

IBDAUTORAL contribuição a consulta publica de IA

Allan Rocha de Souza <allan.rocha@ibdautorial.org.br>

qui 30/06/2022 20:45

Para: CJSUBIA <CJSUBIA@senado.leg.br>;

Cc: Instituto Brasileiro de Direitos Autorais <ibdautorial@ibdautorial.org.br>; Allan Rocha de Souza <allan@rochadesouza.com>;

 1 anexo

Contribuição à consulta pública PL 21_2020_v30.06.22.pdf;

Você não costuma receber emails de allan.rocha@ibdautorial.org.br. [Saiba por que isso é importante](#)

Boa noite,

Encaminho em anexo a Contribuição do Instituto Brasileiro de Direitos Autorais à Consulta Pública promovida pela estimada Comissão de Juristas do Senado Federal.

Seguimos à disposição para qualquer esclarecimento adicional,

Allan Rocha de Souza
Presidente e Diretor Científico
IBDAutorial

Rio de Janeiro, 30 de Junho de 2022.

Ref.: Consulta Pública

À Comissão de Juristas responsável por subsidiar a elaboração de substitutivo sobre inteligência artificial no Brasil (CJSUBIA)

Prezados(as),

Queremos, em primeiro lugar, felicitá-los pelos trabalhos dessa Comissão, essenciais a uma regulação adequada do tema.

O Instituto Brasileiro de Direitos Autorais (IBDAutoral) é uma associação de pesquisadores cujas finalidades centrais são desenvolver pesquisas aplicadas, cursos temáticos e contribuir para a construção de políticas públicas adequadas sobre direitos autorais.

Muito nos honra a possibilidade de contribuir para a solução de questões que entendemos estruturais ao tema. E dentre os vários aspectos que requerem atenção, centramos nossa atenção, nessa contribuição, na Mineração de Textos e Dados, Bases de Dados e Direitos Autorais.

INSTITUTO BRASILEIRO DE DIREITOS AUTORAIS

www.ibdautor.org.br

ibdautor@ibdautor.org.br

INTELIGÊNCIA ARTIFICIAL, DIREITOS AUTORAIS e MINERAÇÃO DE TEXTOS e DADOS

1. Inúmeras atividades requerem hoje o uso intensivo de dados. Entre elas, destacamos a pesquisa contemporânea e a inovação tecnológica. O próprio desenvolvimento e o treinamento de sistemas de inteligência artificial requerem a utilização de grandes volumes de dados, que podem ser estruturados ou não, de diversas tipologias e originados de fontes variadas. Os dados aqui referenciados podem assumir a forma de conjuntos de textos, imagens, sons e dados brutos, os quais, individualmente ou quando considerados em seu conjunto, podem vir a ser protegidos por direitos autorais e conexos.
2. Os direitos autorais, [regidos infraconstitucionalmente pela Lei 9.610/98](#), na essência, protegem *apenas* as *expressões* humanas, criativas e originais, que se enquadrem nas categorias literária, artística ou científica. (art. 7º). Ao fim do século XX, os programas de computador foram incorporados aos direitos autorais, expandindo seu objeto para incluir esta nova forma de expressão.
3. Em qualquer dos casos, a proteção recai unicamente sobre a *forma* que determinado conteúdo é expresso, e não alcança as ideias e concepções que lhe deram origem. Assim, ideias, abstrações, fatos, informações ou percepções não são alcançadas pela proteção por direitos autorais. Conseqüentemente, o aproveitamento econômico das ideias contidas nas obras por quaisquer pessoas é também excluído da seara do direito autoral (art. 8º).
4. Sob a Lei de Direitos Autorais, as compilações, bases e bancos de dados são protegidos desde que sejam originais. Em outras palavras, a organização ou disposição de seu conteúdo deve sair do "lugar-comum", apresentando caráter singular e alguma contribuição criativa. Desta maneira, o direito autoral não alcança bases que organizem seus dados seguindo padrões comumente adotados - seja por razões de acessibilidade e compatibilidade, seja por questões mercadológicas. Por outro lado, os dados ali

compreendidos, quando individualmente considerados, não seriam protegidos por direitos autorais apenas pelo fato de comporem a base de dados.

5. Pode estar se tratando, por exemplo, de dados, que são meros fatos ou abstrações e não expressões tangíveis, a legislação explicitamente os exclui da proteção estendida aos bancos de dados (art. 7º, § 2º) e, mesmo no caso de dados contidos em obras, nenhuma proteção recai sobre os mesmos enquanto não satisfizerem todos os requisitos do direito autoral, como fixação em meio tangível e originalidade.
6. O art. 29 da lei nº 9.610/98 garante ao titular de direitos sobre a obra o controle sobre diversos usos da obra, incluindo diversas atividades intrínsecas ao processo de mineração de dados - tais como a reprodução de conteúdo, sua adaptação (por exemplo, para leitura por máquinas) e a inclusão do referido material em bases de dados.
7. Ao mesmo tempo, o art. 87 da mesma lei confere amplos poderes ao titular de um banco de dados protegido, inclusive proibindo a reprodução, total ou parcial (inciso I), e a divulgação dos resultados de processos de adaptação ou transformação do material (inciso IV).
8. Este nível de proteção se torna problemático quando não se encontra acompanhado de uma regulação que estabeleça situações nas quais seja possível fazer uso do conteúdo das bases de dados para fins que não comprometam o aproveitamento econômico das obras protegidas por direitos autorais - sejam elas obras convencionais, como imagens, textos ou áudio, ou compilações de dados originais.
9. Diante desta carência regulatória, atividades como o desenvolvimento e treinamento de IA se encontram em situação vulnerável: sem termos claros que permitam o uso de textos e dados para fins de mineração, desenvolvedores e pesquisadores podem se encontrar privados de material de treinamento, presos em relações contratuais extremamente desiguais com titulares ou forçados a operar à margem da lei.

MINERAÇÃO DE TEXTOS E DADOS

10. A *mineração de textos e dados*, também conhecida como descoberta de conhecimento a partir de dados, consiste em um processo de extração e análise de grandes quantidades de dados, a partir dos quais são extraídos padrões e correlações que gerarão informações relevantes para o analista ou usuário. Está, portanto, inserida nos processos de análise de dados, sendo crucial para o treinamento e operação de inteligências artificiais. Importante destacar que a mineração de textos e dados aqui abordada em nada se comunica com a “mineração” relacionada a criptomoedas ou às redes *blockchain*.
11. No que diz respeito ao treinamento e operação de sistemas de IA, o foco principal é ajudar a máquina a compreender objetos, modelos, e padrões que serão importantes para a realização de suas atividades. Exemplos práticos de destaque podem ser vistos nas artes, ciências, checagem de fatos e na área da saúde, inclusive no combate à COVID-19.
12. A mineração de textos e dados visa apenas extrair informações e dados contidos em um determinado material (as quais, como vimos, não recebem proteção por direito autoral) e utilizá-los para uso interno. Portanto, não visa utilizar a expressão de obras artísticas, literárias ou científicas (estas, sim, protegidas) ou reproduzir tais obras no produto final - não afetando, portanto, a exploração normal da obra por seu autor ou titular, pois com ela não concorre.

MINERAÇÃO DE TEXTOS e DADOS: PANORAMA INTERNACIONAL

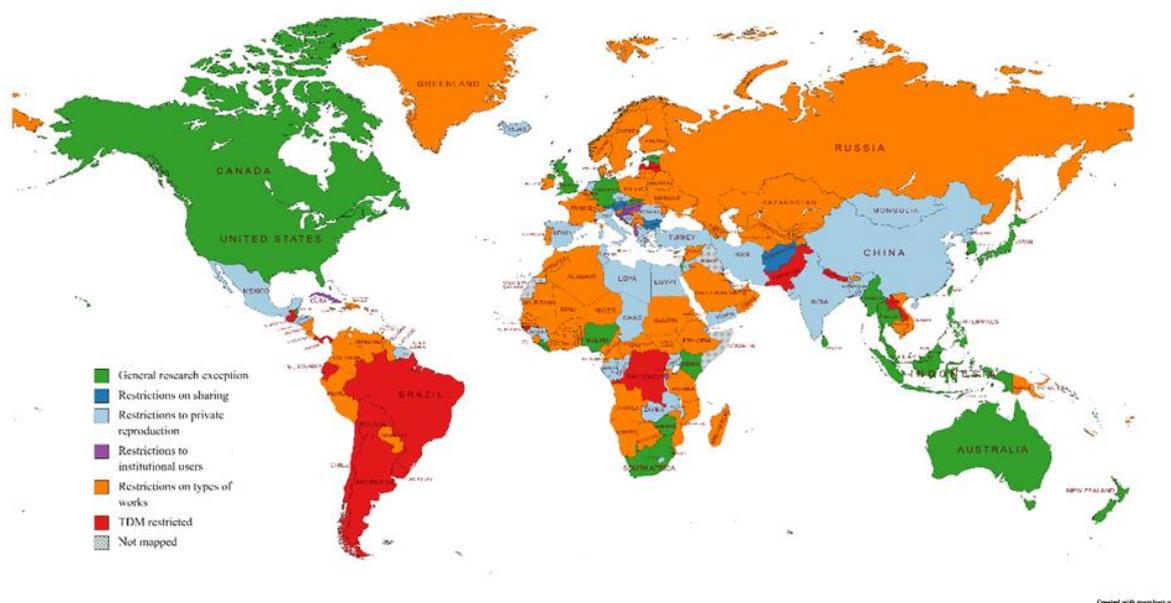
13. Os Estados Unidos, desde o início desse século, estabeleceram diversos precedentes em sua jurisprudência no que se refere à licitude da mineração de textos e dados perante o direito autoral, para fins que incluem o uso acadêmico e comercial.
14. O Japão foi o primeiro país a incluir uma limitação específica para mineração de dados em 2009. [Tal dispositivo recebeu ampliações até sua versão atual](#), que permite quaisquer usos que não envolvam a exploração da expressão de uma obra, no seu artigo 30-4. Outras limitações similares também foram introduzidas na ocasião.
15. A União Europeia, [em sua Diretiva 2019/790](#), estabeleceu dois artigos sobre o tema, mandatórios para todos os seus países-membros: o art. 3º permite a mineração de textos e dados por instituições de pesquisa e de conservação, a partir de obras adquiridas licitamente, desde que estritamente para fins de pesquisa. O art. 4º, por sua vez, estabelece uma limitação mais ampla, que permite reproduções temporárias para fins de mineração, inclusive para fins comerciais - salvo se o titular do material houver expressamente reservado para si o direito de fazê-lo.
16. Na América Latina, [o Código Ingenios do Equador](#), que possui diversas limitações de direito autoral, especialmente no âmbito da pesquisa e educação, também é o único país da América Latina a possuir uma limitação específica para mineração de dados, embora restrita a usos e agentes específicos.
17. Em 2021, [Singapura introduziu uma limitação similar em seu art. 244](#), no qual se permite o uso de obras protegidas para análise computacional de dados, uma vez satisfeitas certas condições - por exemplo, obtenção lícita das obras, não divulgação das mesmas, exceto para fins de verificação ou colaboração em pesquisas conjuntas.
18. Mais recentemente, em 28 de junho de 2022, o [Reino Unido \(2022\) publicou os resultados de sua consulta pública](#) a respeito de como tratar a inteligência artificial no campo da propriedade intelectual. [O documento revela que, embora o país já possua uma limitação para a mineração de textos e dados sem fins comerciais desde 2014 \(REINO UNIDO,](#)

[2014](#)), muitos respondentes a consideraram insuficiente em face de novas limitações aprovadas por diversos locais desde então.

19. Desta forma, foram consideradas diversas opções de ação regulatória, dentre as quais [decidiu-se pela adoção de uma limitação mais ampla, permitindo o uso de obras protegidas e bases de dados em TDM para quaisquer fins \(comerciais ou não\), não permitindo aos titulares o direito de exclusão dessa regra \(opt-out\)](#).

20. Não obstante esse movimento internacional, principalmente de algumas das maiores potências tecnológicas, o Brasil e outros países da América Latina ainda possuem uma legislação que não garante a mínima segurança jurídica para pesquisadores e demais agentes envolvidos em atividades intensivos em dados, como as startups.

21. Este atraso nos coloca em situação de grande desvantagem em relação a outros países do mundo, como se observa da Figura abaixo, extraída de [recente estudo](#) que buscou avaliar legislações de direito autoral ao redor do mundo e as categorizar de acordo com o seu grau de “abertura” no que diz respeito aos usos necessários para fins de pesquisa, com maior foco em práticas de mineração de textos e dados.



Flynn, Schirru, Palmedo, Izquierdo (2022)

22. Na figura acima, os países em verde possuem uma legislação capaz de autorizar distintos usos (por ex. reprodução e comunicação) de todo e qualquer tipo de obra para fins de pesquisa utilizando práticas de mineração de textos e dados. Os países em laranja possuem

alguma limitação no que diz respeito ao uso de um determinado tipo de obra (por ex: software). Já os países em azul claro autorizam o uso de obras para fins de pesquisa apenas quando este uso é conduzido por um indivíduo para seu uso privado. Os poucos países vermelhos, como é o caso do Brasil, possuem uma legislação tão restrita que apenas viabilizam a reprodução de citações e pequenos trechos de obras para fins de pesquisa, o que, por sua vez, inviabilizaria qualquer prática de mineração de textos e dados.

23. A incerteza regulatória que expõe cientistas, instituições de ensino e pesquisa e negócios intensivos em dados - desde startups até multinacionais - pode impactar de maneira substancial o desenvolvimento tecnológico e econômico do país, notadamente no que diz respeito aos aspectos abaixo:

- a. Considerando o caráter transfronteiriço de negócios e pesquisas intensivas em dados, o Brasil não seria visto como um parceiro adequado, por conta de sua inadequação regulatória, o que contribuiria para a perpetuação de uma posição de mero consumidor de tecnologia em alguns setores, e não de um país capaz de explorar o seu potencial de desenvolvimento de tecnologias e sistemas de IA;
- b. Como dito, a mineração de textos e dados é parte essencial do treinamento de máquinas. Assim, ao manter uma regulação que inviabiliza tais práticas, compromete-se todos os projetos cujo treinamento de máquinas envolve o uso de dados não pessoais que possam vir a ser protegidos por direitos autorais e conexos.

24. Ademais, é importante destacar que a Portaria GM Nº 4.617, de 6 de abril de 2021 - a qual instituiu a Estratégia Brasileira de Inteligência Artificial - estabeleceu, em sua página 8, a importância da cópia e edição de dados de treinamento de IA e a necessidade de uma limitação de direito autoral, em território nacional, que permita a mineração de textos e dados.

25. Para que haja uma mudança efetiva neste cenário, é fundamental que as práticas de mineração de textos e dados sejam endereçadas pela legislação vigente, sendo as limitações e exceções o instrumento ideal para a sua regulação e, por conseguinte, da promoção do desenvolvimento, da pesquisa e da inovação nas searas pública e privada.

LIMITAÇÕES E EXCEÇÕES: O SEU PAPEL NO SISTEMA DE DIREITOS AUTORAIS

26. Neste contexto, é importante ressaltar o papel das limitações e exceções no sistema de direito autoral enquanto mecanismos internos, pois visam garantir o equilíbrio entre os interesses do autor e outros direitos fundamentais - como liberdade de expressão, educação, pesquisa, privacidade e acesso à cultura. As limitações e exceções são o reconhecimento explícito de uma ordem constitucional de convivência harmoniosa entre direitos fundamentais e de respeito à função social da propriedade, nos termos do art. 5º, XXIII da Constituição Federal.
27. No âmbito internacional, [o Tratado de Marrakesh](#) ilustra bem a necessidade premente de harmonização ao destacar a importância das limitações como mecanismo de equilíbrio entre a proteção efetiva dos direitos dos autores e o interesse público mais amplo, sobretudo em relação a questões de educação, pesquisa e acesso ao conhecimento e informação.
28. No âmbito nacional, o Superior Tribunal de Justiça, a partir do julgamento do paradigmático [Recurso Especial nº 964.404](#) e em reconhecimento ao papel das limitações e exceções enquanto instrumento de tutela de direitos fundamentais perante os interesses do autor, decidiu que o rol de limitações presente na lei 9.610/98 é exemplificativo, estabelecendo a regra dos três passos da Convenção de Berna e do Acordo TRIPS e a função social da propriedade como parâmetros para identificação de novos usos livres. Tal entendimento foi consolidado no Poder Judiciário com o [Enunciado 115 da III Jornada de Direito Comercial, em 2019](#).
29. Oportuno também destacar alguns pontos centrais do [preâmbulo da Diretiva 2019/790](#), no qual a União Europeia instituiu suas limitações para mineração de textos e dados. O primeiro item em destaque, o de número 8, afirma: "A prospecção de textos e dados torna possível o tratamento de grandes quantidades de informação para obter novos conhecimentos e descobrir novas tendências. Embora as tecnologias de prospecção de textos e dados sejam predominantes em toda a economia digital, existe um amplo reconhecimento de que esta prospecção pode beneficiar, nomeadamente, a comunidade

científica e, ao fazê-lo, apoiar a inovação. (...)". Igualmente relevante é o Item 11 do preâmbulo: "(11) A insegurança jurídica no âmbito da prospecção de textos e dados deverá ser eliminada através da previsão de uma exceção obrigatória para as universidades e outros organismos de investigação, bem como para as instituições responsáveis pelo patrimônio cultural, ao direito exclusivo de reprodução e ao direito de impedir a extração a partir de bases de dados. (...)”

30. Para além das entidades de pesquisa, a Diretiva também aborda a mineração de textos e dados de forma mais ampla ainda no item 3 do preâmbulo: "A rápida evolução tecnológica continua a mudar a forma como as obras e outro material protegido são criados, produzidos, distribuídos e explorados. Continuam a surgir novos modelos empresariais e novos intervenientes. Cumpre que a legislação aplicável esteja orientada para o futuro, para não limitar a evolução tecnológica. (...)”
31. Nesse sentido, o item 17 impõe a necessidade de uma limitação que facilite atividades de TDM para entidades públicas e privadas, as quais utilizam tais técnicas "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 tecnologias", com o objetivo de "proporcionar maior segurança jurídica em tais casos e incentivar a inovação também no setor privado, a presente diretiva deverá prever, em determinadas condições, uma exceção ou limitação para as reproduções e extrações de obras ou outro material protegido, para efeitos de prospecção de textos e dados e permitir que as cópias sejam conservadas durante tanto tempo quanto necessário para fins dessa prospecção de textos e dados.”
32. Mesmo considerando o rol exemplificativo das limitações e exceções e sua interpretação extensiva, a legislação vigente se revela insuficiente para garantir a segurança jurídica e estabilidade necessária para a condução de projetos envolvendo mineração de textos e dados, em razão dos riscos que apresenta.

PROPOSTA DE AJUSTES NO PL 21/2020

33. O resultado da operação de um sistema de IA não depende apenas do algoritmo, mas sim também dos dados que são utilizados para treinar aquele sistema. Uma base de dados limitada ou mal selecionada pode aumentar substancialmente o risco de enviesamento (*bias*) do produto da operação daquele sistema.

Basta pensar em um sistema de reconhecimento do conteúdo de imagens. Para que o sistema seja capaz de identificar que existe, por exemplo, um gato em uma determinada imagem, ele precisa ser treinado com fotos e imagens de gatos para conseguir identificar os padrões necessários e, assim, executar com precisão a função para o qual foi desenvolvido. A quantidade e diversidade de fotos de gatos contribuirá para o treinamento daquele sistema e, por conseguinte, para a precisão de seus resultados.

Como se observa do exemplo acima, é comum que muitos dos dados que são utilizados para treinar sistemas de IA possam vir a ser protegidos por Direitos Autorais (textos, imagens, fotos, códigos-fonte, etc), razão pela qual, não há como tratar de sistemas de IA cujo produto de sua operação depende de um treinamento prévio e/ou de bases de dados, sem considerar a necessária intersecção com a legislação autoral e os eventuais ajustes nas leis competentes.

Assim, entendemos como essencial a existência de uma limitação expressa aos direitos autorais que assegure a possibilidade de mineração de textos e dados para fins de inovação e pesquisa.

34. O art. 5º, VIII do projeto em análise representa um avanço na concretização de uma limitação ao direito autoral que permita o uso de dados no âmbito da IA, tal como discutido na Estratégia Brasileira de Inteligência Artificial. Contudo, entendemos que o dispositivo, tal como apresentado, necessita de uma redação mais apurada para garantir maior segurança jurídica e resultados efetivos no campo das práticas de IA. Destacamos, a seguir, alguns pontos a serem considerados nesse sentido.

35. Dentre os aspectos identificados, primeiramente, observa-se que não é feita menção explícita a "obras protegidas". Uma vez que constituem o objeto central da proteção por direito autoral, a sua omissão do inciso pode excluí-las da possibilidade outorgada pelo dispositivo, tornando-o inerte. Sugerimos, nesse sentido, a substituição do termo "textos" por "obras" (as quais abrangem todo tipo de criação protegida que possa ser usada por uma IA, incluindo textos, imagens e áudio).
36. Em relação aos usos permitidos, entendemos que a simples menção da palavra "uso", embora abrangente e necessária, peca por não expressar, com suficiente clareza, que tipo de operações poderiam ser realizadas no contexto da inteligência artificial. Assim, convém adicionar outras formas de utilização, tais como a coleta, reprodução, tratamento, adaptação, armazenamento, compartilhamento e outras formas de utilização inerentes a essas operações.
37. Quanto à finalidade, é relevante ressaltar que sistemas de inteligência artificial não necessitam de dados única e exclusivamente para seu treinamento. A simples operação e atualização de uma IA requer o uso constante de novos dados. Então entendemos ser fundamental assegurar esta finalidade ampla, não restringindo a possibilidade apenas para o treinamento, mas também o seu aperfeiçoamento e atualização.
38. Ainda com relação às finalidades, entendemos ser essencial assegurar o direito de mineração de textos e dados, a partir de obras protegidas ou não por direitos autorais, tanto para fins de pesquisa e inovação, independente de finalidade comercial ou destinação comercial dos resultados da operação, da inovação em si e nesse sentido propomos a seguinte alteração da Lei de Direitos Autorais, com a inclusão do inciso IX no art. 46:
- “Art. 46. Não constitui ofensa aos direitos autorais:
(...)
IX – o acesso, coleta, extração, reprodução, armazenamento, transformação, apresentação, utilização e reuso de qualquer obra em processos de mineração de dados e textos e em todos os atos necessários para sua realização, , na medida justificada para o seu fim, para quaisquer propósitos que não tenham como objetivo final a simples reprodução, exibição ou disseminação da obra original em si.”
39. Ademais, resta garantir segurança adicional aos processos de mineração de textos e dados, conforme salientado na própria Estratégia Brasileira de Inteligência Artificial.

40. E nesse sentido propomos adicionar ao projeto sob análise, o seguinte: Art. 3º A aplicação de inteligência artificial no Brasil tem por objetivo o desenvolvimento científico e tecnológico, bem como: ... XV – a harmonização com as Leis nºs 13.709, de 14 de agosto de 2018 (Lei Geral de Proteção de Dados Pessoais), 12.965, de 23 de abril de 2014, 12.529, de 30 de novembro de 2011, 8.078, de 11 de setembro de 1990 (Código de Defesa do Consumidor), **9.609, de 19 de Fevereiro e 1998 (Lei de Software), 9.610, de 19 de Fevereiro de 1998 (Lei de Direitos Autorais)** e 12.527 de 18 de novembro de 2011.
41. E também alterar o art. 5 do mesmo projeto da seguinte forma: Art. 5º São princípios para o desenvolvimento e a aplicação da inteligência artificial no Brasil: [...] VIII – disponibilidade de dados: não violação do direito de autor **pela coleta, reprodução, armazenamento, adaptação, compartilhamento e uso de obras**, dados e bancos de dados por ele protegidos, para fins de **mineração de textos e dados** e de treinamento e **operação** de sistemas de inteligência artificial, desde que não seja impactada a exploração normal da obra por seu titular.

DIREITOS AUTORAIS E INTELIGÊNCIA ARTIFICIAL: ASPECTOS RELACIONADOS

42. Não diretamente vinculados à IA, mas ainda assim pressupostos essenciais às atividades de inovação são as questões da preservação e digitalização dos acervos das instituições de memória, as pesquisas transfronteiriças, as políticas de acesso, dados e ciência abertos e a apropriação de produtos gerados autonomamente por sistema de IA.
43. Ainda que não possuam relação direta com o tema da Inteligência Artificial, a obsolescência da legislação de direitos autorais no Brasil afeta atividades fundamentais à preservação da memória e da viabilização do acesso à cultura, notadamente pela ausência de uma limitação que permita aos museus, bibliotecas, arquivos e instituições de preservação fazerem usos específicos de obras de seus acervos (ex: digitalização) para fins de preservação. Os incêndios ocorridos no [Museu Nacional](#) e no [Museu de História Nacional da UFMG](#) são apenas alguns exemplos trágicos de perdas irreversíveis para a nossa cultura e memória. O tema foi um dos pontos centrais de discussão na última reunião

do [Standing Committee on Copyright and Related Rights da Organização Mundial da Propriedade Intelectual](#), e, junto do tema abaixo, é considerado como questão prioritária dentro dos direitos autorais. Sem a existência de uma limitação aos direitos autorais que assegure o exercício direito de preservação, o acervo destas instituições não poderá ser usado para fins de mineração de textos e dados, reduzindo substancialmente seu potencial para fins de inovação e desenvolvimento de sistemas de IA, além do enorme dano social que a ausência desta limitação traz à própria atividade destas instituições.

44. A pandemia do novo coronavírus trouxe consigo a emergência da adoção de modelos de trabalho, ensino e pesquisa à distância. Atividades que antes eram territorialmente restritas podem ser executadas em diversas partes do mundo, o que reforça a necessidade de enfrentar a questão dos usos e projetos transfronteiriços. Projetos em colaboração com instituições estrangeiras reforçam a necessidade de inclusão de um dispositivo na lei de direitos autorais que traga segurança jurídica sobre a legislação de qual país será aplicável às atividades de mineração de textos e dados desenvolvidas a partir de projetos conjuntos internacionais.
45. Pelos motivos já expostos e não menos importante é a necessidade de aprofundar e avançar na política de acesso, dados, ciência e inovação abertas, que podem imensamente contribuir para a inovação tecnológica e científica.
46. Outro importante tema a ser considerado quando da intersecção entre a Inteligência Artificial e os Direitos Autorais é o tratamento jurídico a ser concedido aos produtos gerados a partir do emprego de sistemas de IA e que, se fossem criados por seres humanos, estariam cobertos pela proteção garantida pela LDA. A questão nevrálgica não é o uso desses sistemas com controle substancial do ser humano, situação em que seriam meras ferramentas, mas sim quando a operação de um sistema de IA é determinante ao conteúdo criativo daquela determinada obra. Já foram mapeadas na literatura [inúmeras propostas de modelos regulatórios para lidar com esses casos](#). Não obstante, ao contrário do que ocorre com a mineração de textos e dados, ainda não parece existir um consenso a respeito de uma alternativa ideal, e o tema continua a ser debatido em fóruns internacionais, [como é o caso da Organização Mundial da Propriedade Intelectual \(OMPI\)](#).

CONSIDERAÇÕES FINAIS

47. Os termos “limitações” e “exceções” (aos direitos autorais) comumente são objeto de repúdio desinformado sob uma falsa concepção de que tais usos livres, autorizados pela lei, seriam danosos aos autores, artistas e demais titulares, sem trazerem benefícios sociais correspondentes. Mas, muito pelo contrário, as limitações e exceções aos direitos autorais representam a garantia de harmonização com outros direitos e garantias fundamentais com os quais a exclusividade autoral deve conviver, como os direitos à pesquisa, educação, inovação, liberdade de expressão, o acesso à cultura, conhecimento e informação, dentre outros.
48. Hoje, a maior parte das potências econômicas e países que são conhecidos como polos geradores de tecnologias de inteligência artificial já reconheceram, seja em seus tribunais como em sua legislação, por meio da criação e adequação das limitações e exceções aos direitos autorais, a importância da utilização de obras protegidas para extração de dados e informações a seu respeito, que não são protegidos pelos direitos autorais e em nada impactam sua fruição e exploração normal, para o desenvolvimento de sistemas de IA. Como reiterado ao longo dessa contribuição, o objetivo da mineração não é a fruição da obra, mas a extração de ideias e padrões de um conjunto de centenas ou milhares de textos e dados, aspectos expressamente excluídos da proteção autoral,
49. Tais usos não visam privilegiar apenas a pesquisa e inovação, mas sim todo e qualquer projeto intensivo em dados e todo ecossistema da inovação. Tanto é que a recente alteração na Diretiva da União Europeia que trata sobre o tema foi impulsionada, dentre outros fatores, pela necessidade de viabilizar e garantir segurança jurídica a startups e negócios intensivos em dados.
50. O Brasil, ao continuar a negligenciar a importância de tais ajustes em sua legislação de direitos autorais, inevitavelmente restará prejudicado e fadado a ser um mero consumidor de tais tecnologias, expondo pesquisadores e empreendedores a severas consequências legais quando optarem por trabalharem com sistemas de inteligência artificial que demandem dados para o seu treinamento.

51. As alterações aqui sugeridas são parte de um processo mais abrangente de revisão e reforma da legislação aplicável e que será crucial para a inovação em matéria de inteligência artificial. Sem um arcabouço institucional que garanta o mínimo de segurança para os interessados em conduzir projetos intensivos em dados, não há como sequer considerar outras questões como a regulação dos produtos gerados mediante o emprego de sistemas de IA e tantos outros temas da maior relevância e que partem do princípio que já existe um sistema de IA em operação.
52. É inadiável a definição da questão da mineração de textos e dados pelo Congresso Nacional, tornando-se ainda mais urgente pela constante emergência de novas possibilidades tecnológicas. De todo, nos parece inexorável a urgência em instituir uma limitação que permita mineração de textos e dados na lei de direitos autorais, sob pena de, ao mantermos a inércia, afogarmos o ecossistema de inovação e aprofundarmos o nosso atraso tecnológico.
53. Finalmente, gostaríamos de reiterar as nossas congratulações pelos trabalhos dessa Comissão, e nos colocar à vossa plena disposição para o esclarecimento de qualquer dúvida, bem como para o fornecimento de qualquer informação ou documento que julguem essenciais.

Atenciosamente,



Allan Rocha de Souza

Diretor Científico

Luca Schirru

Diretor Executivo

Alexandre Fairbanks

Diretor Administrativo

Leon Queiroz Ramos

Iniciação Científica

Luís Felipe Majeste

Iniciação Científica

Miguel Alvarenga

Pesquisador

Patrícia Porto

Pesquisadora

Marcelo Franco

Iniciação Científica

REFERÊNCIAS

ABHARI, R. S.; MARINI, M.; CHOKANI, N. COVID-19 Epidemic in Switzerland: Growth Prediction and Containment Strategy Using Artificial Intelligence and Big Data. medRxiv. 2020. Disponível em: <https://doi.org/10.1101/2020.03.30.20047472>. Acesso em 18 abr. 2022.

ALVARENGA, M. B. Mineração de dados, Big Data e direitos autorais no Brasil. 2019. Dissertação (Mestrado em Políticas Públicas, Estratégias e Desenvolvimento) – Instituto de Economia, Universidade Federal do Rio de Janeiro, Rio de Janeiro.

ARNOLD, L.G., BAUER, C. On the growth and welfare effects of monopolistic distortions. *Journal of Economics*, vol. 97, 2009, pp. 19–40.

BATISTA, A.F.M.; MIRAGLIA, J.L.; DONATO, T.H.R.; FILHO, A.D.P.C. COVID-19 diagnosis prediction in emergency care patients: a machine learning approach. medRxiv. 2020. Disponível em: <https://doi.org/10.1101/2020.04.04.20052092>. Acesso em 18 abr. 2022.

BITTENCOURT, P.; RAUEN, A. Políticas de inovação: racionalidade, instrumentos e coordenação. In: RAPINI, M., RUFFONI, J.; SILVA, L., MOTA E ALBUQUERQUE, E. (Org) *Economia da ciência, tecnologia e inovação: fundamentos teóricos e a economia global*. 2a. edição Belo Horizonte: FACE-UFMG, 2021, Capítulo 20.

BRASIL. Conselho da Justiça Federal. III Jornada de Direito Comercial: Enunciados aprovados em 7/6/2019. 2019. Disponível em: https://www.cjf.jus.br/cjf/noticias/2019/06-junho/iii-jornada-de-direito-comercial-e-encerrada-no-cjf-com-aprovacao-de-enunciados/copy_of_EnunciadosaprovadosIIIJDCREVISADOS004.pdf. Acesso em 15 jul. 2021.

BRASIL. Lei nº 9.610, de 19 de fevereiro de 1998. Altera, atualiza e consolida a legislação sobre direitos autorais e dá outras providências. 1998. Disponível em: http://www.planalto.gov.br/ccivil_03/leis/19610.htm. Acesso em 24 ago. 2021.

BRASIL. Portaria MCTI nº 4.617, de 6 de abril de 2021. Institui a Estratégia Brasileira de Inteligência Artificial e seus eixos temáticos. 2021. Brasil: Diário Oficial da União. Publicado no D.O.U de 12 de abril de 2021.

BRASIL. Superior Tribunal de Justiça. 3ª Turma. Recurso Especial nº 964404/ES (2007/0144450-5). Recorrente: Mitra Arquidiocesana de Vitória. Recorrido: Escritório

Central de Arrecadação e Distribuição (ECAD). Relator: Min. Paulo de Tarso Sanseverino. Brasília, 15 de março de 2011. Lex: Diário de Justiça Eletrônico, Brasília, v. 815, 23 mai. 2011.

BROWN, K. Digital Rights Management: Trafficking in Technology That Can Be Used to Circumvent the Intellectual Property Clause. 40 Houston Law Review, vol. 803, 2003, p. 803-836.

CARROLL, M. W. Copyright and the Progress of Science: Why Text and Data Mining is Lawful. Washington College of Law Research Paper No. 2020-15. American University Washington College of Law, 2020.

CHARQUERO, P. D. Flexibilidades al derecho de autor en América Latina. Datsoc e Fundación Karisma, abr. 2021.

CHEN, H; CHIANG, R. H. L.; STOREY, V. C. Business Intelligence and Analytics: from Big Data to Big Impact. MIS Quarterly: Management Information Systems, vol. 36 (4), p. 1165-1188, dez. 2012.

CHILE. Ley N° 17.336, de agosto de 1970 sobre la Propiedad Intelectual (modificada por la Ley N° 21045 del 3 de noviembre de 2017). 2017. Disponível em: <https://wipolex.wipo.int/en/text/511570>. Acesso em 18 abr. 2022.

COLÔMBIA. Ley N° 23 de 28 de enero de 1982, sobre Derechos de Autor (modificada por la Ley N° 1915 de 12 de julio de 2018). 2018. Disponível em: <https://wipolex.wipo.int/en/text/584747>. Acesso em 18 abr. 2022.

CORIAT, Benjamin. From Natural-Resource Commons to knowledge Commons: Common Traits and Differences. In: International Seminar on Property and Commons: new issues of shared access and innovation. Université de Paris, Sorbonne, França, 2013.

DEAN, Jared. Big Data, Data Mining, and Machine Learning: Value Creation for Business Leaders and Practitioners. Wiley, 2014.

DERCLAYE, E. The Legal Protection of Databases: A Comparative Analysis. Edward Elgar, 2008.

DREXL, J.; HILTY, R. M.; BENEKE, F.; DESAUNETTES, L. FINCK, M.; GLOBOCNIK, J.; OTERO, B. G.; HOFFMANN, J.; HOLLANDER, L.; KIM, D.; RICHTER, H.; SCHEUERER, S.; SLOWINSKI, P. R.; THONEMANN, J. Technical Aspects of Artificial Intelligence: An Understanding from an Intellectual Property Law Perspective. Max Planck Institute for Innovation and Competition Research Paper Series – Research Paper No. 19-13 (Research Group on the Regulation of the Digital Economy, 2019).

DURO, L. D. New Text and Data Mining Exceptions: Implementation in the EU and the possible outcome of Brazil adopting a similar approach. 2019. Dissertação (Master in Intellectual Property and Knowledge Management) – Faculty of Law, Maastricht University, Maastricht, Holanda.

EICHER, T., PENALOSA, C. G. Endogenous strength of intellectual property rights: implications for economic development and growth. *European Economic Review*, vol. 52 (2), 2008, pp. 237–258.

EQUADOR. Código Orgánico de la Economía Social de los Conocimientos, Creatividad e Innovación. Año IV – N° 899. Quito, 9 de dezembro de 2016. Disponível em: <https://wipolex.wipo.int/en/text/439750>. Acesso em 18 abr. 2022.

FERRAZ, J.C.; de PAULA, G.; KUPFER, D. Política industrial. In: KUPFER, D., HASENCLEVER, L. *Economia industrial: fundamentos teóricos e práticas no Brasil*. 3a. ed. São Paulo: Atlas, 2002, Capítulo 24.

FLYNN, S.; GEIGER, C.; QUINTAIS, J. P.; MARGONI, T.; SAG, M.; GUIBAULT, L.; CARROLL, M. W. Implementing User Rights for Research in the Field of Artificial Intelligence: A Call for International Action. *PIJIP/TLS Research Paper Series*, 48, 2020.

FLYNN, S.; PALMEDO, M. The User Rights Database: Measuring the Impact of Copyright Balance. 4 dez. 2017. Disponível em: <https://ssrn.com/abstract=3082371>. Acesso em: 18 abr. 2022.

FURUKAWA, Y. The protection of intellectual property rights and endogenous growth: Is stronger always better? *Journal of Economic Dynamics & Control*, vol. 31, pp. 3644–3670, 27 fev. 2007.

GANGOPADHYAY, K.; MONDAL, D. Does stronger protection of intellectual property stimulate innovation? *Economics Letters*, vol. 116, pp. 80-82, 24 jan. 2012.

GEIGER, C.; FROSIO, G.; BULAYENKO, O. The Exception for Text and Data Mining (TDM) in the Proposed Directive on Copyright in the Digital Single Market - Legal Aspects (March 2, 2018). Centre for International Intellectual Property Studies (CEIPI) Research Paper No. 2018-02, Disponível em: <https://ssrn.com/abstract=3160586> or <http://dx.doi.org/10.2139/ssrn.3160586>. Acesso em 18 abr. 2022.

GERVAIS, D. J. Exploring the Interfaces Between Big Data and Intellectual Property Law. *Journal of Intellectual Property, Information Technology and Electronic Commerce Law*, vol. 10 (1), pp. 22-38, 2019.

HAN, J.; PEI, J.; KAMBER, M. *Data mining: concepts and techniques*. [s.l.] Elsevier, 2011.

HO, Dean. Addressing COVID-19 Drug Development with Artificial Intelligence. *Advanced Intelligent Systems*. vol. 2. 5. 2020. Publicado por WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim. Disponível em: <https://onlinelibrary.wiley.com/doi/full/10.1002/aisy.202000070>. Acesso em 18 abr. 2022.

HUGENHOLTZ, P. B. Data property: Unwelcome Guest in the house of IP. In: REDA, J. (ed.). *Better Regulation for Copyright: Academics meet Policy Makers*. TheGreens/EFA, p. 65-77, 2017. Disponível em: https://juliareda.eu/wp-content/uploads/2017/09/2017-09-06_Better-Regulation-for-Copyright-Academics-meet-Policy-Makers_Proceedings.pdf. Acesso em 18 abr. 2022.

JAPÃO. Ministério da Justiça. Lei nº 38, de 6 de maio de 1970. Emenda nº 73 de 2009. Disponível em: <http://www.japaneselawtranslation.go.jp/law/detail/?printID=&ft=1&re=02&dn=1&x=0&y=0&co=01&ia=03&ky=copyright+act&page=24&vm=02&lvm=02&id=3379>. Acesso em 18 abr. 2022.

JAPÃO. Ministério da Justiça. Lei nº 38, de 6 de maio de 1970. Emenda nº 30 de 2018. Disponível em: <http://www.japaneselawtranslation.go.jp/law/detail/?printID=&ft=1&re=02&dn=1&x=0&y=0&co=01&ia=03&ky=copyright+act&page=24&vm=02&lvm=02&id=3379>. Acesso em 18 abr. 2022.

KAPLAN, A.; HAENLEIN, M. Siri, Siri, in my hand: Who's the fairest in the land? On the interpretations, illustrations, and implications of artificial intelligence. *Business Horizons*, vol. 62, pp. 15-25, jan-fev. 2019.

KELLEHER, J. D.; TIERNEY, B. *Data Science*. Cambridge: MIT Press, 2018.

KROENKE, D. M.; AUER, D. J.; VANDENBERG, S. L.; YODER, R. C. *Database Concepts*. 8ª ed. Nova York: Pearson, 2016.

KWAN, Y.K., LAI, E.L.-C. Intellectual property rights protection and endogenous economic growth. *Journal of Economic Dynamics and Control* 27 (5), 2003, pp. 853–873.

LUNDVALL, B-A; BORRÁS, S. Science, technology and innovation policy. In: FARGERBERG, J., MOWERY, D., NELSON, R. (Ed.). *The Oxford handbook of innovation*. New York: Oxford University Press, 2007.

MCAFEE, A.; BRYNJOLFSSON, E. Big Data: The Management Revolution. *Harvard Business Review*, pp. 1-9, out. 2012.

MURRAY, F.; STERN, S. Do formal intellectual property rights hinder the free flow of scientific knowledge? An empirical test of the anti-commons hypothesis. *Journal of Economic Behavior & Organization*, vol. 63, pp. 648–687, 16 jan. 2007.

NYHAN, J., & DUKE-WILLIAMS, O. Joint and multi-authored publication patterns in digital humanities. *Literary and Linguistic Computing*, 29(3), 387-399, 2014. Disponível em: <https://academic.oup.com/dsh/article/29/3/387/986317>. Acesso em 18 abr. 2022.

OMC. Acordo sobre Aspectos dos Direitos de Propriedade Intelectual Relacionados ao Comércio (TRIPS). 1994. Disponível em: https://www.wto.org/english/docs_e/legal_e/27-trips.pdf. Acesso em 18 abr. 2022.

OMPI. World Intellectual Property Report: The Direction of Innovation. Genebra, 2022.

OSTROM, E.; HESS, C. Private and Common Property Rights. In: Workshop in Political Theory and Policy Analysis. Research Paper no. 2008-11-01. Indiana University, 2007.

REINO UNIDO. Artificial Intelligence and Intellectual Property: copyright and patents: Government response to consultation. 28 jun. 2022. Disponível em <https://www.gov.uk/government/consultations/artificial-intelligence-and-ip-copyright-and-patents/outcome/artificial-intelligence-and-intellectual-property-copyright-and-patents-government-response-to-consultation#text-and-data-mining>. Acesso em 29 jun. 2022.

REINO UNIDO. The Copyright and Rights in Performances (Research, Education, Libraries and Archives) Regulations 2014 (S.I. 2014/1372), regs. 1, 3(2). 2014. Disponível em <https://www.legislation.gov.uk/ukxi/2014/1372/regulation/3/made>. Acesso em 29 jun. 2022.

SOUZA, Allan Rocha de. A Função Social dos Direitos Autorais. 1. ed. Campos dos Goytacazes: Editora da Faculdade de Direito de Campos, 2006.

SOUZA, Allan Rocha de; SCHIRRU, L. A Nova Fronteira Tecnológica do Direito Autoral. In: Manoel Pereira dos Santos; Flávia Mansur Murad Schaal; Rubeny Goulart.. (Org.). Propriedade Intelectual no Mundo da Inteligência Artificial. 1ed. Rio de Janeiro: PI & NewTech Editores, 2021, v. 1, p. 151-194.

SOUZA, A. R.; SCHIRRU, L.; ALVARENGA, M. B. Direitos autorais e mineração de dados e textos no combate à COVID-19 no Brasil. *LiinC em Revista*, vol. 16 (2), p. 1-15, 2020.

SU, F. Cross-national digital humanities research collaborations: structure, patterns and themes. *The Journal of documentation*. 76(6), p. 1295-1312, 2020.

VALENTE, M. G.; ALVARENGA, M. B. **Mineração de dados no descompasso jurídico**. In: SANTOS, Manoel Joaquim Pereira dos; SCHAAL, Flavia Mansur Murad; GOULART, Rubeny (orgs.). **A Propriedade Intelectual e o Mundo da Inteligência Artificial**. 1ed. PI & NewTech Editores, 2021.

WANG, Q. Distribution features and intellectual structures of digital humanities: A bibliometric analysis. *Journal of Documentation*, 74(1), p. 223-246, 2018.

WANG, S.; KANG, B.; MA, J.; ZENG, X.; XIAO, M.; GUO, J.; CAI, M.; YANG, J.; LI, Y. MENG, X.; XU, B. A deep learning algorithm using CT images to screen for Corona Virus Disease (COVID-19). *medRxiv*. 2020. Disponível em: <https://doi.org/10.1101/2020.02.14.20023028>.

WU, T. Law and Economics of Information. In: In: PARISI, F. (ed.). *The Oxford Handbook of Law and Economics: Volume 2: Private and Commercial Law*. Oxford Handbooks, 2017.