
TRADUÇÃO

Os artilheiros que estão faltando no time da inteligência artificial: Big Data, o direito fundamental à pesquisa, e as (insuficientes) limitações para mineração de textos e dados na Diretiva (UE) 2019/790 relativa aos direitos de autor no mercado único digital.

Christophe Geiger

cgeiger@luiss.it

Professor de direito da Universidade Luiss
Guido Carli, em Roma.

Os artilheiros que estão faltando no time da inteligência artificial: Big Data, o direito fundamental à pesquisa, e as (insuficientes) limitações para mineração de textos e dados na Diretiva (UE) 2019/790 relativa aos direitos de autor no mercado único digital.

Palavras-chave

direitos autorais

inteligência artificial

“big data”

direito à pesquisa

limitações à mineração de textos e de dados

Resumo

Este artigo propõe que as recentes estratégias da União Europeia no campo da Inteligência Artificial (IA) se assemelham a um time de futebol que não dispõe de um artilheiro para vencer qualquer uma das competições com outras jurisdições que possuem limitações mais flexíveis aos direitos autorais, em particular aquelas que permitem uma robusta mineração de textos e dados (TDM). Analisam-se aqui as limitações de TDM recentemente introduzidas na legislação europeia em direitos autorais pela diretiva relativa aos direitos de autor e direitos conexos no mercado único digital, de modo a demonstrar que essas disposições não apenas falham ao levar em conta o direito à pesquisa alicerçado no direito fundamental à informação, como também não permitirão que a União Europeia crie um ambiente competitivo para o desenvolvimento de inovações em IA e aquelas intensivas em dados. O artigo conclui pela necessidade de uma revisão imediata da estrutura dos direitos de autor para as atividades de TDM na UE e, no plano internacional, da implementação da diretiva pelos Estados-Membros que esteja em conformidade com o arcabouço dos direitos fundamentais da UE e com o objetivo estabelecido pelos legisladores europeus.

Em 19 de fevereiro de 2020, a Comissão Europeia anunciou uma ambiciosa estratégia digital para a União Europeia, definindo, em duas comunicações, os objetivos a serem atingidos – uma sobre o futuro digital da Europa¹ e outra sobre dados², ambos complementados pelo *White Paper* sobre inteligência artificial³. Ao ler estes documentos estratégicos, seria de se esperar que fosse finalmente elaborado um quadro jurídico sólido e ponderado para o ambiente digital e que as políticas da UE nesta área constituíssem, no futuro, mais do que uma simples colcha de retalhos de interesses setoriais. Com efeito, no que diz respeito à revolução que a tecnologia digital trouxe a muitos setores da UE, lê-se:

Esta profunda transformação da sociedade apela a uma reflexão de fundo a todos os níveis sobre a melhor forma de a Europa responder a estes riscos e desafios.

As dificuldades serão enormes, mas a Europa dispõe, inquestionavelmente, dos meios necessários para concretizar um futuro digital melhor para todos.⁴

Ou: “Criar uma Europa preparada para a era digital é um quebra-cabeças complexo, cujas diversas peças estão interligadas; tal como em qualquer quebra-cabeças, só é possível ter uma visão do conjunto uma vez inseridas todas as peças”⁵. A vontade de finalmente modernizar e adaptar o quadro jurídico da propriedade intelectual (PI) aos desafios colocados pelo ambiente digital foi reiterada no plano de ação da Comissão Europeia sobre a PI, publicado no final de novembro de 2020, onde se destacou que:

A revolução tecnológica – a economia e a sociedade dos dados, a transição para a inteligência artificial (IA), a importância crescente das novas tecnologias como

a tecnologia das cadeias de blocos, a impressão 3D e a Internet das Coisas, bem como o desenvolvimento de novos modelos de negócios como a economia das plataformas e a economia de dados e a economia circular – oferece uma janela de oportunidade única para modernizar a nossa abordagem de proteção dos nossos ativos incorpóreos.⁶

Em resumo: uma reflexão horizontal sobre as questões digitais é desejada pela Comissão Europeia antes de se (re)definir as regras do jogo por meio de uma legislação adequada. No entanto, em um exame mais atento, nem no plano de ação proposto, nem na agenda de implementação para essas estratégias, aparece qualquer questionamento sobre os espaços livres deixados pela legislação em PI para permitir o desenvolvimento de um ecossistema digital equilibrado na UE⁷. A questão também foi totalmente ignorada pela Proposta de Regulamento do Parlamento Europeu e do Conselho que estabelece regras harmonizadas em matéria de inteligência artificial (regulamento inteligência artificial) e altera determinados atos legislativos da União, publicada em 21 de Abril de 2021, apesar do fato de um dos principais objetivos propostos pela Comissão Europeia ser o de “garantir a segurança jurídica para facilitar os investimentos e a inovação no domínio da IA”⁸. A menos que a Comissão considere que a questão já foi resolvida, nomeadamente pela diretiva de 17 de abril de 2019 relativa aos direitos de autor e direitos conexos no mercado único digital (Diretiva 2019/790)⁹, é difícil entender como uma reflexão horizontal sobre inovação digital pode ser conduzida sem incluir grande parte do direito digital, como o direito autoral (incluindo suas limitações e exceções) e a propriedade intelectual em geral¹⁰. Ainda pior, é provável que as soluções que foram adotadas anteriormente na Diretiva 2019/790 estejam

em contradição direta com os objetivos agora apresentados, e que, como resultado, as muitas ambições recentemente apresentadas provavelmente permaneçam letra morta.

Muitos exemplos poderiam ser fornecidos, mas gostaríamos de destacar apenas um nesta contribuição: uma política proativa no campo da inteligência artificial e da inovação digital exige, sem dúvida, propor um quadro jurídico moderno que reconheça o risco de obstáculos ou proibições injustificados para o desenvolvimento da inteligência artificial na União Europeia¹¹. Nesse contexto, sabe-se que para permitir o aprendizado de máquina, essencial à inteligência artificial, é necessário ter exceções robustas para mineração de textos e dados para que a máquina possa reproduzir, armazenar e processar dados existentes e propor novas soluções¹². A questão também é bastante atual no campo da saúde pública, porque é óbvio que grandes estoques de dados sobre a COVID-19 precisam ser processados e analisados em nível global para permitir o desenvolvimento de novos tratamentos¹³.

A mineração de textos e dados tornou-se, de fato, uma ferramenta fundamental para a pesquisa, seja aquela realizada por órgãos públicos ou por agentes privados¹⁴. Contudo, como foi argumentado por Bernt Hugenholtz e outros¹⁵, a solução proposta pela diretiva de 2019 é amplamente inadequada; é muito restritiva no que diz respeito às exceções para mineração de textos e dados. Assim, para usar uma analogia “futebolística”, a estratégia de Inteligência Artificial da União Europeia pode se assemelhar a um time de futebol que ficaria sem atacantes para marcar gols com sucesso e vencer qualquer uma das competições com outras jurisdições que possam ter a vantagem de disposições legais mais flexíveis, que permitam atividades mais abrangentes de TDM.

Por que este é o caso? Voltemos rapidamente às duas exceções introduzidas na legislação europeia pela diretiva de 2019.

Em primeiro lugar, deve-se notar que a Comissão Europeia claramente não entendeu, no início, a importância das exceções de mineração de textos e dados para o desenvolvimento da inteligência artificial. É certo que a proposta de diretiva de 14 de setembro de 2016¹⁶ previu uma exceção mandatória¹⁷, permitindo a mineração de textos e de dados, que não poderia ser afastada por contratos¹⁸. Contudo, foi limitada às “[...] reproduções e extrações efetuadas por organismos de investigação para a realização de prospecção de textos e dados de obras ou outro material protegido a que tenham acesso lícito para efeitos de investigação científica”¹⁹. Essas organizações beneficiárias foram compreendidas de forma restritiva, uma vez que o Considerando II especificou que:

[...] os organismos de investigação dos Estados-Membros têm geralmente em comum o fato de agirem sem fins lucrativos ou no âmbito de uma missão de interesse público reconhecida pelo Estado. Esta missão de interesse público pode refletir-se, por exemplo, no financiamento público, em disposições da legislação nacional ou em contratos públicos²⁰.

Isso abrange apenas organizações de pesquisa e universidades públicas (ou aquelas que realizam uma missão de serviço público).

Mas o que dizer das importantes atividades de pesquisa conduzidas por *startups* que operam no ambiente digital, que são a fonte de importantes inovações, particularmente no domínio da inteligência artificial, e cujo potencial de crescimento foi fortemente visado na nova estratégia da Comissão?²¹ As *startups* não foram consideradas e, portanto, suas atividades de mineração de dados permanecem sujeitas aos direitos exclusivos. Em linhas semelhantes, o que dizer das parcerias público-privadas tão incentivadas na apresentação de candidaturas

em concursos para a União Europeia? É certo que o considerando 10 afirma, de forma relativamente vaga, que “os organismos de investigação devem também beneficiar da exceção ao participarem em parcerias público-privadas”, sem, contudo, explicar como isso funcionaria. Isso sem contar os pesquisadores individuais sem vínculo com uma instituição, cujas atividades também não são contempladas, e os jornalistas que, por definição, trabalham para estruturas privadas, que também não se beneficiarão da exceção.

Esta situação é altamente problemática em termos de direitos fundamentais e, em particular, no que diz respeito ao direito à pesquisa. De fato, o direito à pesquisa tem uma forte base nos direitos humanos e é protegido em âmbito internacional, europeu e nacional²². Baseado no direito à informação, inclui o direito ativo de buscar informações efetivas e objetivas por meio do uso de fontes existentes, o que implica na capacidade de utilizar legalmente, no ambiente digital, técnicas de mineração de textos e dados para a realização de pesquisas²³. Para os jornalistas, é de crucial importância, pois certas investigações só são possíveis hoje por meio da mineração, em larga escala, de textos e dados. A título de exemplo, há que se recordar do escândalo dos “*Panama Papers*”, que evidenciou uma evasão fiscal em larga escala de políticos, bilionários, celebridades e desportistas de alto nível, e que só pôde ser revelado através da busca e mineração de informações e de técnicas de busca automatizada, de mais de um milhão de documentos bancários de *offshores*.

Ao restringir desta forma o escopo da exceção, a Comissão correu o risco de criar uma disposição ineficaz e, portanto, rapidamente obsoleta, notadamente no que diz respeito ao desenvolvimento da inteligência artificial, mas também quanto a outras atividades de investigação e inovação essenciais e que não são conduzidas por entidades públicas. Isso é especialmente verdadeiro porque a proposta de

diretiva só permitiu que organizações de pesquisa buscassem textos e dados “sobre obras ou outros assuntos aos quais tenham acesso legal para fins de pesquisa científica”, o que parece excluir grande parte da pesquisa on-line, na medida em que a licitude das fontes permaneceria incerta²⁴.

Por todas estas razões, a proposta de diretiva foi alvo de muitas críticas durante a fase parlamentar²⁵, levando a mudanças em seu texto final²⁶. Como resultado, o círculo de beneficiários da exceção foi estendido às “Instituições responsáveis pelo patrimônio cultural”, que, de acordo com o considerando 13 da diretiva, inclui, principalmente, bibliotecas, museus e arquivos²⁷. Além disso, foi incluída a possibilidade de armazenar obras para fins de busca e mineração, o que é importante porque pesquisa leva tempo: deve, portanto, ser possível voltar às obras em que os dados estão contidos sem ter que reproduzi-las todas as vezes. Por último, foi introduzida uma nova exceção no corpo da diretiva, em que “Os Estados-Membros devem prever uma exceção ou limitação aos direitos previstos [...] a fim de permitir a utilização digital de obras e outro material protegido...”²⁸.

O objetivo da introdução de um novo dispositivo era claramente responder às críticas dirigidas ao texto inicial e permitir aos Estados-Membros prever uma exceção para os casos não abrangidos pelo artigo 3 da diretiva, nomeadamente para autorizar a mineração de textos e dados para entidades excluídas do escopo do artigo 3. Isso decorre claramente do considerando 18 da diretiva 2019/790, que afirma que:

[...] Para além da sua importância no contexto da investigação científica, as técnicas de prospecção de textos e dados são amplamente utilizadas, tanto pelas entidades públicas, como privadas para analisar grandes quantidades de dados em diferentes domínios da vida

e para vários fins, nomeadamente para serviços públicos, decisões empresariais complexas e para o desenvolvimento de novas aplicações ou tecnologia.

Essa nova disposição seria um grande avanço se a exceção não tivesse sido acompanhada de um mecanismo de “*opt-out*”, permitindo aos titulares de direitos reservar expressamente a utilização de obras “[...] de forma adequada, em particular por meio de leitura ótica no caso de conteúdos disponibilizados ao público em linha”²⁹. Ao condicionar a nova exceção à boa-fé dos titulares de direitos, corre-se o risco de inviabilizar esta exceção na prática, uma vez que a sua eficácia dependerá da implementação (ou não) do “*opt-out*”³⁰.

O objetivo do legislador foi o de levar em conta os interesses legítimos dos titulares de direitos e, em particular, dos produtores de bases de dados, também abrangidos pela exceção, que têm interesse em impedir a superexploração das suas bases de dados que seriam afetadas pela disposição³¹. No entanto, se a intenção fosse promover a pesquisa e, em particular, criar um arcabouço legal que estimulasse a inovação, seria possível sujeitar as atividades de mineração de textos e dados exploradas para fins comerciais a um direito à remuneração³². As incertezas criadas pela diretiva a esse respeito, que se repetirão na fase de transposição a nível nacional³³, podem conduzir a um grande atraso para a Europa no desenvolvimento de novas tecnologias e no campo da inteligência artificial³⁴, enquanto, em outros lugares, essas atividades com alto potencial de inovação possivelmente já estão cobertas por normas abertas, como o uso justo (“*fair use*”)³⁵, ou exceções mais flexíveis³⁶. Isso sem mencionar o fato de que o legislador europeu também achou por bem especificar que o famoso teste dos três passos é aplicável às duas exceções recém-criadas³⁷, bem como o enigmático artigo 6(4) da

diretiva InfoSoc, de 2001³⁸, que, recorde-se, visa à resolução de conflitos entre usuários e titulares de direitos em caso de implementação de “medidas técnicas de proteção” (MTP), mas que ainda hoje permanece incerto na sua compreensão e aplicação³⁹. Por último, ainda que o escopo das exceções tenha sido ampliado, muitas questões pendentes no âmbito da proposta de diretiva permanecem, no texto final, por resolver, dando origem a uma grande insegurança jurídica sobre as atividades de mineração de textos e dados, correndo, assim, o risco de prejudicar a competitividade da União Europeia como espaço de pesquisa e de criatividade no cenário mundial⁴⁰.

Em suma, se a ambição da Comissão é propor uma política proativa que posicione a União Europeia na vanguarda da inovação digital e da inteligência artificial, receia-se que tal tarefa se revele simplesmente impossível se o debate sobre o alcance das exceções de mineração de textos e dados não for reaberto, com urgência, em um futuro muito próximo⁴¹. O impulso pode vir da Organização Mundial da Propriedade Intelectual, que recentemente abriu uma grande consulta sobre a propriedade intelectual e a inteligência artificial, incluindo o tema da mineração de textos e dados⁴². Nesse contexto, estudiosos convidaram a OMPI a assumir a liderança nessa área, por exemplo, por meio de uma proposta legislativa destinada a criar uma nova exceção aos direitos autorais no modelo do Tratado de Marrakesh⁴³. Fundamentalmente, isso só será possível se repensarmos o pressuposto subjacente e, ainda, amplamente aceito, de que apenas fortes direitos exclusivos podem servir à inovação e à criatividade, e reconhecendo que limitações e exceções podem, igualmente, desempenhar essa tarefa⁴⁴.

Notas Finais

1 Comunicação da Comissão ao Parlamento Europeu, ao Conselho, ao Comitê Econômico e Social Europeu e ao Comitê das Regiões, Construir o futuro digital da Europa, COM/2020/67 final. Documento 52020DC0067. Disponível em <<https://op.europa.eu/pt/publication-detail/-/publication/db95106e-53ca-11ea-aece-01aa75ed71a1>>.

2 Comunicação da Comissão ao Parlamento Europeu, ao Conselho, ao Comitê Econômico e Social Europeu e ao Comitê das Regiões, Uma estratégia europeia para os dados, COM/2020/66 final. Documento 52020DC0066. Disponível em <<https://eur-lex.europa.eu/legal-content/PT/TXT/PDF/?uri=CELEX:52020DC0066&from=PT>>.

3 Livro Branco sobre a inteligência artificial - Uma abordagem europeia virada para a excelência e a confiança, Bruxelas, 19.2.2020 COM (2020) 65 final. Disponível em <<https://op.europa.eu/pt/publication-detail/-/publication/ac957f13-53c6-11ea-aece-01aa75ed71a1>>.

4 Comunicação da Comissão ao Parlamento Europeu, ao Conselho, ao Comitê Econômico e Social Europeu e ao Comitê das Regiões, Construir o futuro digital da Europa, *supra*, p. 1.

5 *Ibid.*, p. 3.

6 Comunicação da Comissão ao Parlamento Europeu, ao Conselho, ao Comitê Econômico e Social Europeu e ao Comitê das Regiões, Tirar pleno partido do potencial de inovação da EU: Um plano de ação em matéria de propriedade intelectual para apoiar a recuperação e resiliência da EU, COM/2020/760

final. Documento 52020DC0760, p. 2 (destacamos). Disponível em <<https://eur-lex.europa.eu/legal-content/PT/TXT/PDF/?uri=CELEX:52020DC0760&from=PT>>.

7 Para uma crítica inicial da falta de ambição do legislador da UE no campo dos direitos autorais, veja as reflexões fundamentais em: HUGENHOLTZ, P. Bernt (ed.), *The Future of Copyright in a Digital Environment*. Den Haag: Kluwer, 1996, em particular o capítulo do próprio editor, *Adapting Copyright to the Information Superhighway*, p. 81 *et seq.*

8 Proposta de Regulação do Parlamento Europeu e do Conselho que estabelece regras harmonizadas em matéria de inteligência artificial (Lei da inteligência artificial) e altera determinados atos legislativos da União, COM/2021/206 final. Documento 52021PC0206, p. 3. Disponível em <https://eur-lex.europa.eu/resource.html?uri=cellar:e0649735-a372-11eb-9585-01aa75ed71a1.0004.02/DOC_1&format=PDF>.

9 Diretiva (UE) 2019/790 do Parlamento Europeu e do Conselho, de 17 de abril de 2019, relativa aos direitos de autor e direitos conexos no mercado único digital e que altera as Diretivas 96/9/CE e 2001/29/CE, OJ L 130, 17.5.2019, p. 92–125 (diretiva 2019/790). Disponível em <<https://eur-lex.europa.eu/legal-content/PT/TXT/?uri=CELEX%3A32019L0790>>. Para comentários, veja as várias opiniões desenvolvidas pela Sociedade Europeia de Direitos Autorais (“European Copyright Society”), em <<https://europeancopyrightsociety.org/>>.

10 Sobre a importância crucial das exceções e limitações no direito autoral europeu, em particular para a contínua inovação e da criatividade, ver, *i.e.*: HUGENTHOLTZ, P. Bernt, *Fierce Creatures, Copyright Exemptions: Towards Extinction ?*, IFLA/IMPRIMATUR,

Rights, Limitations and Exceptions: Striking a Proper Balance, Amsterdam, Oct. 1997, republicado em VAVER, David (ed.), *Intellectual Property Rights: Critical Concepts in Law*, v. 2, London: Routledge, 2006, p. 231; GEIGER, Christophe, Promoting Creativity through Copyright Limitations, Reflections on the Concept of Exclusivity in Copyright Law, *Vanderbilt Journal of Entertainment & Technology Law*, v. 12, Issue 3, 2010, p. 515; GEIGER, Christophe; SCHÖNHERR, Franciska, Defining the Scope of Protection of Copyright in the EU: The Need to Reconsider the Acquis regarding Limitations and Exceptions, in: SYNODINOU, Tatiana-Eleni (ed.). *Codification of European Copyright Law, Challenges and Perspectives*. The Hague/London/New York: Kluwer Law International, 2012, p. 133, e as várias excelentes contribuições em: OKEDIJI, Ruth L. (ed.), *Copyright Law in an Age of Limitations and Exceptions*. Cambridge: CUP, 2017.

11 Sobre os desafios legais colocados pela IA na UE, veja o excelente relatório redigido por HUGENHOLTZ, P. Bernt; QUINTAIS, João Pedro; GERVAIS, Daniel para a Comissão Europeia: *IViR and JIPP, Trends and Developments in Artificial Intelligence Challenges to the Intellectual Property Rights Framework*, Final Report, Sept. 2020, disponível em <<https://ec.europa.eu/digital-single-market/en/news/trends-and-developments-artificial-intelligence-challenges-intellectual-property-rights>>, deixando de fora, contudo, do escopo da análise de “questões legais concernentes à inclusão do objeto de proteção aos sistemas de IA (e.g. para *text-and-data mining*)”.

12 Ver HACKER, Philipp, A Legal Framework for AI Training Data, *13 Law, Innovation and Technology* (2021, no prelo), demonstrando que “os dados de treinamento para a IA não desempenham apenas um papel fundamental no desenvolvimento de aplicações de IA, mas atualmente são capturados de forma inadequada pela UE”; KOP, Mauritz. The Right

to Process Data for Machine Learning Purposes in the EU, *Harvard Journal of Law & Technology*, v. 34, Digest, Spring 2021, p. 1. Sobre a importância da mineração de textos e dados para a inteligência artificial, veja, por exemplo,: GEIGER, Christophe. The Artificial Intelligence and Data-led Revolution of Copyright and its Wider Implications, *Digital Encounters*, relatório final, disponível em <https://thinkfide.com/wp-content/uploads/2021/01/Encounter_8_Report_Final.pdf>DU-CATO, Rosana; STROWEL, Alain. Limitations to text and Data Mining and Consumer Empowerment: Making the Case for a Right to ‘Machine Legibility’, *IIC*, v. 50, n. 6, 2019, p. 649; BINCTIN, Nicolas. TDM: un enjeu de l’intelligence artificielle, *RIDA*, n. 262, 2019, p. 7; CHIOU, Theodoros. Copyright lessons on Machine learning: what impact on algorithmic art? *JIPITEC*, n° 10, 2019, p. 398; BENSAMOUN, Alexandra; FARCHY, Joelle; SCHIRA, Paul-François. Intelligence artificielle et culture, relatório para a CSPLA, janeiro 2020, p. 54; GEIGER, Christophe; FROSIO, Giancarlo; Frosio; BULAYENKO, Oleksandr. Crafting a Text and Data Mining Exception for Machine Learning and Big Data in the Digital Single Market, in: SEUBA, Xavier; GEIGER, Christophe; PÉNIN, Julien (eds.). *Intellectual Property and Digital Trade in the Age of Artificial Intelligence and Big Data*. Collection CEIPI/ICTSD, Global Perspectives and Challenges for the Intellectual Property System: Geneva/Strasbourg, 2018, n. 5, p. 95.

13 Sobre a relação entre os dois tópicos, ver, em particular: HACKETT, Teresa. COVID and Copyright: The Right to Research, 17 de Agosto de 2020, <www.eifl.net>; FLYNN, Sean; GEIGER, Christophe; QUINTAIS, Joao Pedro (com a colaboração de T. Margoni, M. Sag, L. Guibault, M. Carroll), Implementing User Rights for Research in the Field of Artificial Intelligence: A Call for International Action, *EIPR*, n. 7, 2020, p. 393. Ver também Statement

on Copyright and Proposal of a Waiver from Certain Provisions of the Trade Related Aspects of Intellectual Property Rights (TRIPS) Agreement for the Prevention, Containment and Treatment of COVID-19 (IP/C/W/669), 22 March 2021, endorsed by 250 organizations and prominent researchers calling for the reduction of copyright barriers to COVID-19 prevention, containment and treatment, disponível em <<http://infojustice.org/archives/43020>>: “Em muitos países, os pesquisadores não têm os direitos necessários para usar as metodologias de pesquisa mais avançadas, como mineração de texto e dados, para ajudar a encontrar e desenvolver tratamentos para o COVID-19. De fato, o próprio vírus foi descoberto por um projeto de pesquisa de mineração de textos e dados que não seria legal em muitos países”.

14 Ver o Considerando 8, da Diretiva 2019/790, de 17 de abril de 2019, segundo o qual: “[...] há um reconhecimento generalizado de que a mineração de textos e dados pode, em particular, beneficiar a comunidade de pesquisa e, ao fazê-lo, apoiar a inovação”. Ver também MARGONI, Thomas, Text and Data Mining in Intellectual Property Law: Towards an Autonomous Classification of Computational Legal Methods, *CREATE*, working paper 01/2020: “O impacto que a TDM pode ter na ciência, artes e humanidades é inestimável. Isso porque, ao identificar as correlações e padrões que muitas vezes estão ocultos aos olhos de um observador humano devido à quantidade, complexidade ou variedade de dados pesquisados, a TDM permite a descoberta de conceitos ou a formulação de correlações que, de outra forma, permaneceriam ocultos ou não descobertos. Considerando esse ponto de vista, pode-se argumentar efetivamente que a TDM cria um novo conhecimento a partir de dados antigos”.

15 Para uma análise (crítica), ver também: HUGENHOLTZ, P. Bernt, The New Copyright Directive: Text and Data Mining (Articles 3 and 4), *Kluwer Copyright Blog*, 24 de Júlio de 2019, disponível em <<http://copyrightblog.kluweriplaw.com/2019/07/24/the-new-copyright-directive-text-and-data-mining-articles-3-and-4/>>; GEIGER, Christophe; FROSIO, Giancarlo; BULAYENKO, Oleksandr, Text and Data Mining: Articles 3 and 4 of the Directive 2019/790/EU, in: SAIZ GARCÍA, Concepción; EVANGELIO LLORCA, Raquel (eds.). *Propiedad intelectual y mercado único digital europeo*. Valencia: Tirant lo Blanch, 2019, p. 27; KOP, *supra*, The Right to Process Data for Machine Learning Purposes in the EU, p. 7.

16 Veja a Proposta de Diretiva do Parlamento Europeu e do Conselho relativa aos direitos de autor no mercado único digital, COM/2016/0593 final - 2016/0280 (COD). Disponível em <<https://eur-lex.europa.eu/legal-content/PT/TXT/PDF/?uri=CELEX:52016PC0593&from=PT>>.

17 O que, ao contrário da Diretiva de 2001, constitui uma melhora, considerando que a lista de 20 exceções e limitações *opcionais* possuía apenas um impacto limitado em termos de harmonização. Sobre essa questão, ver, por exemplo: GEIGER, Christophe; SCHÖNHERR, Franciska. The Information Society Directive (with updates from Stavroula Karapapa), in: STAMATOUDI, Irini; TORREMANS, Paul (eds.), *EU Copyright Law*. 2^a ed. Cheltenham, UK/Northampton, MA: Edward Elgar, 2021, p. 279.

18 De acordo com os Considerandos da proposta de diretiva, a exceção destinava-se a endereçar a grande insegurança jurídica relativa às atividades de mineração de textos e dados (Considerando 10), porque “a prospecção de textos e dados pode envolver atos protegidos por direitos de autor e/ou o direito sobre bases de dados *sui generis*, nomeadamente a

reprodução de obras ou outro material protegido e/ou a extração do conteúdo de uma base de dados. Quando não existe qualquer exceção ou limitação aplicável, seria exigida aos titulares de direitos uma autorização para efetuar tais atos” (Considerando 8).

19 Art. 3, §1, da Proposta.

20 Ver também o art. 2 (“Definições”) da Proposta de Diretiva: “«Organismo de investigação», uma universidade, um instituto de investigação ou qualquer outro organismo cujo principal objetivo seja a realização de investigação científica ou a realização de investigação científica e prestação de serviços de ensino: a) sem fins lucrativos ou por reinvestir a totalidade dos lucros na investigação científica; ou b) em conformidade com uma missão de interesse público reconhecida por um Estado-Membro”. Ver também art. 2º, parágrafo 3º, segundo o qual uma «Instituição responsável pelo património cultural» significa “uma biblioteca ou um museu acessíveis ao público, um arquivo ou uma instituição responsável pelo património cinematográfico ou sonoro”.

21 Ver o Livro Branco sobre a inteligência artificial - Uma abordagem europeia virada para a excelência e a confiança, Bruxelas, 19.2.2020 COM (2020) 65 final, *supra*, p. 3, segundo o qual: “A Europa está numa boa posição para beneficiar do potencial da IA, não só como utilizador, mas também como criador e produtor desta tecnologia. A Europa tem excelentes centros de investigação, *empresas em fase de arranque inovadoras*, uma posição de liderança mundial em robótica e setores da indústria transformadora e de serviços competitivos, desde o setor automóvel aos cuidados de saúde, à energia, aos serviços financeiros e à agricultura” (destacamos).

22 GEIGER, Christophe; JÜTTE, Bernd Justin. Conceptualizing the Right to Research and its Implications for Copyright Law, An International and European perspective, *American University International Law Review* 2022, v. 38 (no prelo). Segundo o art. 19 da Declaração Universal dos Direitos Humanos de 1948: “Todo ser humano tem direito à liberdade de opinião e expressão; esse direito inclui a liberdade de, sem interferência, ter opiniões e de procurar, receber e transmitir informações e ideias por quaisquer meios e independentemente de fronteiras” (destacamos). Da mesma forma, o art. 19, n. 2, do Pacto Internacional sobre Direitos Civis e Políticos, de 1966, especifica que: “Toda pessoa terá direito à liberdade de expressão; esse direito *incluirá a liberdade de procurar, receber e difundir informações e ideias de qualquer natureza, independentemente de considerações de fronteiras*, verbalmente ou por escrito, em forma impressa ou artística, ou por qualquer outro meio de sua escolha” (destacamos). Na Europa, o direito à informação é derivado do art. 10(1) da Convenção Europeia de Direitos Humanos e Art. 11 da Carta dos Direitos Fundamentais da União Europeia, e é incluído em várias constituições nacionais, tais como o art. 5(1) da Lei Básica alemã, art. 16(3) da Constituição Federal da Suíça e Art. 11 da Declaração de Direitos Humanos francesa. Na UE, um “direito à pesquisa” pode também encontrar fundamento no Art. 13 da Carta dos Direitos Fundamentais da União Europeia (liberdade das artes e das ciências), que estipula que “as artes e a pesquisa científica são livres. A liberdade acadêmica deverá ser respeitada”. Ver, por exemplo, FLYNN; GEIGER; QUINTAIS, *supra*.

23 O direito à informação consiste em uma liberdade passiva de receber informações existentes e em um direito ativo de buscar informações efetivas e objetivas pelo uso de fontes existentes. Sobre o direito fundamental à informação, em particular no contexto da lei

de direitos autorais, ver, com mais detalhes: GEIGER, Christophe. *Droit d'auteur et droit du public à l'information, approche de droit comparé*, Paris: Litec, 2004; Author's Right, Copyright and the Public's Right to Information: A Complex Relationship, in: MACMILLAN, F. (ed.): *New Directions in Copyright Law*. Cheltenham, UK: Edward Elgar, 2007, v. 5, p. 24. E, é claro: HUGENHOLTZ, Bernt P. *Auteursrecht op informatie*. Deventer: Kluwer, 1989, iniciando a discussão sobre a interface entre o direito de autor e a liberdade de informação.

24 O requisito da “legalidade da fonte” pode ser encontrado, por exemplo, na exceção francesa de mineração de texto e dados introduzida no Código de Propriedade Intelectual pela lei “para uma república digital” de 7 de outubro de 2016 (Art. 122-5 10º do CPI, e para bases de dados, Art. L. 342-3, 5). Deve-se notar, no entanto, que a Alemanha, que também introduziu tal exceção em sua lei de direitos autorais em 2017, não acrescentou esse critério adicional (UrhG, par. 60d (I)).

25 Neste sentido, ver GEIGER, Christophe; FROSIO, Giancarlo; BULAYENKO, Oleksandr. The Exception for Text and Data Mining (TDM) in: the Proposed Directive on Copyright in the Digital Single Market - Legal Aspects, Study for the Directorate-General for Internal Policies of the Union, Department of Citizens' Rights and Constitutional Affairs, European Parliament, February 2018. Para uma avaliação crítica da proposta de diretiva, ver também GEIGER, Christophe; FROSIO, Giancarlo; BULAYENKO, Oleksandr, Text and Data Mining in the Proposed Copyright Reform: Making the EU Ready for an Age of Big Data? *IIC*, v. 49, n. 7, 2018, p. 814, e dos mesmos autores: The EU Commission's Proposal to Reform Copyright Limitations: A Good but Far Too Timid Step in the Right Direction, *EIPR*, v. 40, 2018, p. 4; European Copyright Society,

General Opinion on the EU Copyright Reform Package, 24 January 2017, p. 5; HILTY, Reto M.; RICHTER, Heiko, in: HILTY, Reto M.; MOSCON, Valeria (eds.). *Modernisation of the EU Copyright Rules, Position Statement, MPI for Innovation and Competition*, Research Paper nº 17-12, p. 25 et sq; JONDET, Nicolas, 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 Etats membres, *Propr. intell.*, n. 67, 2018, p. 25.

26 Para comentários sobre as disposições definitivas da diretiva, ver: GEIGER; FROSIO; BULAYENKO, Text and Data Mining: Articles 3 and 4 of the Directive 2019/790/EU, p. 27; HUGENHOLTZ, P. Bernt, Artikel 3 en 4 DSM-richtlijn: tekst- en datamining, *AMI*, n. 5, 2019, p. 167; DUCATO; STROWEL, Limitations to text and Data Mining and Consumer Empowerment: Making the Case for a Right to 'Machine Legibility', p. 649; ALLEAUME, Christophe. Les exceptions en faveur de l'enseignement et de la recherche scientifique, *Comm. com. électr.*, n. 10, 2019, p. 9; BERNAULT, Carine. Exceptions - Fouille de texte, *Propr. intell.*, n. 72, 2019, p. 39; BALLETT, L. L'exception TDM dans la directive droit d'auteur, une volonté européenne, *Dalloz IP/IT*, n. 5, 2019, p. 304.

27 O texto completo do considerando 13 é o seguinte: “O conceito de instituições responsáveis pelo património cultural deverá abranger as bibliotecas acessíveis ao público e os museus, independentemente do tipo de obras ou de outro material protegido que tenham nas suas coleções permanentes, bem como arquivos e instituições responsáveis pelo património cinematográfico ou sonoro. O referido conceito deverá ainda incluir, nomeadamente, as bibliotecas nacionais e os arquivos nacionais, bem como estabelecimentos de ensino, organismos

de investigação e de radiodifusão do setor público, no que diz respeito aos seus arquivos e bibliotecas acessíveis ao público”.

28 Art. 4.

29 Art. 4(3). No entanto, não fica claro quais são essas ferramentas “apropriadas” para exercer o *opt-out*: tecnológicas, contratuais ou ambas?

30 Algumas obras estarão disponíveis para pesquisa, outras não. Na verdade, é muito provável que os beneficiários, geralmente hostis a essa nova exceção, usem sistematicamente o “*opt-out*” padrão para impedir a sua implementação.

31 Isso é destacado por BINCTIN, TDM: un enjeu de l’intelligence artificielle, *supra*, p. 7.

32 Ver, nesse sentido: GEIGER; FROSIO; BULAYENKO, Text and Data Mining in the Proposed Copyright Reform: Making the EU Ready for an Age of Big Data?, p. 838. A questão da remuneração é particularmente discutida no campo do aprendizado de máquina, que exige que a máquina reproduza e armazene um grande conjunto de obras potencialmente protegidas por direitos autorais para produzir resultados valiosos. Como a liberação de direitos traria uma enorme insegurança jurídica devido à grande quantidade de obras em jogo e, portanto, custos de transação muito altos, um direito de remuneração previsto em lei poderia facilitar o desenvolvimento do aprendizado de máquina e permitir a remuneração dos criadores, quando apropriada (por exemplo, quando a IA está produzindo uma obra derivada em um contexto comercial). Nesse sentido, ver: KOP, The Right to Process Data for Machine Learning Purposes in the EU, *supra*, p.7, propondo “a criação de um ambiente “*one-stop shop*” de liberação de direitos online, com o licenciamento obrigatório ou estatutário para conjuntos de dados

de treinamento de aprendizado de máquina semelhantes a uma agência de direitos coletivos multiterritorial e paneuropeia”. De uma forma mais geral, escrevendo sobre os direitos de remuneração estatutários como uma solução de compromisso viável no ambiente digital, ver: GEIGER, Christophe; BULAYENKO, Oleksandr, Creating Statutory Remuneration Rights in Copyright law: What Policy Options under the International Framework? *CEIPI Research Paper*, nº 2020-05. Disponível em <SSRN: <https://ssrn.com/abstract=3722108>>.

33 Ver também: DUSOLLIER, Séverine. The 2019 Directive on Copyright in the Digital Single Market: Some progress, a few bad choices, and overall a failed ambition, *Common Market Law Review*, v. 57, nº 4, 2020, p. 987: “A exceção para TDM para fins não relacionados à pesquisa é, portanto, bastante precária e é subserviente à sua proibição pelos titulares de direitos”.

34 Ver também, nesse sentido: DUCATO; STROWEL, Limitations to text and Data Mining and Consumer Empowerment: Making the Case for a Right to ‘Machine Legibility, *supra*, p. 649 *et seq.*: “Nem a proposta inicial da Comissão Europeia com foco no contexto de pesquisa, nem as disposições finais da Diretiva “CDSM” parecem suficientes para facilitar o uso da TDM para melhorar a divulgação inteligente e, mais amplamente, para aplicativos de IA”.

35 Ver, nesse sentido, SAG, Matthew. The New Legal Landscape for Text Mining and Machine Learning, *Journal of the Copyright Society of the USA*, v. 66, 2019, p. 291; CARROL, Michael. Copyright and the Progress of Science: Why Text and Data Mining Is Lawful, *U.C. Davis L. Rev.*, v. 53, 2019, p. 893; GEIGER; FROSIO; BULAYENKO, Text and Data Mining in the Proposed Copyright Reform: Making the EU Ready for an Age of Big Data? p. 821. Como corretamente dispõe Bernt Hugenholtz (em:

The New Copyright Directive: Text and Data Mining (Articles 3 and 4), *supra*: “as disposições de TDM da Diretiva “CDSM” garantem consideravelmente menos liberdade para a mineração de textos e dados do que inicialmente parecem fazer crer. A cláusula “opt-out” do art. 4, em particular, deixa os mineradores com fins lucrativos na UE à mercê dos proprietários de conteúdo. Isso coloca desenvolvedores de IA, jornalistas, laboratórios de pesquisa comerciais e outros inovadores em desvantagem competitiva em comparação com os Estados Unidos, onde a mineração de textos e dados é considerada um uso justo, mesmo que feita com fins lucrativos”.

36 Ver, nesse sentido: SENFTLEBEN, Martin *et al.*, Ensuring the Visibility and Accessibility of European Creative Content on the World Market: The Need for Copyright Data Improvement in the Light of New Technologies (12 Feb. 2021), disponível em <<https://ssrn.com/abstract=3785272>>; DREXL, Josef; HILTY, Reto *et al.* Artificial Intelligence and Intellectual Property Law, Position Statement of the MPI for Innovation and Competition of 9 April 2021 on the Current Debate, *Max Planck Institute for Innovation and Competition*, Research Paper n° 21-10, p. 3. O Japão, por exemplo, introduziu em 2019 uma nova exceção que permite de forma abrangente o uso de uma obra que não visa a desfrutar nem fazer com que outra pessoa desfrute da obra (art. 30-4, da Lei de Direitos Autorais Japonesa), que inclui o uso de obras protegidas por direitos autorais para atividades de mineração de textos e dados também para fins comerciais. Sobre esta disposição, ver: UENO, Tatsuhiro. The Flexible Copyright Exception for ‘Non-Enjoyment’ Purposes Recent Amendment in Japan and Its Implication, *GRUR Int.*, v. 70, 2021, p. 145.

37 Sobre a (ainda) incerta interpretação da regra dos três passos, ver, *i.e.*: GEIGER, Christophe. From Berne to National Law, via the Copyright Directive: The Dangerous Mutations of the Three-Step Test, *EIPR*, v. 29, 2007, p. 486.

38 Ver o art. 7, parágrafo 2º, da Diretiva 2019/790, que estabelece: “O artigo 5.º, no 5, da Diretiva 2001/29/CE é aplicável às exceções e limitações previstas no presente título. O artigo 6.º, no 4, primeiro, terceiro e quinto parágrafos, da Diretiva 2001/29/CE é aplicável aos artigos 3.º a 6.º, da presente diretiva”.

39 Sobre esse artigo, ver, por exemplo: GEIGER, Christophe. The Answer to the Machine should not be the Machine, Safeguarding the Private Copy Exception in the Digital Environment, *EIPR*, v. 30, 2008, p. 121, respondendo à CLARK, Charles, The Answer to the Machine is in the Machine, in: HUGENHOLTZ, B.P (ed.). *The Future of Copyright in a Digital Environment*, p. 139; HUGENHOLTZ, P. Bernt. Copyright, Contract and Code: What Will Remain of the Public Domain, *Brook. J. Int’l L.*, v. 26, 2000, p. 83.

40 GEIGER, Christophe; JÜTTE, Bernd Justin. The Right to Research as Guarantor for Sustainability, Innovation and Justice in EU Copyright Law, in: PIHLAJARINNE, Taina; MÄHÖNEN, Jukka; UPRETI, Pratyush (eds.). *Rethinking the Role of Intellectual Property Rights in the Post Pandemic World: An Integrated Framework of Sustainability, Innovation and Global Justice*. Edward Elgar, 2022. Ver, em especial, a condição de a fonte ser legal, que foi mantida (Art. 3(1) e Art. 4(1)), bem como as incertezas quanto às atividades de TDM em parcerias público/privadas, embora fortemente incentivadas pela União Europeia em sua política de pesquisa. A este respeito, o considerando 11 da diretiva permanece bastante vago: “[...] Embora

os organismos de investigação e as instituições responsáveis pelo património cultural devam continuar a beneficiar dessa exceção, deverão também poder contar com os seus parceiros privados para proceder à prospecção de textos e dados, inclusive através do recurso às suas ferramentas tecnológicas”. Isso significa que haverá extensão da exceção ao operador privado no caso de contrato de pesquisa com órgão público? (Ver, neste sentido, BINCTIN, TDM: un enjeu de l’intelligence artificielle, *supra*, p. 23). Isso permanece obscuro; é muito provável que tenhamos de esperar por uma decisão do Tribunal de Justiça da União Europeia (TJUE) para especificar os contornos exatos da exceção. Enquanto isso, a incerteza persistirá, arriscando aumentar ainda mais a distância entre a UE e outras regiões do mundo em inovação e pesquisa digital. Criticando essa incerteza, ver também: DUCATO, Rosanna; STROWEL, Alain. Ensuring Text and Data Mining: Remaining Issues with the EU Copyright Exceptions and Possible Ways Out, *EIPR*, v. 43, 2021, p. 322 *et sq.*

41 Entretanto, espera-se que os legisladores nacionais aproveitem ao máximo a margem de manobra deixada pela diretiva na sua transposição. Eles podem, por exemplo, implementar o “*opt out*” de uma maneira que proteja a eficácia da exceção. Para esse fim, os Estados Membros podem interpretar o “expressamente reservada pelos respetivos titulares de direitos de forma adequada” de maneira muito restritiva (impondo, por exemplo, algumas formalidades aos titulares de direitos sobre a reserva, como requisito de registro etc.), tornando insuficiente a mera implementação de Medidas Técnicas de Proteção (TPM) para a reserva da obra. Alternativamente, a “adequação” do “*opt-out*” pode incluir um elemento de “razoabilidade” da reserva. Por fim, uma interpretação restritiva da possibilidade de “*opt-out*” poderia ser exigida pelo “*effet utile*” do dispositivo e pela

necessidade de interpretá-lo à luz do direito fundamental à informação, protegido pelo art. 10, da (CEDH).

42 WIPO Conversation on Intellectual Property (IP) and Artificial Intelligence (AI), May 21, 2020, 2nd session, WIPO/IP/AI/2/GE/20/1 REV, para. 24, levantando a seguinte pergunta fundamental: “O uso dos dados que subsistem em obras de direitos autorais sem autorização para aprendizado de máquina constitui uma violação de direitos autorais?” (§24, i).

43 Ver FLYNN; GEIGER; QUINTAIS *et al.*, Implementing User Rights for Research in the Field of Artificial Intelligence: A Call for International Action, *supra*.

44 Ver, por exemplo: GEIGER, Christophe, Copyright as an Access Right, Securing Cultural Participation through the Protection of Creators’ Interests, *in*: GIBLIN, Rebecca; WEATHERALL, Kim G. (eds.). *What if we could reimagine copyright?* Acton: Australian National University (ANU) Press, 2016, p. 73 *et sq.*; Taking the Right to Culture Seriously: Time to Rethink Copyright Law, *in*: GEIGER, Ch. (ed.), Intellectual Property and Access to Science and Culture: Convergence or Conflict?, *CEIPI/ICTSD publication series on “Global Perspectives and Challenges for the Intellectual Property System”*, n° 3, Geneva/Strasbourg, 2016, p. 90; Promoting Creativity through Copyright Limitations, Reflections on the Concept of Exclusivity in Copyright Law, *supra*.