

Relatório da OCDE sobre Inteligência Artificial, Data Scraping e os Desafios para a Propriedade Intelectual

abreuadvogados.com



No passado dia 9 de fevereiro de 2025, a Organização para a Cooperação e Desenvolvimento Económico (OCDE) publicou um relatório intitulado "*Intellectual Property Issues in Artificial Intelligence Trained on Scraped Data*", que aborda as implicações da inteligência artificial (IA) no âmbito da propriedade intelectual (PI).

Este relatório foca-se especialmente nos métodos de extração de dados, nomeadamente de *data scraping*, que tem vindo a ser recorrentemente utilizado no treino de sistemas de IA. O uso deste tipo de métodos tem vindo a colocar desafios e problemas jurídicos, em particular, mas não se limitando a, proteção do direito de autor, de marcas, segredos de negócio e bases de dados.

Avanços da IA e desafios que estes avanços colocam

A inteligência artificial depende da utilização de grandes volumes de dados para treinar os seus modelos, sendo que a qualidade desses dados é fundamental para a eficácia dos sistemas. Estes dados podem ter origem em fontes específicas, como bases de dados licenciadas, ou ser recolhidos através de técnicas que têm vindo a ser amplamente utilizadas, como o *data scraping*, que permite a obtenção massiva de informações.

Embora o *scraping* possibilite uma recolha mais ampla e diversificada de dados, ao abranger todos os dados disponíveis, aumenta o risco de violação de direitos de propriedade intelectual (ainda que de forma não intencional, uma vez que o *scraping* pode incluir obras protegidas que se encontram ilicitamente disponíveis). Muitos dos dados recolhidos provêm de obras protegidas por direitos de autor, pertencentes a criadores como fotógrafos, escritores e artistas.

O relatório da OCDE salienta que, apesar da existência desses direitos, técnicas como o *scraping* dificultam o seu exercício pelos autores. Com efeito, a falta de transparência de muitos sistemas é assinalada como um dos principais obstáculos ao conhecimento da violação por parte dos respetivos titulares e, conseqüentemente, ao exercício dos seus direitos.

No referido relatório é ainda mencionado que, de uma forma geral, as legislações de propriedade intelectual, frequentemente anteriores ao desenvolvimento destas novas técnicas e tecnologias, não estão, em muitos casos, preparadas para enfrentar os desafios impostos pelo *scraping*. Como resultado, a aplicação eficaz das normas de PI torna-se difícil, aumentando o risco de violação, mesmo que não intencional, dos direitos de exclusivo.

O Papel da Propriedade Intelectual na Inovação e os Riscos do Data Scraping

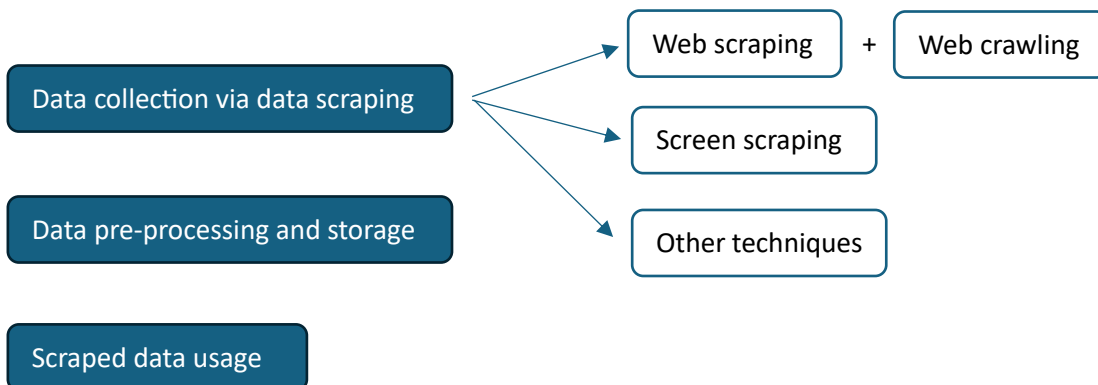
Dado o papel crucial que a PI desempenha na proteção da criatividade e da inovação, o relatório torna evidente a urgência na atualização das legislações nesta nova era digital. Porém, note-se que o equilíbrio entre o desenvolvimento tecnológico e a salvaguarda dos direitos de PI será determinante para assegurar uma inovação responsável e sustentável.

Definição e Importância dos Conceitos Técnicos

Tendo em conta que um dos problemas identificados pelo relatório é precisamente a ausência, por um lado, e a disparidade entre diferentes ordens jurisdicionais, por outro, de conceitos-chave no contexto da IA e da PI, o relatório debruça-se sobre os mesmos e procede a uma proposta dessas definições, dentre as quais a definição de *data scraping*.

No que respeita à definição de *data scraping*, o relatório propõe uma definição de trabalho, que pode ser completada com outras definições focadas em técnicas e/ou atividades específicas utilizadas na recolha de dados. Assim, no presente relatório, o *data scraping* refere-se “à extração automática de dados de treino de IA da internet, de bases de dados em linha e de outras fontes, utilizando ferramentas de software ou scripts automatizados” (tradução nossa).

O relatório densifica, ainda, definições sobre as diferentes técnicas usadas em “Data Scraping”. São elas:



Estes termos são frequentemente utilizados de forma intercambiável, mas, como o relatório destaca, é essencial estabelecer uma terminologia comum para promover a clareza e a consistência, especialmente em termos legais.

Propostas para uma Abordagem Responsável

Para responder aos desafios apresentados pelo *data scraping* e seus impactos na PI, o relatório propõe várias medidas políticas – sublinhando as que já haviam sido propostas

no EASD Recommendation (*OECD, Recommendation of the Council on Enhancing Access to and Sharing of Data, 2021*). Dentre elas, destacam-se:

- **Código de conduta relativo ao *data scraping***
Desenvolver medidas flexíveis e voluntárias que considerem as diferentes abordagens jurídicas e regulamentares entre jurisdições, incluindo a adoção de um "código de conduta" transfronteiriço para a extração de dados e o apoio ao desenvolvimento de ferramentas técnicas e termos contratuais normalizados;
- **Cláusulas contratuais tipo**
Incentivar o desenvolvimento de ferramentas técnicas que protejam os direitos de propriedade intelectual, permitam o controlo de acesso aos dados pelos titulares de direitos e apoiem os mecanismos de licenciamento, incluindo tecnologias de reforço da privacidade;
- **Ferramentas técnicas**
Implementar iniciativas de sensibilização para informar as partes interessadas sobre os seus direitos e responsabilidades, educando os utilizadores de IA para uma utilização responsável;
- **Sensibilização para o *data scraping***
Estabelecer definições e terminologia harmonizadas para as atividades de extração de dados, promovendo a coerência e a compreensão partilhada entre as partes interessadas;

Conclusão

O relatório da OCDE evidencia os desafios legais emergentes no cruzamento entre a evolução dos sistemas de IA e a proteção dos direitos de PI, especialmente no que se refere a técnicas de extração de dados como o *data scraping*.

Este estudo segue os princípios da OCDE para o desenvolvimento de IA, que defendem a defesa da inovação e a promoção de um uso confiável desta tecnologia, respeitando os direitos de PI e promovendo um ambiente jurídico mais coeso e transparente.



Thinking about tomorrow? Let's talk today.

António Andrade – Sócio

antonio.andrade@abreuadvogados.com

Manuel Durães Rocha – Sócio

manuel.rocha@abreuadvogados.com

Ricardo Henriques – Sócio

ricardo.henriques@abreuadvogados.com

Margarida Castillo Silva – Associada

margarida.c.silva@abreuadvogados.com

Catarina Rocha Rodrigues – Advogada Estagiária

catarina.r.rodrigues@abreuadvogados.com