THE TEXT AND DATA MINING EXCEPTION IN THE DIRECTIVE ON COPYRIGHT IN THE DIGITAL SINGLE MARKET: A STORMY OUTLOOK FOR AUTHOR'S RIGHTS

Andrée Glancia MADINDA

Winner of the 2022 edition of the GESAC/ALAI European Authors' Right Award

Doctoral student, Institut François Gény laboratory, University of Lorraine

Directive 2019/790 of 17th April 2019, known as the Directive on Copyright in the Digital Single Market,[1] offers fertile ground for many lines of thought to flourish. Enacted on the same day as another, equally interesting directive,[2] Directive (EU) 2019/790 on Copyright and Related Rights in the Digital Single Market has attracted particular attention. In addressing the many developments it contains, the *Text and Data Mining* (TDM) exception will be the focus of our reflections here. Whether in terms of lauding the enforceability of author's rights on the Internet or the revival of collective management, the new significance accorded to the text and data mining exception inevitably attracts attention. It transpires that this aspect of the Directive provokes fears or enthusiasm depending on the point of view adopted.

The Birth of the Directive. The Directive on Copyright in the Digital Single Market is adopted in the spirit of harmonising European copyright law and ensuring it functions effectively.[3] While campaigning for more vigorous Community copyright provisions, the Community legislator took the opportunity to initiate reconciliation with collective management organisations (CMOs) through legislation. It should be noted that, although it was previously accepted that copyright and related rights are inextricably linked to the contractual sphere, the Community legislator seemed unenthusiastic about the way in which CMOs turn the exercise of author's rights into a collective undertaking.[4] Such misgivings can be explained by the differing conceptions of copyright in the European Union's Member States.

In keeping with the French view of what should constitute author's rights, the initial reflex is to be suspicious of this exception which seems, once again, to target moral rights.[5] In fact, moral rights are

[1] Dir. (EU) 2019/790 of the European Parliament and of the Council of 17th Apr. 2019 on copyright and related rights in the digital single market and amending Dir. 96/9/EC and 2001/29/EC, OJEU No. L130 of 17th May 2019 p. 92: Lexbase hebdo éd. Affaires, 11th July 2019, special file; *Propr. Intell.* 2019, n° 72, pp. 33 ff., chron. by A. Lucas, J.-M Bruguière and C. Bernault; *JCP* G 2019, 26, 693, p.1236, obs. V. L. Benabou; *JCP* E 2019, 27, study, p. 1343, by E. Treppoz; RTDcom 2019, p. 648, note F. Pollaud-Dulian; CCE Oct. 2019, 10, special file.

[2] The Directive (EU) 2019/789 of 17th April 2019, known as the CabSat 2 Directive, laying down rules on the exercise of copyright and related rights applicable to certain online transmissions by broadcasters and retransmissions of television and radio programmes, adopted on the same day, will not be the focus in these reflections. For more on this point, see G. Vercken and V. Tabary, "L'autre' directive: la nouvelle directive dite 'CabSat 2' du 17 April 2019", Légipresse, 2019, No. 372.

[3] As stated in the introduction to the Directive "Whereas (1): The Treaty on the Functioning of the European Union provides for the establishment of an internal market and the institution of a system ensuring that competition in the internal market is not distorted. Further harmonisation of the laws of the Member States on copyright and related rights should contribute to the achievement of those objectives.

[4] S. Carré, "Gestion collective: regards au prisme de la DAMUN", *in* RFPI, special issue Dec. 2021, p. 47.

[5] The author's moral right is without doubt an aspect specific to French provisions on author's rights. However, this monumental structure could be compared to a giant with clay feet due to the legal exceptions, customs and

not the only type of rights to suffer. However, it is important to recall that this exception appeared in French law a few years ago and that it is not entirely unexpected that it has been enshrined in European legislation. Discreet calls for "value sharing"[6] have been heard from France for several years now.[7] What does this exception entail in tangible terms?

Definitions. The law is based on a system of principles and exceptions. The principles entail scope to exercise certain rights and the exceptions concern limitation of such rights by means of a waiver. The TDM exception entails authorising a kind of reading grid, turning the spotlight on use,[8] ultimately,[9] of texts and works by researchers in order to make research carried out by another individual more concrete, extend it or add the finishing touch to it.[10] *Stricto sensu,* one might venture to state that it is above all a question of bringing to light a new aspect based on analysis of the relations between the works "mined". Critics of this exception, however, point out that it undermines the author's moral right, in addition to denying the financial remuneration to which he or she could lay claim. Advocates of the exception, on the other hand, emphasise that it contributes to research, innovation and, of course, to the objective of sharing knowledge. Some even claim that Community law needed to be more ambitious than national legislation in this field to be better placed to compete with Anglo-American law.[11] Is reusing the works of others, utilising their content and deriving a "new intellectual work" from this really different from producing a composite work? How could the data mining exception be adopted, given that author's rights provisions are so specific and particular? The question merits attention whichever side of the Channel one may find oneself and indeed even across the Atlantic, for the approach to dealing with literary and artistic property is rather specific. We of course acknowledge the similarities between what are dubbed "French-style" author's rights and the Anglo-American copyright system. However, while in one case the author takes centre-stage, in the other it is the work that is in the limelight.

Article L. 122-5 of the French Intellectual Property Code is amended as follows: after the ninth paragraph, a paragraph is inserted which states that the author may not prohibit digital copies or reproductions for the purpose of mining texts and data.[12] In addition to this passage, after the fifth

practices developed by contract law. Further reading: H. Raizon, *La contractualisation du droit moral de l'auteur*, thesis, Avignon, 2014, p. 141ff., 211ff. and 229ff.

[6] "Copyright and value sharing are addressed at the European level. It is absolutely necessary to defend them. We are the country of Beaumarchais and of author's rights, let's not forget that. It is an aspect of soft power, of which we are the world leader. We must convey this message in the interest of attaining a balance", Françoise Nyssen, Emmanuel Macron's former Minister of Culture declared when she appeared before the Senate's Cultural Committee in 2018. She also referenced the European copyright reform under discussion at the time, mentioning certain measures which, in her words, "exacerbate rightsholders' opposition" to such provisions.

[7] In this regard, it is worth recalling work done in the context of the Lemaire Act in 2016, application of which was halted due to the work on the future EU directive 2019/789 of 17th April 2019.

[8] The French term "exploitation" here in the original text is intended in its original meaning of "highlighting" (the data collected), rather than to the intellectual property sense of exploitation of intellectual property rights.

[9] CPI [French Intellectual Property Code], Art. L. 122-5-3.-I. "Text and data mining, within the meaning of Article L. 122-5, 10°, means implementation of a technique for automated analysis of texts and data in digital form in order to extract information, in particular patterns, trends and correlations. "

[10] Jean-Michel Bruguière uses the terms "forage de données, prospections de données" [data mining, data prospection], "extraction d'un savoir ou d'une connaissance" [extraction of knowledge], J.-M. BRUGUIERE, *Le droit du copyright anglo-américain*, Dalloz, Connaissance du droit, 2017, p.193.

[11] N. JONDET, "L'exception pour le data mining dans le projet de directive sur le droit d'auteur- Pourquoi l'Union européenne doit aller plus loin que les législations des États membres", *Propr. Intel.*, Apr. 2018, n°67, Doctrine.

[12] CPI, Art. L. 122-5-I "10° Digital copies or reproductions of a work for the purpose of text and data mining under the conditions provided for in Article L. 122-5-3; (...)

paragraph of Article L. 342-3 of the same Code, a paragraph is inserted, authorising users to store copies of works that are mined, provided that access thereto has been lawful.[13] Reading these two articles of the Intellectual Property Code would seem to suggest that the author's economic rights are infringed rather than his or her moral rights. However, as we shall see below, the author's moral rights are likewise affected by text and data mining.

Certainly, the exception (concerning the right) to text and data mining raises issues for both the author's moral and economic rights. Moreover, the notion of data[14] is of interest here, as it is sufficiently broad to involve the work as a whole and the various components from which it is formed, in all their particularity. This presupposes an equally broad contractual approach to the issue. That also applies to exploitation of rights through collective management.

Collective Management. Collective management of author's rights is undoubtedly the most widely practised, if not the most representative, form in which the author's economic rights are exercised. In this light, one might consider that collective management organisations would be the most vigilant with regard to the text and data mining exception. Working on this assumption, the first question to address would be which specific aspect of this exception would give rise to misgivings on their part? In the light of concerns about the author's economic rights, it would be fair to say that collective management organisations may have objections concerning the texts used. Is it possible that the result of the mining would be a new work or would it simply be an item not eligible for copyright protection?

Finally, whether one considers the author's rights, the general public's rights, not to mention the interests of research, it should be borne in mind that each stakeholder always defends their own interests. In tangible terms, what can we learn from this (new) exception to copyright? Answering this question will lead us to examine the text of the directive in terms of the positions adopted by its supporters (I) and by its critics (II).

I. THE RIGHT TO TEXT AND DATA MINING AS AN ALLY OF RESEARCH AND INNOVATION

Ordinance No. 2021-1518. In French law, the Directive on Copyright in the Digital Single Market has taken the form of Ordinance n°2021-1518 in order to integrate it into the Intellectual Property Code. Amendment of Article L. 122-5 of the Intellectual Property Code cited above reveals an extension of the research and teaching exception (A); an extension that may have been welcomed by practitioners but which erodes the scope of national legislation (B).

A. A Broader Text and Data Mining Exception in the Interest of Science

The exceptions listed in this Article shall not conflict with normal exploitation of the work and shall not unreasonably prejudice the author's legitimate interests.

The detailed rules for application of this article (...) shall be specified by decree in the Council of State."

[13] CPI, Art. L. 342-3 para. 6 "5° Digital copies or reproductions of the database made by a person who has lawful access to it, with a view to text and data mining in a research context, excluding any commercial purpose. Retention and communication of technical copies resulting from processing, at the end of the research activities for which they were produced, shall be ensured by bodies designated by decree."

[14] Data is information. The most familiar category of data is personal data, the fate of which is governed by the equally well-known GDPR. In the case considered here, we consider that any information concerning the identification or production of the work of art falls within the scope of TDM.

Research for Everyone and by Everyone. In the world of science, it is not uncommon for work that has a global impact to come into being thanks (in part) to work (or research) by previous scientists that they did not complete or conducted differently. The idea of being able to use the groundwork done by others is therefore not extraordinary or immoral. Positive feedback for the Directive on Copyright in the Digital Single Market comes mainly from scope to bring more researchers together. In addition, making science attractive and comprehensible to the general public gives rise to the notion of popularisation as a herald-bearer of contemporary science. The idea of (more or less) free access to the work of other researchers is invigorating for the market from an economic perspective. Thinking only in terms of profit and scientific progress, the pros seem more visible than the cons. However, getting more and more people on board to push a project forward is not so simple.

Circular Search. To a certain extent, TDM is reminiscent of the blockchain phenomenon. The latter consists of networking information, offering members of the network an opportunity to simultaneously share information and use it immediately.[15] Instantaneous sharing is not the goal that the Directive promotes, but opening up sharing to as many people as possible seems intended to place individuals within a circle and on the same level. Without being categorical, a distinction can be drawn between the text or data mining exception aimed at researchers in a predefined field[16] and the form of the exception directed to the general public[17] in a more global sense. Knowledge circulation ultimately occurs in a system similar to that of the aforementioned network, albeit not simultaneously: members of a community are brought together. Data represent the known foundation of a body of knowledge, so does sharing data presume that they are transformed or that their original state is maintained? TDM advocates protecting data from modification, yet the idea of innovation may undermine this conception. The phenomenon of innovation leads to transformation of the raw material.

EU legislation stipulates that rightsholders are responsible for implementing "proportionate" protection measures to ensure the security and integrity of the networks and databases on which their works are hosted.[18] This provision has two practical consequences. Firstly, the protection that the author may deploy against TDM is upstream protection. The logical consequence is that, as with all matters relating to exploitation of rights, copyright holders must ensure the security of any hosting structure that could be envisaged in the state of the art. Of course, the technical measures permitted are those that are effective[19] and do not interfere with users' rights.[20] The second consequence is that the data mined must not be altered.[21]In this sense, the Directive provides a neutral zone for the protagonists involved and offers realistic prospects of achieving the goal of popularising scientific discoveries.

---

[15] Rapport de la mission d'information commune sur la blockchain (chaîne de blocs) et ses usages: un enjeu de souveraineté, Dec. 2018.

[16] *Directives 2019/790 et 2019/789 sur le droit d'auteur dans le marché unique numérique Commentaire article par article*, edited by N. BINCTIN and X. Près, Bruylant, 2021, *esp.* p. 28.

[17] *Directives 2019/790 et 2019/789 sur le droit d'auteur dans le marché unique numérique Commentaire article par article*, edited by N. BINCTIN and X. Près, *op. cit. esp.* p. 32.

[18] CPI, Art. L. 122-5-3 II 4°.

[19] The effectiveness of the technological measures studied lies in the fact that the rightsholder genuinely controls use of the work. A. Lucas, A. Lucas-Schloetter and C. Bernault, *Traité de la propriété littéraire et artistique*, 5th ed., LexisNexis, 2017, n°1134.

[20] CPI, Art. L331-5 to L331-11.

[21] It is intriguing to note that it is possible to drill down, dismember a work of the mind (to extract data) and to insist that the work of the mind retain its integrity (by prohibiting alteration of the data). This point will be discussed below.

Popularisation and Scientific Progress. Popularisation entails being able to make the content of research and scientific results accessible to non-experts. It involves making scientific knowledge accessible and adapting it so that any audience outside the field can understand its quintessence. There is a paradox here however, since the aim of industrial property, in particular, is to protect scientific research, inter alia through patents. On the other hand, the intellectual property system is intended to encourage creation. The TDM exception is therefore a necessity of our era.[22] Making it possible for the general public to understand insights produced by science helps explain the need for scientific research and to promote it. The law, which is itself a response adapted to the society for which it is enacted, does not disregard socio-economic concerns. The process of popularisation of science leads to the emergence of research models and the definition of TDM is certainly compatible with this.

Modelling Science. The notion of originality[23] is indubitably still high on the agenda when it comes to the criteria that render works of the mind eligible for protection. The possibilities opened up by TDM or its exploratory method may well revive this (endless) debate. For example, on the basis of the analysis upheld by various courts in their rulings, Olivier Laligant has considered whether novelty should be maintained as an aspect determining originality.[24] Copyright practitioners and mainstream scholarship would fiercely refute such a position. Nevertheless, originality is generally assessed in relation to what is unknown, what has not yet been done, what has not been seen. This gives credence to the idea of considering, even partially, the idea of novelty in the quest to determine originality, without however encouraging any confusion between the two terms.

With regard to the Directive, the proposed exploratory method presupposes a mechanism that would be repeated; the only fluctuating criteria would be the subject-matter of the text or data mining, together with the author. Taking this as the point of departure, could upstream repetition affect the "original" nature of the work of the mind and render its protection questionable? This question arises as a function of the legal categorisation accorded to the result of the text and data mining. If the result of the mining is *merely a* scientific *discovery*,[25] modelling is a tool that simply facilitates this undertaking. However, if the result of the text or data mining were to be categorised as a work of the mind, would automating the search process make the ideas of originality and novelty inappropriate to categorise what would be created? This, in our view, is a genuine dilemma. By analogy with the system applied to composite works and the *sui generis* database provisions, the idea of an original work could undoubtedly be upheld because what is created bears the stamp of the

[22] A. Lucas, A. Lucas-Schloetter and C. Bernault, *Traité de la propriété littéraire et artistique*, 5th ed. , *op. cit*.

[23] For Olivier Laligant, the originality of a work lies in the fact that its author has put his or her stamp on it (his or her personality) and, above all, has given it "sufficient" individuality. This should be distinguished from the notion of novelty. O. Laligant, *La véritable condition d'application du droit d'auteur: originalité ou création ?* PUAM, 1999, n°161.

[24] *Ibid*.

[25] CA Paris, 2e , Grotte Chauvet case, 6 Feb 2015. In this case, the inventors were denied the status of being the first to disclose the cave paintings in the cave discovered. To begin with, the term inventor is interesting because it refers, in property law, to the discoverer of a treasure and in industrial property law to the person who has made an invention. In other words, in terms of literary and artistic property law, this does not constitute the creation of a work of the mind. Secondly, the Court substantiated its position by recalling that the discovery demonstrated that many people spent time at the site during the period when the paintings were made, thus refuting the idea that the inventors (Chauvet and allies) were the first to make these paintings public. Hence the rejection of the appellants' claims.

Finally, the reason why we reference this decision is because of the strict definition of the term "discovery". It is about bringing to light what was hidden, ignored. Literary and artistic property is a creative process, bringing out of nothing what does not exist rather than bringing to light something that has been lost or concealed.

author's personality. However, the Chauvet cave case shows that French case law considers that works of the mind must fulfil at least three criteria: originality, novelty and first publication (or disclosure). This complicates the situation and highlights the way in which the scope of national legislation has shrunk in the face of Community law.

B. A Text and Data Mining Exception Beyond the Scope of National Legislation

It should be noted that the directive is an example of the to and fro between strengthening and weakening the protection afforded to creators of works of the mind. Collective management of copyright appears to be something of a hybrid creature. Should we now simply accept that everything that is not contractually authorised is, potentially, outside the scope of protection afforded by collective management? On the one hand, the Community legislator offers a carrot by encouraging use of CMOs through granting of licences and reaffirmation of fairer remuneration for authors.[26] The ease of implementation of such practices chimes with the development and operation of CMOs. On the other hand, the legislation applies the stick by insisting on users' rights to such an extent that it arouses misgivings among rightsholders. However, using the stick can backfire as rightsholders can renegotiate the terms of use for their protected works.[27]

Some might however simply point out no one could really object to this approach if a strict stance is adopted when interpreting assignment of author's rights. Such an approach would nonetheless be bothersome, for although the co-contractor only has a say in the realm that falls within the scope of contractual provisions, the exception contained in the TDM Directive seems to ignore two important aspects. Firstly, the exception is addressed to the author; should the conclusion therefore be drawn that the author loses any right that he or she has not included within the scope of contractual provisions or assigned to a collective management organisation? Secondly, not being able to claim a right does not mean that exploitation, or even use, of such a right is possible in violation of the rights acquired through a contract or through collective management.

At this stage, the Directive places the protagonists in a deadlock. The impression arises that the TDM provisions would tend to establish a right, a subjective (full and complete) privilege rather than a (one-off) exception. Insistence that this is an exception and extension of the Directive's provisions perform a curious balancing act between the author's right to a monopoly and the exception to that monopoly.

Residual Content Within the Scope of National Legislation. National legislation deals with subject-matter that does not fall within the scope of European law and stipulates provisions that are not addressed at the European level. That is the principle as it should be applied. Is it possible in this case to state that this process has, objectively, been established? It is difficult to say, as national legislation is responsible for stipulating certain exceptions and their implementation, yet impeding TDM does not seem entirely straightforward.

The text and data mining exception is in the spotlight precisely because the Directive on Copyright in the Digital Single Market seems to accord it a degree of importance that would tend to foster broader application of a waiver to the monopoly. Before the Community legislator took a position, rightsholders were proposing introduction of an exception for the benefit of digital protagonists. This exception would comply with contractual constraints in order to preserve the exclusive rights related

[26] S. Carré, "Gestion collective: regards au prisme de la DAMUN", *in* RFPI, special issue, Dec. 2021, p. 47.

[27] A.-E. KAHN, "Les exceptions de fouille de textes et de données dans la directive 2019/790 du 17 avril 2019 : la fragilité d'un équilibre apparent", RFPI, special issue, Dec. 2021, p. 42.

to such provisions, while allowing innovation to flourish through a system of self-regulation. The wording of the TDM Directive demonstrates that this approach did not prevail, thus determining the national legislator's room for manoeuvre. That gives rise to concerns that a new method would erode the author's exclusive rights.

II. THE "RIGHT TO TEXT AND DATA MINING", A NEW MANIFESTATION OF THE EROSION OF AUTHOR'S RIGHTS

The European legislator seems to be clear about asserting that TDM constitutes an exception. However, the way in which it is presented and TDM's reception by users suggest that a new *sui generis* right could take shape through a process of transformation not covered by the law.

The exceptions stipulated certainly make it possible to constrain the monopoly conferred by copyright and render it less overwhelming. However, provisions on literary and artistic property reveal a rather conflict-ridden relationship between these two sides of the same legal framework.[28] The upheaval sparked by the TDM provision therefore affords an opportunity to analyse its impact on the author's moral rights (A) and the tendencies that can be detected in the context of economic rights (B).

A. Moral Rights, A Giant with Clay Feet

A Novel Work, a New Work. The idea of creating a work of the mind underpinned by the text and data mining exception provokes a sense of déjà vu. Just think of the controversy surrounding artificial intelligence that is still disrupting the world of intellectual property.[29] The distinction between robots and artificial intelligence makes it possible to state that creation by human action has occurred as a result of controlling the machine's actions or conversely to assert that the machine is autonomous during the process of creation. Legal scholarship takes the view that the author of a work of the mind cannot be a machine or an animal because they do not fulfil the first criterion of being a legal person. However, if we reconsider automated text and data mining, how would the result obtained be categorised in legal terms? Firstly, if we consider that this is only the groundwork for creation of a work of the mind, that would imply that a new work does not come into being solely on the basis of text and data mining. The automated nature of the mining process would support this argument. Secondly, if we consider that the mining mechanism, despite its automated aspect, also entails a form of originality because of the reading grid 'created' by the author of the text or data mining process, could we affirm that a work of the mind has come into being? If so, would it be a composite work? That is a legitimate question, as the work would be created from components that exist in other works. However, denying any rights to the creators of previous works excludes the possibility of classifying this as a composite work. In our view, there are two inconsistencies in this approach.

The first is financial. Greater protection for remuneration of authors of works of the mind cannot be asserted on the one hand while denying them any rights due to an exception with unprecedented scope, given the number of works of the mind that could be involved in each mining process. It should be recalled that text and data mining processes may be repetitive, a point on which the

[28] For further analysis: D. Piatek, *La crise des exceptions en droit d'auteur: étude paradigmatique*, thesis, Orléans, 2016, 540 p.

[29] On the subject of artificial intelligence: - Report by C. Villani, "Donner un sens à l'intelligence artificielle", 28th March 2018, https://www.aiforhumanity.fr/. - Report of the CSPLA Mission, "Intelligence artificielle et Culture", A. BENSAMOUN and J. FARCHY, 27th Jan. 2020. *esp.* p. 138 Quelle éthique de l'IA.

Directive remains enigmatic, merely authorising mining if it is conducted with respect for lawful use provisions.[30] This recalls the origin of the extension, namely academic research. However, asserting that text and data mining is conducted purely for the good of society, or even for altruistic ends, with no personal gain intended, would be a somewhat partial view.

The second inconsistency arises from the moral right, which is precisely the issue addressed here. The status of the work and its creator's wishes seem to be completely disregarded in favour of the wishes of the individual engaging in text or data mining. This is strange, given that previous exceptions and the general exercise of any right must respect such considerations. Although the legislative texts recall the respect due to the work and its creator, implementation of the exception does give rise to some risks. Compilation of works of the mind turns the spotlight on this independence of the individual engaging in the text or data mining, which in our view is ambiguous.

Compilation of Works of the Mind, the Jean Ferrat Case. By way of illustration, the case of artist Jean Ferrat[31] is one example of exercise of the author's moral rights. In this case, the singer was able to challenge a catalogue of songs because of his personal history and political beliefs. In the TDM era, would he still be able to raise the same arguments and win his case? The lack of scope for an author to assert their rights because text and data mining is an "automated" process could prevent the author from mounting any defence. Respect for the work and its integrity may be affected. The dense nature of the mining process may pose a problem. Rightsholders could be completely unaware of the use that is made of their rights. Although the obligation to respect the conditions of use is specified and although access to the works mined must be lawful, the Directive does not offer sufficiently convincing scope to assert moral rights. The advantage granted to researchers seems to take very little account of the wishes of creators of works of the mind.

Copyright and CMOs in the Face of TDM. Another disadvantage of automation directly affects CMOs. Two scenarios could be cited. In the first scenario, the CMO owns the rights to the work that is mined. It seems that in this configuration, it would be up to the CMO to secure the work. However, the Directive on Copyright in the Digital Single Market contains aspects that may make it difficult to understand the CMO's right of action. In an infringement case, the court before which the case was brought asked the CMO to demonstrate ownership of the rights held in the work it claimed.[32] By placing this aspect at the heart of the TDM system, it becomes clear that such reference to ownership of rights can lead to a lack of clarity as to who is responsible for protecting the work of the mind. The second scenario is based on the idea that the intellectual work being mined is currently or has been in the CMO's catalogue. Scope for copyright holders to renegotiate the terms of use of works of the mind can, in practice, lead to confusion among users. Some who believe they are availing themselves lawfully of the exception may in fact be infringing the rights of copyright holders. Access to the work is determined by the will of the rightsholders, which is bound to have an impact on the mining process.

[30] The CPI does not mention the issue of repetition of text or data mining any more than the Directive does (Art. L. 122-5-3.-I). This omission may be justified as provisions authorising storage are included.

[31] Court of Cassation (Social Division), 10th July 2002, B., V, n° 241, Jean Ferrat case. If this ruling is transposed to the present case, there is some doubt as to whether the author would be able to defend the right to respect for his work.

[32] J. Daleau, "Recevabilité de l'action en contrefaçon d'une société de gestion collective", *D.* IP/IT, 20th Apr. 2021. In this article, the author analyses a decision (Civ. 1st , 11th March 2020, FS-P+B, n° 19-11.532) in which the Court obliged a collecting society that wished to bring an infringement case to furnish proof concerning the rights it intended to claim.

Data Mining. TDM entails drilling down inside a work of the mind to extract data. We have noted that data are pieces of information collected, produced and used in the context of a research project (particularly in the world of science). That means that the aspects explored encompass texts, images, sounds, etc. Anything that can ultimately be considered a work of the mind can potentially be "mined". This also assumes that it is possible to reproduce the work in a digital format.[33] That means that a work that cannot be digitised cannot be mined.[34] The data extracted is not extracted for the sake of preservation of such data. Firstly, research progress suggest that the data may potentially be transformed. That route is opened up by use of the data collected. Furthermore, the extent to which respect is shown towards the work may be questionable as it may be 'dismantled' in the mining process. Highlighting the fragility of moral rights' contemporary foundations cannot conceal the impact of TDM on the author's economic rights.

There is a genuine link between moral and economic rights, irrespective of whether the economic rights are held by the author or by other parties. TDM keeps rightsholders at arm's length by refusing additional remuneration as a consequence of text and data mining.

B. Economic Rights, Defeated without a Battle?

Social Interest. Starting from a broad definition, it can be said that text and data mining involves

"exploring, without a priori, a very large number of data (images, texts, sounds). This massive and automated exploration subsequently makes it possible to highlight correlations within these data pools, making it possible to identify and understand links between apparently distinct phenomena, to devise, confirm or invalidate scientific hypotheses, to acquire new knowledge, to open up new avenues or to anticipate trends".[35]

This method is clearly geared towards what the general public and, more broadly, society as a whole can gain from the creative work of others. The diverse origins of works can be a problem. Confusion may also arise due to the use of algorithms. In other words, protected works could very well exist that are separated by a thin line from created works not protected by author's rights provisions.[36] In this respect, stipulating a uniform system is complicated by the interests of the general public, which appear to be more important to the legislator than rightsholders' interests. Generally speaking, a work that is not protected by copyright can be understood by means of other material; unless it is an object without an owner. There is potential for clashes concerning use and compilation of a large number of works – all the more so as economic concerns are the driving force behind research, even when it is conducted for noble ends.

Economic Antagonism. Law concerning literary and artistic property fosters creation and economic antagonism through its sincere affirmation of the wish that creators of works of the mind should be able to live from the fruits of what they have created. It provides very fertile ground for competition

[33] Directive on Copyright in the Digital Single Market, Recital. 3, CPI Art. L. 122-5-I. Text and data mining is directly dependent on the possibility of digitising intellectual works.

[34] The TDM exception could hardly be applied to what cannot be seen, heard or even touched. The sculpture *Io sono* by Italian artist Salvatore Garau, sold for €148,020 in 2021, is an example of an invisible work that it would be virtually impossible to mine.

[35] A.-E. KAHN, "Les exceptions de fouille de textes et de données dans la directive 2019/790 du 17 avril 2019 : la fragilité d'un équilibre apparent", RFPI, special issue, Dec. 2021, p. 39 .

[36] Directive on Copyright in the Digital Single Market, Recital 9. The Directive simply rules out control by means of copyright, but does not consider the possibility that another field (e.g. contract law) may exercise extensive control over the material mined, irrespective of its legal category.

and presents itself as the guardian of the wealth generated by works of the mind. However, with TDM, Community legislation takes an unexpected turn for authors and rightsholders; the law seems to move from being their protector to a stance that is hostile towards them. Market competitiveness is the justification presented.

Defence of Economic Rights Over and Above Moral Rights. Exploring data can be defended as a process that does not alter the work. In practice, this can give rise to difficulties. First of all, the data that is explored is usually stored, even if only temporarily, for the purposes of such study.[37] It should be recalled that mining serves a purpose and is not an end in itself. That is the example we noted concerning artificial intelligence: before it is able to produce, it feeds itself by storing information selected via algorithms. Subsequently, if the works thus explored and stored have open-access status, respect for intellectual property rights must nonetheless still be considered, at least from the point of view of database producers, for example. However, if the works are not open access, from the point of view of the collective management organisations, would this step not correspond to use in the meaning of the rights transferred to the CMO with regard to the works? The phenomenon is not a textbook case since lawful access by one person may be used by several people (single subscription and multiple accounts). For example, the notions of household and family circle are broad enough to make it difficult to detect fraud. If research is combined with a race to produce, it seems possible that CMO rights may be violated. The lack of precision in the legislative text affords room for doubt. A competitive spirit gives rise to this type of use.

Competition and Intellectual Property Rights. Although advocates of the directive emphasise scientific progress, it is also likewise clear that the legislator is influenced by international competition on all fronts. Europe wishes to be competitive and thus prioritises scientific research. This directive, which lies at the heart of many copyright disputes, also relates to competition law. Competition law defends liberalism, which can put it at odds with intellectual property law. However, the two areas are similar in one respect: seeking to strike a (market) balance. Whenever the product of text and data mining has economic value, competition and free riding issues will take centre-stage. Praise for scientific innovation does not mean that science lacks a monetary dimension.

The text and data mining exception is clearly changing the rules, shifting the dividing lines between the protagonists. In this respect, the work of the mind and its creator seem to find a scrap of salvation only in the context of exploitation of the work. Compliance with intellectual property provisions cannot run counter to ordinary contract law. We would therefore contend that TDM is counterbalanced by the rules of contract law because the decisive issue is access to works. Since this is the opening that the Directive leaves to rightsholders, the battle to be waged will unfurl largely in terms of technical protection measures. The key will thus be to highlight the opt-out that the Directive permits for rightsholders.[38]

Conclusion:

[37] This storage may prove to be a source of conflict in the light of sharing and framing. Framing is a "technique involving dividing a page on a website into several frames and displaying in one of them, by means of a clickable link or embedded web link, an element from another site in order to hide from the users of that site the original environment to which that element belongs", https://www.gazette-du-palais.fr/actualites-juridiques/jur-cjue-la-technique-de-la-transclusion-et-le-droit-dauteur/.

[38] Directive on Copyright in the Digital Single Market, Art. 4.3 and Recital 18. The reservation must be express and must not hinder scientific research or the exception concerning temporary reproduction. It should be noted, however, that applying these provisions is not necessarily straightforward.

There is no doubt that text and data mining delights economic protagonists and receives the blessings of advocates of social solidarity. When it comes to literary and artistic property law, this exception does not present a united front. The dual nature of copyright is once again highlighted. The author must be able to assert his or her rights and live, if they so wish, from their creative work. However, drawing a livelihood from work one has created means sharing it with the public to some extent. In practice, it is fair to say that there is some clumsiness in the exception for text and data mining: under the guise of ensuring the greatest number of people benefit from it, individuals engaging in text or data mining ultimately benefit the most from this measure, to the detriment of the general public (for nothing obliges those doing the mining to publish their results) and authors (the law puts those engaging in text or data mining beyond the reach of any claim).

In the words of Anne-Emmanuelle Kahn, the text and data mining exception presents an "apparent balance" that is rather fragile.[39] European legislation has paved the way, but not always for the better. Stating that the work takes precedence over everything could thus lead one to conclude that the end (perhaps) justifies the methods used to create it. That would delight researchers and leave rightsholders feeling cheated. We are aware that the aim is not to foster such a caricatural approach. On closer inspection, European legislation is granting a favour that can be withdrawn even before it reaches the recipients. The Directive, through the reservation or opt-out granted to rightsholders, gives with one hand and takes away with the other.[40] Ultimately, the text and data mining exception seems to leave a bitter-sweet odour in its wake; perhaps future developments will make it sweeter.

(English translation by Helen FERGUSON)

[39] A.-E. KAHN, "Les exceptions de fouille de textes et de données dans la directive 2019/790 du 17 avril 2019 : la fragilité d'un équilibre apparent", RFPI, special issue, Dec. 2021, p. 39.

[40] A.-E. KAHN, "Les exceptions de fouille de textes et de données dans la directive 2019/790 du 17 avril 2019 : la fragilité d'un équilibre apparent", RFPI, special issue, Dec. 2021, p. 45.