer if created by an employee in the course of his/her employment upon instructions by the employer. Additionally, in the case of software, the Publications Office clarified in their summary of the Enforcement Directive that EU member states may lay down that legal persons or entities may also be the right holder. In some jurisdictions, not all rights may always be transferable from the author to another.

18.3.2 *Sui Generis* Database Protection

The *sui generis* database right was created with the investor in mind, so just being the factual maker is insufficient to be the right holder. According to the Database Directive, the **right holder** is the person who takes the initiative and the risk of investing. Subcontractors and work for hire are explicitly excluded from this definition. If a database is made by an employee, the allocation of the rights depends on the criteria in national law. If multiple people or entities have contributed to a database, there might be joint rights. Unlike copyright, the *sui generis* database right is fully transferable. Like in copyright, the *sui generis* database right is not a single right. When a database is covered by *sui generis* database protection, the right holder has the exclusive rights to (1) extraction and (2) reutilization. These rights should be understood as follows.

Extraction refers to the transfer of the database or a substantial part thereof. This transfer may be permanent or temporary. Moreover, the means through which it is transferred do not matter. It is also irrelevant where the database is transferred to (type of medium). What matters is that the database or a substantial part thereof is transferred. This means that any person other than the right holder in principle requires the authorization from the right holder to perform this act lawfully. However, authorization is also required for systematic extraction of insubstantial parts. This is included in the definition of extraction to combat “milking.” This is the process of repeatedly transferring small parts of a database until the entire database or a substantial part thereof is transferred.

The other type of use—**reutilization**—refers to making the database or a substantial part thereof available to the public. This includes the distribution or renting of copies, online transmission of the database, and other types of transmissions. Any way in which the database is made public falls under this definition. In essence, this right thus gives the right holder the sole right to conduct an incidental reutilization of (a substantial part of) the database. However, just like the right to extraction, the right to reutilization also protects against the systematic reutilization of insubstantial parts. Again, if this definition were limited to substantial

parts or the entire database, this would provide third parties the opportunity to still communicate (a substantial part of) the database, just a smaller part at a time. Finally, there is one last instance in which there is reutilization. It involves the use of a meta search engine with certain functionalities.

A meta search engine is a search engine that makes it possible to search through a number of other databases. Generally, it transfers the search query that is inserted by a visitor of the meta search engine to other search engines. It does not copy anything from the databases through which it searches, but shows the results from the search, including those from other databases. It was established in *Innoweb BV v Wegener* that such a meta search engine is *likely* to reutilize (a substantial part of) the database if the three following functionalities are present. First, the search forms offered to the end user by the meta search engine and the other database function essentially the same. Second, the queries are translated for the end user in real time to other search engines. This means that all the information of the other database is searched through in real time after the end user of the meta search engine has initiated the search. Third and finally, the results are presented all together in an order that reflects similar criteria to those used by the other database. To this end, the format of the meta search engine's own website is used in showing the results, showing duplicates together as a block item. To reiterate, if a meta search engine that searches other databases functions in the aforementioned way, the operator of this meta search engine is likely to engage in the reutilization of (substantial parts of) another database. Of course, this does not mean that if a meta search engine does not have these characteristics, there could not nevertheless be reutilization.

For both of these rights, the word **substantial** plays a role again. For the purposes of extraction and reutilization, the term "substantial" refers to the volume of data from a database, more specifically, the volume of data that is extracted or reutilized in relation to the whole database (see *BHB v William Hill*). There is a link here between the investment and the two rights. The easy way to approach this is quantitatively. Consider the following example. There was substantial investment in the collection, verification, and/or presentation of the data, but no significant differences in the investment across the data. A third party now extracts half the data of the database. That means half of the investment is represented by the extracted part. The part that is extracted is thus likely to be substantial. However, whether the extraction or reutilization is substantial can also be tested qualitatively. This is a bit more ambiguous. The circumstances of our example change somewhat. Now, there is certain data in the database that has required much more investment in their collection, verification, and/or presentation than the rest of the data. The more "expensive data" is only a small part of the entire database. A third party now reutilizes only the part of the database that contains the "expensive data." Even though it is less data, it represents a bigger part of the investment. This means that it is likely that such reutilization by a third party would be qualitatively substantial. In both examples, the third party probably cannot conduct these acts without authorization from the right holder or by law.

18.3.3 Trade Secret Right

The Trade Secret Directive stipulates that the **trade secret holder** is any natural or legal person lawfully controlling the trade secret. Like the *sui generis* database right, this right can be fully transferred. The trade secret right protects against the unlawful acquisition, use, and/or disclosure of protected subject matter. These acts are to be construed very broadly. Any act contrary to honest commercial practices, unauthorized access, and/or appropriation of any material that contains the protected subject matter falls under unlawful **acquisition**. The same is true for material from which the trade secret information can be derived. Of course, if a person then proceeds to **use and/or disclose** the trade secret, this too would be unlawful. Use or disclosure of protected subject matter in breach of a contractual duty—including a confidentiality agreement—or any other duty imposing limits on those acts is also unlawful. Moreover, unlawful use includes the production of infringing goods, or offering or placing those on the market. Storing, importing, and exporting infringing goods to that end also fall within that definition. A good is infringing if the unlawfully acquired, used, or disclosed protected subject matter contributes in a meaningful way to (the production process or marketing of) a product.

The trade secret right is arguably the most fragile intellectual property right. When copyright or *sui generis* database right is infringed, these intellectual property rights will continue to exist. Once the data covered by a trade secret right is misappropriated in a way that it no longer satisfies the conditions regarding its secrecy, the right lapses. However, it is important to reiterate that the trade secret protects against *unauthorized* acts. Consider the following example. Data covered by a trade secret is disclosed under a nondisclosure agreement against payment. If the duties of the provider and acquirer—contractual and otherwise—do not prevent this transfer of data under the circumstances, it is likely to be lawful. Such disclosure presumably leaves the trade secret intact. Trading under a nondisclosure agreement does not necessarily result in the loss of the trade secret right. Contracts such as employee contracts with confidentiality clauses and nondisclosure agreements are thus vital tools for the holder of the trade secret right.

18.3.4 Summary

If a dataset qualifies for copyright, *sui generis* database protection, and/or a trade secret right, the protection is still limited to certain subject matter. Moreover, it is only protected against certain unlawful acts performed by someone other than the right holder (see ■ Table 18.3). Such acts are unlawful without authorization provided by the right holder or law (i.e., exception).

■ **Table 18.3** Scope of protection

	Copyright	<i>Sui generis</i> database right	Trade secret right
Protected subject matter	Original expression	Database	Trade secret
Protected against unlawful	<ul style="list-style-type: none"> • Reproduction • Making public 	<ul style="list-style-type: none"> • Extraction • Reutilization 	<ul style="list-style-type: none"> • Acquisition • Use • Disclosure

Note to table: Author's own table

18.4 Exceptions and Limitations

18.4.1 Limitations of the Rights

In some cases, a use by a third party falls outside the scope of the right. Limitations, as the word suggests, limit the protection. For instance, intellectual property rights do not last indefinitely. In the EU, copyright lasts up to 70 years after the death of the author following the Term Directive. According to the Database Directive, *sui generis* database protection lasts for 15 years starting from the day of completion of the database, but the clock restarts with every new substantial change and/or investment. Trade secret rights are the exception here: there is no **maximum term of protection** inserted in the Trade Secret Directive. The trade secret right will last until its protected subject matter no longer satisfies the criteria for this right.

As aforementioned, copyright does not extend to **facts and ideas**. Moreover, even subject matter that is neither fact nor idea can fall outside the scope of the protection when it is not part of the original expression. Furthermore, originality means that the creative choices are made by the author, not that it should be new. This means that it does not protect against **independent creation**.

For the *sui generis* database right, protection revolves around the investment. If a third party incidentally **extracts** and/or **reutilizes insubstantial parts**, in principle, that would be lawful. However, there are some boundaries there as well. In doing so, the Database Directive requires the third party to take care that its acts do not conflict with the normal exploitation of the database by the right holder or unreasonably prejudice his/her interests. In short, acts by a third party should not “harm” the investment.

The scope of protection offered by the trade secret right also has its limitations under the Trade Secret Directive. The trade secret right only protects against unlawful acts. This means that **independent creation** or **discovery** does not interfere with trade secret rights. Moreover, **reverse engineering** after lawfully obtaining a prod-

uct would also not breach the trade secret right. This means that it would be lawful for a third party to buy a product that was brought on the market in the EU by the right holder and study its functioning to improve his/her own production process, product, and/or service. A car manufacturer could, for instance, buy a car sensor offered on the market by a competitor to reverse engineer it and use the gained knowledge to improve its own car sensors.

Finally, several references have been made to the transferability of rights allocated to the right holder by these intellectual property rights. What this means is that it is generally possible to contractually “reserve” or transfer such rights or allow acts under certain circumstances. Right holders themselves can thus also contractually limit their own rights. For the reservation of rights, think, for example, of a situation of shared rights. It could be beneficial for parties to lay down contractually that the authorization of *all* right holders must be obtained, not just one. Alternatively, a right holder could transfer the sole right to an exclusive distributor to enforce the intellectual property right against (alleged) infringers, thereby freeing his/her own hands. An example of allowing acts under certain conditions can be found in many terms of service in the gaming industry. Such terms often contain a clause allowing their users to engage in acts such as live streaming themselves playing the game in question. Another fitting example is the use of a threshold, allowing users to use protected material as long as they do not gain profit over a certain established number or reach a set number of clients (■ Table 18.4).

■ Table 18.4 Limitations

	Copyright	<i>Sui generis</i> database right	Trade secret right
Maximum term	70 years after the author's death	15 years, but renewable	–
Outside the scope	<ul style="list-style-type: none"> • Facts • Ideas • Independent creation 	<ul style="list-style-type: none"> • Extraction of insubstantial parts • Reutilization of insubstantial parts 	<ul style="list-style-type: none"> • Independent creation • Independent discovery • Reverse engineering
Contractual limitations possible	<ul style="list-style-type: none"> • Yes, on exploitation rights^a 	<ul style="list-style-type: none"> • Yes 	<ul style="list-style-type: none"> • Yes

Note to table: Author's own table

^a As mentioned earlier, the bundle of rights is not always transferable in its entirety. However, it is important to note that this generally does not apply to exploitation rights

18.4.2 Exceptions: Common Ground

If an act is covered by an exception, it is authorized by the law. This means the right holder cannot authorize nor prevent the act. Following legal instruments such as the Berne Convention, the TRIPS Agreement, and the InfoSoc Directive, the exceptions should be limited to special cases and not interfere with normal exploitation of the work nor unreasonably prejudice the legitimate interests of the author. In general, these exceptions are thus applied narrowly across the EU in favor of high protection of intellectual property rights.

The exceptions vary somewhat per intellectual property right, but there is some common ground. For instance, the exception for **teaching and research** and **public security purposes or an administrative or judicial procedure** exist both in the InfoSoc Directive (on copyright) and the Database Directive. In the former, these are exceptions to the right holder's reproduction right. In the latter, these exceptions target both the extraction and reutilization right. In the *sui generis* database regime, however, these exceptions can only be relied upon by a "lawful user." Think, for instance, of circumventing the requirement of a subscription to gain access to a nonpublic database without authorization. Extraction and/or reutilization by such a user cannot fall within the scope of these exceptions.

An example of teaching and research could be the showing of clips, (preparatory design material for) software code, small texts, or parts of a database for illustration to students or researchers. In order to qualify, both regimes require that third parties must not perform such uses for commercial purposes. Where possible, the source should be referenced and the use should not go beyond what is required for the noncommercial purpose pursued. For the exception for public security purposes or an administrative or judicial procedure, an example could be the copying of a work or certain data from a database to verify imported goods. Another could be the inclusion of such materials in the written decision of a court case revolving around questions of infringement of copyright and/or *sui generis* database protection. Again, such acts may not have been conducted for commercial purposes.

18.4.3 Exceptions Specific to the Right

The most common and relevant exceptions specific to the EU copyright regime are journalism, citation for review and criticism, and caricature, parody, and pastiche. EU law, more specifically the InfoSoc Directive, does not provide any conditions for any of these exceptions. This means, for example, that member states were free to limit exceptions to certain circumstances or uses only.

Finally—and perhaps most importantly—there is the recently introduced **text and data mining** exception in the Digital Single Market Directive. This concept is best understood as any analytical technique that is automated. It is used to derive information by analyzing text and data in digital form. This exercise could, for example, be performed to discover patterns, trends, and correlations in a dataset. Two variations of this right have been introduced, one focusing on text and data

■ **Table 18.5** Summary of exceptions

	Copyright	<i>Sui generis</i> database right	Trade secret right
Exceptions	<ul style="list-style-type: none"> • Teaching and research • Public security purposes and administrative or judicial procedure • Journalism • Citation for review and criticism • Caricature, parody, and pastiche • Text and data mining 	<ul style="list-style-type: none"> • Teaching and research • Public security purposes and/or administrative or judicial procedure 	<ul style="list-style-type: none"> • Freedom of expression and right to information • Revealing misconduct, wrongdoing, or illegal activity • Legitimate tasks of worker(s) (representatives)

Note to table: Author's own table

mining for scientific purposes and a general one. Both require that there was lawful access to the works that is to be subjected to text and data mining. The first type allows for the storage and retention of reproductions of the works for scientific research. However, there should be an appropriate level of security present on the storage of the copies of the works. For the general exception, there is the precondition that the right holder has not explicitly reserved the use of his/her work. In the absence of such a reservation, the works may be “mined,” kept, and stored as long as is required for the aim pursued with the text and data mining.

Under the regime of trade secret rights, the most relevant exceptions are the following three. First, the act may not infringe the trade secret right if the **freedom of expression and right to information** can successfully be invoked. Moreover, the acquisition, use, or disclosure of subject matter protected by a trade secret right in the pursuit of **revealing misconduct, wrongdoing, or illegal activity**, such as whistleblowing, may also be lawful. Furthermore, linked to one of the previously mentioned limitations, if their **legitimate tasks as workers or workers' representatives** necessitated the disclosure, the holder of a trade secret right may also not be able to apply for remedies against them (■ Table 18.5).

18.5 Alternative Sources

These intellectual property rights may vary in scope and purposes, but it is entirely possible that several may be applicable to (parts of) the same dataset or code. The exceptions to these various rights are limited to specific rights and purposes. Therefore, it is possible that an act that falls under an exception for one intellectual property right is not allowed due to the presence of another right. If a third party requires access to datasets (partially) covered by these rights, there are several options to gain lawful access to datasets.

The most straightforward option is to obtain a **license** from the right holder to use his/her datasets. A license allows the licensee to use protected subject matter in accordance with the agreed-upon terms, usually against payment. Protected subject matter can be licensed for some or all of the uses covered by copyright and/or *sui generis* database protection, but the author or maker remains the owner. Another similar option here would be to enter into an **ad hoc agreement or partnership** either by paying a sum or by offering something in return.

Alternatively, it is sometimes also possible to gain access to comparable datasets via other sources, such as **public sector information** or “PSI.” This is a very interesting and useful opportunity to consider because the state possesses many data—think, for instance, of maps, court decisions, company data, citizen statistics, etc.—and might have an obligation to release that data and allow its reuse (i.e., Freedom of Information Acts), although not necessarily for free. Data are likely to be subject to a PSI regime when (1) linked to the execution of state activities, (2) there are no intellectual property rights owned by third parties on them, and (3) the data are not kept secret for reasons of public policy (including data protection).

Depending on the business model, another option to consider might be using software or data subject to **open licensing schemes** (“open source”). The use of such data or software is free, but, depending on the type of license, there may be other types of restrictions. The most common division made is between *permissive licenses* and (weak or strong) *copyleft licenses*. These types of licenses are best envisioned as a spectrum from least prescriptive to most prescriptive license. Both types of licenses do not restrict the use of the subject matter in terms of use, modification, and redistribution, but permissive licenses allow proprietary derivative works while copyleft licenses do not. This means, for instance, that a third party can make modifications to the subject matter under the permissive license and license and distribute it under a different type of license. A weak copyleft license, on the other hand, would not permit this. Such licenses contain a clause proscribing making material derived from its subject matter proprietary or relicensing this derived material. Strong copyleft licenses additionally require that its subject matter can also not be licensed against a different license than the original. This means that a work subject to a “normal” proprietary license cannot be combined with another work subject to a copyleft license.

The provision of **complementary services or products** on the market to create or gain access to a similar dataset is also a possibility. For instance, a third party wants similar data as generated by sensors brought on the market by a competitor. The third party could decide to offer software that could operate the sensors from the competitor or offer competing sensors. Another option here would be to turn your own clients into data collectors themselves by having them correct or report certain data. Think, for instance, of reporting additions to a map or modifications to a street.

Finally, if these datasets contain personal data, you can ask those individuals to use their **right to data portability** via a promotion for new or existing customers of their own services or products. According to the General Data Protection Regulation, this right gives natural persons the opportunity to move their personal data from one online service to another. The requirements that have to be met here are that the

data (1) are personal data and (2) have been provided to the controller by the person whom the personal data are about. For instance, an insurance company or municipality could offer benefits in exchange for their personal data transfer to you, such as a discount on the insurance fee or on services offered by the municipality.

In short, if intellectual property rights exist on a dataset and none of the exceptions are applicable, there are still several avenues to gain legal access. Moreover, alternative sources can be explored as a complementary source or substitute for the protected dataset.

Conclusion

To summarize, when dealing with subject matter such as datasets and software, it is important to first establish whether intellectual property rights may exist on them. If so, the use of such subject matter by third parties may be restricted. Which uses are restricted and under what conditions depend on which right applies and, in varying degrees, on which regime applies (i.e., EU or USA).

Second, it should be established who the right holder is. If you are the right holder, this means that you may be able to restrict the access and use of others of the protected subject matter. If it is someone else, there are several possible routes to lawful use of that party's protected subject matter or alternatives to this subject matter, from obtaining consent from the right holder to acting within the scope of limitations or exceptions to finding or creating alternative sources.

Take-Home Messages

- Copyright on the contents, selection, or arrangement of a dataset gives the right holder the sole right to reproduce and make the protected material public.
- The *sui generis* database right gives the right holder the sole right to extract and reutilize substantial parts of the database.
- Trade secret rights on data protect the right holder against unlawful acquisition, use, and disclosure of the protected material.
- A third party can only engage in lawful use of subject matter protected by these rights if it is authorized by the right holder or by law (if exceptions are applicable).
- In the absence of authorization, there are several ways in which legal access can alternatively be gained to (parts of) the dataset or software or to comparable sources.

? Discussion Questions

1. Why do we have intellectual property rights?
2. How would you explain the distinction between ideas and expressions in copyright law?
3. Please define all types of investments relevant for *sui generis* database protection, including the means through which such investments can be made.
4. Please briefly explain the status of reverse engineering under the trade secrecy protection.

Further Reading

- Agreement on Trade-Related Aspects of Intellectual Property Rights, Annex 1c of the Marrakesh Agreement Establishing the World Trade Organization, 1994.
- Berne Convention for the Protection of Literary and Artistic Works of 1886 as amended on September 28, 1979 ('Berne Convention').
- Case C-202/12 *Innoweb BV v Wegener* [2013] ECLI:EU:C:2013:850.
- Case C-203/02 *British Horseracing Board v William Hill* [2004] ECLI:EU:C:2004:695.
- Case C-393/09 *Bezpečnostní softwarová asociace* [2010] ECLI:EU:C:2010:816.
- Case C-5/08, *Infopaq v Danske Dagblades Forening* Case C-5/08, *Infopaq v Danske Dagblades Forening* [2009] ECLI:EU:C:2009:465.
- Convention on the Grant of European Patents of 1973 as amended last on 29 November 2000 ('European Patent Convention').
- Council Directive 93/98/EEC of 29 October 1993 harmonizing the term of protection of copyright and certain related rights [1993] OJ L290/9 ('Term Directive').
- Estelle Derclaye, 'Database sui generis right: what is a substantial investment? A tentative definition' (2005) IIC 36(1).
- Directive 2001/29/EC of the European Parliament and of the Council on the harmonisation of certain aspects of copyright and related rights in the information society [2001] OJ L167/10 ('InfoSoc Directive').
- Directive 2004/48/EC of the European Parliament and of the Council on the enforcement of intellectual property rights [2004] OJ L157/45 ('Enforcement Directive').
- Directive 2009/24/EC of the European Parliament and of the Council of 23 April 2009 on the legal protection of computer programs [2009] OJ L111/16 ('Software Directive').
- Directive (EU) 2016/943 of the European Parliament and of the Council on the protection of undisclosed know-how and business information (trade secrets) against their unlawful acquisition, use and disclosure [2016] OJ L 157/1 ('Trade Secret Directive').
- Directive 2019/790 of the European Parliament and of the Council on copyright and related rights in the Digital Single Market and amending Directives 96/9/EC and 2001/29/EC [2019] OJ L130/11 ('Digital Single Market Directive').
- Directive 96/6/EC of the European Parliament and of the Council of 11 March 1996 on the legal protection of databases [1996] OJ L77/20 ('Database Directive').
- Feist Publications, Inc., v. Rural Telephone Service Co.*, 499 U.S. 340 (1991).
- Husovec, M. (2019). How Europe Wants to Redefine Global Online Copyright Enforcement. *TILEC Discussion Paper*, 2019–2016.
- Publications Office in their Summary of Directive 2009/24/EC—the legal protection of computer programs, 23 January 2017.
- Regulation 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC [2016] OJ L 119/1 ('General Data Protection Regulation').
- Rosati, E. (2017). GS Media and its implications for the construction of the right of communication to the public within EU copyright architecture. *Common Market Law Review*, 54(4), 1221–1242.