

A SUSCETIBILIDADE DE PROTEÇÃO DE ALGUNS ARTEFACTOS DIGITAIS: ALGORITMOS E MODELOS, APIS, PROMPTS, AI SLOP, DADOS SINTÉTICOS E PESOS DE REDES NEURONAIS ARTIFICIAIS*

THE IP PROTECTION OF SPECIFIC DIGITAL ARTIFACTS: ALGORITHMS AND MODELS, APIS, PROMPTS, AI SLOP, SYNTHETIC DATA AND NEURAL NETWORK WEIGHTS*

NUNO SOUSA E SILVA**

«The story of the computer is made the more fascinating if we realize, at the outset, that it is the nature of thought itself with which we are dealing. Slowly, and perhaps not inexorably, the technology of the computer has moved forward toward a single goal: to make a machine behave “intelligently”.»

MILDE JR., KARL F., “Can a computer be an “Author” or an “Inventor””, *51 Journal of the Patent and Trademark Office Society*, Patent and Trademark Office Society, Buffalo, 1969, p. 382

“We have trapped lightning in a rock and then taught the rock to think. It is black magic of the highest order, no matter what the size!”
Desconhecido, possivelmente de um ficheiro .sig antigo da Usenet

Abstract: This paper examines the intellectual property protection of emerging digital artifacts in the context of AI, including algorithms, AI models, APIs, AI-generated content, prompts, synthetic data, and neural network weights. The research demonstrates that algorithms and AI models, as abstract mathematical formulas, generally fall outside IP protection. While APIs face uncertain protection status, AI-generated works without substantial human creative control are in the public domain. The paper identifies prompt engineering as a potential new form of creative expression, though its classification remains ambiguous between software and literary works. The study proposes that the EU’s sui generis database right may offer protection for synthetic data and neural network weights. However, this protection mechanism creates potential divergence from non-EU jurisdictions lacking equivalent rights. The analysis concludes that protecting AI artifacts may lead to reconceptualising fundamental copyright principles or developing new specialised regimes.

* Este artigo destina-se aos Estudos em Homenagem ao Prof. Marcos Wachowicz.

** Advogado (PTCS) e Prof. Auxiliar da Universidade Católica Portuguesa. E-mail: nsousaesilva@gmail.com

Summary: 1. Introduction. 2. What is a computer programme? 3. Algorithms and models. 4. Application programming interfaces (APIs). 5. AI-generated creations. 6. Prompts and AI Slop. 7. The *sui generis* right of the database manufacturer. 8. Synthetic data. 9. Neural network weights. 10. Conclusion.

Keywords: Artificial intelligence; intellectual property; copyright; APIs; synthetic data; neural networks; Prompts; software; algorithms; models.

Resumo: Este artigo examina a suscetibilidade de proteção de certos artefactos digitais emergentes no contexto da inteligência artificial (IA), incluindo algoritmos, modelos, APIs, conteúdo gerado por IA, *prompts*, dados sintéticos e pesos de redes neurais. A investigação demonstra que algoritmos e modelos de IA, sendo fórmulas matemáticas abstratas, geralmente não são passíveis de proteção. Existe ainda incerteza quanto ao enquadramento de APIs. As obras geradas por IA sem controlo criativo humano substancial estão no domínio público. O artigo identifica a engenharia de *prompts* como uma nova forma potencial de expressão criativa, embora a sua classificação permaneça ambígua entre *software* e obras literárias. O estudo propõe que o direito *sui generis* da UE sobre bases de dados pode oferecer proteção para dados sintéticos e pesos de redes neurais. No entanto, esta via de proteção cria uma potencial divergência em relação a jurisdições fora da UE que não preveem direitos equivalentes. A análise conclui que a proteção de artefactos de IA pode conduzir a uma reconceptualização de princípios fundamentais da Propriedade Intelectual ou ao desenvolvimento de novos regimes especializados.

Sumário: 1. Introdução. 2. O que é um programa de computador? 3. Algoritmos e modelos. 4. Interfaces de programação de aplicações (APIs). 5. Criações geradas por IA. 6. *Prompts* e *AI Slop*. 7. O direito *sui generis* do fabricante de bases de dados. 8. Dados sintéticos. 9. Pesos de redes neuronais. 10. Conclusão.

Palavras-chave: Inteligência artificial; propriedade intelectual; Direito de Autor; APIs; dados sintéticos; redes neuronais; *prompts*; software; algoritmos; modelos.

1. Introdução

Existe um conjunto de objetos, resultados de atividade intelectual humana que são produzidas no contexto da utilização de computadores – as criações informáticas – que levantam dúvidas quanto à respetiva qualificação e suscetibilidade de proteção por direitos exclusivos. O universo das criações informáticas vai muito para além do *software*. Desde logo, podemos ter representações digitais de obras musicais, cinematográficas, escultóricas ou literárias (*e-books*). Além disso, muitas dessas obras “nascem digitalmente”, sendo concebidas, compostas,

gravadas, escritas e/ou editadas *em* computadores ou, mais recentemente, automaticamente *por* computadores¹.

Depois de décadas de discussão, a proteção jurídica dos programas de computador (*software*) parecia estar sedimentada. Até ao fim dos anos 80 do século XX existiu uma intensa discussão quanto à forma mais apropriada de tutela jurídica do *software*². Acabou por singrar a opção pelo Direito de Autor, integrando a proteção numa importante rede de tratados, com especial destaque para a Convenção de Berna de 1886. Desde 1991 que os Estados-Membros da União Europeia estão obrigados a proteger o *software* como obra literária³. Em Portugal esse regime consta do DL 252/94, de 20 de Outubro e no Brasil da Lei nº 9609/98, de 19 de fevereiro. As principais diferenças deste regime, face ao direito de autor comum, dizem respeito às regras de titularidade, aos direitos morais, aos limites e exceções, incluindo utilizações livres e esgotamento⁴.

Nesta década 20 do século XXI surgem novamente dúvidas quanto ao enquadramento de algumas criações digitais. Assim, neste texto iremos analisar sucintamente o estatuto jurídico de algoritmos e modelos, *APIs*, criações geradas por inteligência artificial, incluindo dados sintéticos, pesos de redes neuronais, *prompts*, *vibe coding* e *AI Slop*. Não trato, propositadamente, das várias questões espinhosas relativas à geração destes artefactos⁵. O objetivo deste texto é apenas apresentar breves reflexões sobre a respetiva suscetibilidade de proteção.

¹ Quando falamos de criações informáticas podemos distinguir obras “em formato digital”, meramente representativas de conteúdo (como ficheiros em mp3 com obras musicais ou ficheiros jpeg com obras visuais), das obras “intrinsecamente” digitais, como aquelas que se dirigem ou se relacionam diretamente com o funcionamento do computador. É certo que cada vez mais a fronteira se esbate e a distinção se torna meramente tendencial – as obras do artista turco REFIK ANADOL demonstram a limitada utilidade desta distinção.

² MARCOS WACHOWICZ, *Propriedade Intelectual do Software & Revolução da Tecnologia da Informação*, Juruá Editora, Curitiba, 2004, pp. 29 e ss. JON BING, “Copyright protection of computer programs”, in: *Research Handbook on the Future of EU Copyright* (ed. ESTELLE DERCLAYE), Edward Elgar, Cheltenham, 2009, pp. 401-426.

³ Tal resultou da Diretiva 91/250/CEE, de 14 de maio, a qual veio a ser substituída pela Diretiva 2009/24/CE, de 23 de abril. Esta solução encontra-se igualmente prevista no art. 10º do Acordo TRIPS e no art. 4º do *WIPO Copyright Treaty* de 1996, sendo hoje um padrão mundial. Para o contexto histórico da Diretiva veja-se, por todos, MICHEL WALTER/SILKE VON LEWINSKY, *European Copyright Law: A Commentary*, Oxford University Press, Oxford, 2010, pp. 89 e ss.

⁴ MARCO ANTONIO MARISCAL MORAZA, *Protección Jurídica del Software*, Reus, Madrid, 2022. Em relação ao Brasil, MANOEL J. PEREIRA SANTOS, “Programas de Computador”, in: *Manual de Propriedade Intelectual* (Coord. Luiz Claudio Garé et al.), Almedina Brasil, São Paulo, 2023 e MARCOS WACHOWICZ, *cit.*, pp. 130 e ss.

⁵ Sobre isso, cfr., por todos, JOÃO PEDRO QUINTAIS, “Generative AI, copyright and the AI Act”, *Computer Law & Security Review*, Vol. 56 (2025), p. 106107.

2. O que é um programa de computador?

No livro Verde da Comissão Europeia de 1988⁶ usava-se a seguinte definição de programa de computador: “um conjunto de instruções cujo propósito é levar um aparelho de processamento de informação a realizar as suas funções”. No entanto, nem o legislador europeu nem o legislador português definiram “programas de computador”. Apesar disso, a jurisprudência portuguesa já definiu programa de computador como uma “pré-listagem de instruções que precede a introdução dos dados, instruções essas que são destinadas a orientar a ação do computador relativamente ao material informativo que se pretende que seja processado”, o que está essencialmente correto⁷. Na mesma linha, o legislador brasileiro define programa de computador como “expressão de um conjunto organizado de instruções em linguagem natural ou codificada, contida em suporte físico de qualquer natureza, de emprego necessário em máquinas automáticas de tratamento da informação, dispositivos, instrumentos ou equipamentos periféricos, baseados em técnica digital ou análoga, para fazê-los funcionar de modo e para fins determinados” (art. 1º da Lei nº 9609/98).

O programa de computador corresponde essencialmente ao código-fonte (escrito por programadores em linguagem de programação) e código-objeto, também designado código-máquina (o conjunto de instruções elétricas que o processador vai executar)⁸. A jurisprudência europeia já esclareceu que a

⁶ *Green Paper on Copyright and the Challenge of Technology – Copyright Issues Requiring Immediate Action* (COM (1988) 172 final) p. 170. A lei norte-americana usa uma definição próxima: “a set of statements or instructions to be used directly or indirectly in a computer in order to bring about a certain result”. 17 USC § 101.

⁷ Ac. TRL, de 16.1.2014, proc. 113/13.9YHLSB-A.L1-6. Uma excelente descrição de programa de computador é a seguinte: “At the core of the concept is the notion of a program which can be executed by a computer. The program will then have to take the form of a set of instructions conforming to the formalism of a programming language. The language consists of a certain predefined set of commands, which semantic is defined in detail with respect to the actions they will cause in a computer. The language will also have a defined syntax which has to be followed strictly for the program to be executed, or executed correctly. The program will permit variables which may be chosen by the programmer with considerable freedom (...) The notion of a computer program is centered on the version in source code. This is a program written in a high-level language (...) Typically, the program in high-level language will be run through another program to replace the high-level instructions by the machine-related detailed instructions necessary to run the computer. This program is called a compiler. In this the simple sentences of high-level language are exploded into a very high number of primitive instructions implied by the definitions. The result is the object program, which may govern the functions of a computer”, cfr. JON BING, *cit.*, pp. 403-405.

⁸ Existem diferentes processos de transformação de código-fonte em código objeto: i) compilação, em que o código-fonte é integralmente traduzido para código-máquina por um compilador (isto é, um programa de computador que vai transpor o código escrito por seres humanos para código binário), gerando um ficheiro executável (exemplos: C, C++, Rust); ii) interpretação, em que o código-fonte é executado linha

proteção jusautorai se encontra limitada ao código-fonte e o código-objeto, únicas formas de expressão de um programa de computador, não se estendendo a outros elementos⁹.

Além do programa de computador propriamente dito, o *material de concepção preparatório* do programa é equiparado ao programa (art.1º/3 do DL 252/94 e art. 1º/1 da Diretiva). No entanto, a lei não define material de concepção preliminar. Em geral, entende-se que está em causa a chamada “lógica de negócio” (*business logic*) e a “arquitetura” de um programa de computador, incluindo esquemas, diagramas e a descrição de sequência de passos, que podem levar à criação de um programa de computador¹⁰. Isto significa que alguém que não domine a linguagem de programação utilizada, mas que conceba o sistema em termos técnicos, poderá ser autor de um programa de computador¹¹.

3. Algoritmos e modelos

Um *algoritmo* é um processo ou conjunto de regras para resolver uma classe de problemas¹². Quando, na escola primária, se ensinam as formas de realizar quatro operações aritméticas básicas, o que se ensina são algoritmos. Os algoritmos para a divisão, multiplicação, soma ou subtração são, precisamente, um conjunto de passos (abstratamente formulados) que nos permite chegar a um resultado.

Em contrapartida, um *modelo* é uma representação abstrata e simplificada de um sistema ou processo, construída para compreender, explicar ou prever comportamentos. No contexto da inteligência artificial, os modelos são fórmulas matemáticas que processam dados de entrada para produzir resultados espe-

por linha por um interpretador, sem gerar código-máquina permanente (exemplos: Python, JavaScript no browser); iii) a chamada abordagem híbrida, em que código é compilado para um formato intermédio (bytecode) que é depois interpretado ou compilado “just-in-time” durante a execução (exemplos: Java, C#) e a iv) transpilação, implicando a conversão de uma linguagem de alto nível para outra (exemplo: TypeScript para JavaScript).

⁹ Confirmando-o numa decisão recente cfr. C-159/23, *Sony* (EU:C:2024:887), §34 e 37.

¹⁰ BENTLY/YIN-HARN, *Concise European Copyright Law*, Wolters Kluwer, Alphe an den Rijn, 2016, p. 244. Em sentido contrário e muito crítico da jurisprudência europeia cfr. JOSÉ ALBERTO VIEIRA, *Direito de Autor: Dogmática Básica*, Almedina, Coimbra, 2020, pp. 514-515.

¹¹ A distinção entre o material de concepção preliminar e uma simples obra literária (sujeita ao regime geral) está na proximidade. O material tem de ser... preliminar, isto é, numa forma próxima e conducente à elaboração de código. Em 2018 o Supremo tribunal holandês (NL:GHDHA:2016:217) explicou que o critério é a quantidade de trabalho ou esforço necessário para gerar código a partir do material preliminar, se for significativo então não estará abrangido pelo regime especial dos programas de computador.

¹² PETER SLOWINSKI, “Rethinking Software Protection”, in: *Artificial Intelligence and Intellectual Property* (eds. JYH-AN LEE/RETO HILTY/KUNG-CHUNG LIU), Oxford University Press, Oxford, 2021, p. 346.

cíficos. Ao contrário dos algoritmos tradicionais, que seguem sequências de instruções determinísticas, os modelos de IA – especialmente os baseados em aprendizagem automática (*machine learning*) – são “treinados” com dados para desenvolver a capacidade de fazer previsões ou classificações¹³. Um modelo de IA contemporâneo, como os *Large Language Models* (p.e., GPT5, Llama, Gemini ou Claude) não é escrito linha por linha por programadores humanos, mas emerge de um processo automatizado de treino sobre vastas quantidades de dados. Este processo resulta numa rede neuronal artificial com milhões ou biliões de parâmetros (pesos) que representam matematicamente as conexões entre os neurónios artificiais. Um modelo não é mais do que uma fórmula matemática¹⁴.

Os algoritmos e modelos são implementados em programas de computador através de código. Este é escrito numa linguagem de programação (código-fonte) e transformado em código-objeto (também designado código-máquina) para ser lido pelo processador e, assim, realizar as operações e funções pretendidas. Os mesmos algoritmos e modelos podem ser implementados com código muito diferente, da mesma forma que podemos todos contar essencialmente a mesma história (os três porquinhos, Romeu e Julieta ou a Cinderela) usando palavras e frases marcadamente distintas.

Ora, um princípio universal de Direito de Autor, consagrado no art. 9º/2 do Acordo TRIPS, é o princípio da *dicotomia ideia/expressão*. O Direito de Autor apenas protege a expressão, não a ideia. Uma ideia, em si mesma, não é passível de apropriação individual. Assim, no caso de um livro ou filme de ficção não é a história que é objeto de proteção, mas antes a forma como ela é contada; não é a composição (abstrata) dos objetos ou pessoas que é protegida numa pintura ou fotografia, mas sim a expressão que lhe é dada. O direito de autor não protege, por isso, teorias, métodos, estilo, materiais, dados ou factos¹⁵. Traçar a linha entre o que constitui uma ideia e a sua expressão pode revelar-se difícil, até porque para comunicar uma ideia

¹³ Sendo que, como é evidente, há algoritmos de treino e os modelos podem representar um algoritmo e vice-versa. O Regulamento Europeu de IA (Reg. 2024/1689) adotou a seguinte definição de “sistemas de Inteligência Artificial”, que se baseia na definição atualizada da OCDE: “um sistema baseado em máquinas concebido para funcionar com níveis de autonomia variáveis, e que pode apresentar capacidade de adaptação após a implantação e que, para objetivos explícitos ou implícitos, e com base nos dados de entrada que recebe, infere a forma de gerar resultados, tais como previsões, conteúdos, recomendações ou decisões que podem influenciar ambientes físicos ou virtuais” (art. 3º/1).

¹⁴ PETER SLOWINSKI, *cit.*, p. 347.

¹⁵ Art. 9º/2 TRIPS e, em Portugal art. 1º/2 CDADC. Para explicar a dicotomia ideia-expressão ANDREAS RAHMATIAN, *Copyright and Creativity. The Making of Property Rights in Creative Works*, Edward Elgar, Cheltenham, 2011, pp. 131 e ss. apresenta exemplos ilustrativos: a descrição verbal de uma pintura refere-se às ideias que lá estão, a pintura é a expressão dessas ideias; uma carta de Mozart explicando ao seu pai como pensa resolver certas questões relativas à composição de uma ópera (designadamente

é sempre necessário dar-lhe alguma expressão. Por isso mesmo, existe a chamada *merger doctrine*, segundo a qual não subsistirão direitos de autor numa dada expressão se esta for a única forma de exprimir a ideia que lhe subjaz¹⁶.

Relativamente aos programas de computador este problema agudiza-se visto que todas as suas expressões são funcionais¹⁷. O *software* é por natureza uma obra utilitária. No entanto, sublinhe-se que o carácter funcional ou utilitário de uma criação não constitui óbice à sua protecção por direito de autor¹⁸. O que se exige é apenas originalidade e exteriorização.

Originalidade quer dizer que foi o autor quem criou a obra, que esta *tem a sua origem* no seu trabalho. Nesse sentido significa apenas que uma obra não foi copiada. Mas também quer dizer que há criatividade: escolhas conscientes e criativas por parte de um ser humano que constituem uma expressão da personalidade de um autor¹⁹. Desde o acórdão *Infopaq* que o TJ vai revisitando a noção de obra, mantendo como critério único de originalidade a fórmula “criação intelectual do seu autor”²⁰. Ou seja, se existir uma margem de criatividade suficiente para que a criação, o conjunto de escolhas feitas por uma ou mais pessoas, apresente uma marca de personalidade (criatividade), então a obra será protegida²¹.

a utilização de música turca, o recurso a mudanças de andamento e de tom) constituem ideias, a pauta conterà a respetiva expressão/concretização.

¹⁶ NOAM SHEMTOV, *Beyond the Code: Protection of Non-Textual Features of Software*, Oxford University Press, Oxford, 2017, pp. 102-135. Um outro conceito relacionado é o de *scenes à faire*: os elementos comuns e estereótipos considerados necessários ou típicos de um determinado género, cenário ou tema não são protegidos por direitos de autor, sendo considerados básicos, comuns ou inevitáveis. Um exemplo pitoresco da aplicação deste conceito encontra-se na decisão *Walker v. Time Life Films, Inc.*, 784 F.2d 44 (2d Cir. 1986), em que o tribunal norte-americano concluiu que um filme sobre o trabalho de polícias no sul do Bronx “would need to feature drunks, prostitutes, vermin, and derelict cars to be perceived as realistic”. Nesse sentido, a repetição desses elementos em obras posteriores não constitui violação de direito de autor.

¹⁷ JOSÉ OLIVEIRA ASCENSÃO, “Direito de Autor e Informática Jurídica”, in: *Estudos sobre Direito da Internet e da Sociedade da Informação* (coord. JOSÉ OLIVEIRA ASCENSÃO), Almedina, Coimbra, 2001, pp. 9-10.

¹⁸ C-833/18, *Brompton Bicycle*, §38: “... a protecção (...) ao abrigo do direito de autor aplica[-se] a um produto cuja forma é, pelo menos em parte, necessária à obtenção de um resultado técnico quando esse produto constitua uma obra original resultante de uma criação intelectual, na medida em que, através dessa forma, o seu autor exprime a sua capacidade criativa de modo original, efetuando escolhas livres e criativas que refletem na referida forma a sua personalidade, o que cabe ao órgão jurisdicional nacional verificar tendo em conta o conjunto dos elementos pertinentes do litígio no processo principal”.

¹⁹ TATIANA-ELENI SYNODINOU, “The foundations of the concept of work in European Copyright Law”, in: *Codification of European Copyright Law* (ed. TATIANA-ELIANA SYNODINOU), Kluwer Law International, Alphen aan den Rijn, 2012, pp. 95-96.

²⁰ C-5/08, *Infopaq* (EU:C:2009:465). Sobre esta jurisprudência cfr. NUNO SOUSA E SILVA, “Uma introdução ao direito de autor europeu”, *Revista da Ordem dos Advogados*, (2013), pp. 1365-1373.

²¹ ESTELLE DERCLAYE, “Infopaq International A/S v Danske Dagblades Forening (C-5/08): wonderful or worrisome? The impact of the ECJ ruling in Infopaq on UK copyright law”, *European Intellectual Property Review*, (2010), p. 249.

É importante realçar que a originalidade se refere à expressão (de um dado tipo) e não às ideias. Nesse sentido, o art. 1º/2 da Diretiva 2009/24, transposto no art. 2º do DL 252/94, estabelece que “...a proteção abrange a expressão, sob qualquer forma, de um programa de computador. As ideias e princípios subjacentes a qualquer elemento de um programa de computador, incluindo os que estão na base das respetivas interfaces, não são protegidos pelos direitos de autor ao abrigo da presente diretiva”²².

Em 2011, no processo C-393/09, *BSA*, o Tribunal de Justiça foi confrontado com a questão de saber se uma interface gráfica poderia ser protegida por via de direitos de autor. Tendo negado a proteção por via do direito de autor “especial” para programas de computador (visto que a interface gráfica não constitui a expressão de um programa de computador) o Tribunal afirmou que a proteção por via do direito de autor “comum” poderia ser concedida desde que essa obra fosse original²³. Referindo-se então à determinação da originalidade afirmou que “este critério não pode ser preenchido pelos componentes da interface gráfica do utilizador que se caracterizam unicamente pela sua função técnica”²⁴ visto que “quando a expressão dos referidos componentes resulta da sua função técnica, o critério da originalidade não se encontra preenchido, porque as diferentes formas de executar uma ideia são tão limitadas que a ideia e a expressão se confundem”²⁵; assim “...os componentes da interface gráfica do utilizador não permitem ao autor exprimir o seu espírito criador de modo original e chegar a um resultado que constitua uma criação intelectual desse autor”²⁶. Por outras palavras, a interface gráfica não constitui, para este efeito, *software* e apenas poderá ser protegida quando a sua conceção não seja determinada apenas por fatores técnicos ou funcionais²⁷.

Na mesma linha, no acórdão de 2012, C-406/10, *SAS* o Tribunal de Justiça da União Europeia concluiu: “no que respeita aos elementos de um programa de computador (...) nem a funcionalidade de um programa nem a linguagem de programação e o formato de ficheiros de dados utilizados no âmbito de um

²² Vide também os considerandos 10 e 11. Curiosamente a transposição portuguesa não refere expressamente interfaces, mas antes “... a liberdade das ideias e dos princípios que estão na base de qualquer elemento do programa ou da sua interoperabilidade” (art. 2º/2 do DL 252/94).

²³ C-393/09, *BSA* (EU:C:2010:816), §42-46. Sobre as interfaces gráficas em detalhe cfr. NOAM SHEMTOV, “Software and graphical user interfaces”, in: *Research Handbook on Intellectual Property and Digital Technologies* (Ed. TANYA APLIN), Edward Elgar, Cheltenham, 2020, pp 2-25.

²⁴ C-393/09, *BSA*, §48.

²⁵ C-393/09, *BSA*, §49. Isto expressa a já mencionada *merger doctrine*. Na mesma linha veja-se C-833/18, *Brompton Bicycle* (EU:C:2020:461), §27.

²⁶ C-393/09, *BSA*, §50.

²⁷ Para um exemplo de jurisprudência que protegeu as interfaces gráficas ao abrigo do direito de autor comum pode ver-se, em Inglaterra, *THJ Systems Limited & Anor v Daniel Sheridan & Anor* [2023] EWCA Civ 1354. Na mesma linha face ao direito brasileiro cfr. MANOEL J. PEREIRA SANTOS, *cit.*, pp. 349-350.

programa de computador para explorar algumas das suas funções constituem uma forma de expressão desse programa”²⁸.

A jurisprudência portuguesa, no acórdão da Relação de Lisboa de 6 de Abril de 2021²⁹, já se pronunciou nesse sentido, concluindo: “o algoritmo não é um «bem» tutelável no domínio dos direitos de autor, estando excluído da proteção conferida aos programas de computador” visto que “o direito de autor não protege funcionalidades”. Como explica JOSÉ ALBERTO VIEIRA, os elementos não expressivos, como as estruturas de dados, a algoritmia ou a lógica de um programa de computador não são protegidos³⁰.

O mesmo acontece no contexto das patentes. O artigo 52^o/2 da Convenção da Patente Europeia esclarece o conteúdo do conceito “invenção”³¹. Este artigo deixa claro que descobertas, teorias e métodos, materiais já existentes na natureza, criações estéticas, jogos, métodos comerciais, programas de computador (enquanto tais). Nessa linha entende-se na Europa que os algoritmos em si, por serem métodos matemáticos abstractos, não são protegidos. E, após alguma hesitação, parece, desde o caso *Alice v. CLS Bank* decidido pelo *Supreme Court* em 2014, que vale igual conclusão nos EUA. Também por isso parece ser assente que os métodos gerais de desenvolvimento de *software* de inteligência artificial não serão patenteáveis.

Em suma, os algoritmos, por serem ideias, métodos e processos abstratos, tipicamente expressos por fórmulas matemáticas não são passíveis de proteção por direito de autor (visto constituírem ideias)³² ou por patente (em virtude da falta de carácter técnico)³³. No entanto, certas implementações de algoritmia podem ser patenteáveis desde que apresentem carácter técnico³⁴. Esta circunstância leva a que a via mais atrativa para a proteção de algoritmos seja o segredo³⁵.

²⁸ C-406/10, *SAS* (EU:C:2012:259), §39. Admite-se, porém, a tutela jusautoral pelo direito de autor comum de algumas destas criações, como a linguagem de programação e o formato de ficheiros (C-406/10 *SAS*, §45), o manual de instruções (C-406/10 *SAS*, §64) e os elementos gráficos e sonoros de um videojogo (C-355/12, *Nintendo* (EU:C:2014:25), §23) desde que sejam originais.

²⁹ Proc. 55/19.4YHLSB.L1-PICRS, rel. Isoleta Costa.

³⁰ JOSÉ ALBERTO VIEIRA, *Direito de Autor...*, cit., p. 528.

³¹ Fala-se a este propósito de uma “delimitação negativa do conceito de invenção”. MAXIMILIAN HAEDICKE /HENRIK TIMMANN, *Patent Law: A handbook*, C.H. Beck Hart Nomos, 2014, p. 96.

³² PETER SLOWINSKI, cit., p. 347.

³³ MAXIMILIAN HAEDICKE/HENRIK TIMMANN, cit., p. 102: “A manner of proceeding, or a procedure in the field of mathematics is also called a (mathematical) algorithm. (...) mathematics is not a technical field in the meaning of patent law”. No caso *Gottschalk v. Benson*, 409 U.S. 63, o *Supreme Court* norte-americano rejeitou a patenteabilidade de um método numérico visto que “the patent would wholly pre-empt the mathematical formula and in practical effect would be a patent on the algorithm itself”.

³⁴ MAXIMILIAN HAEDICKE/HENRIK TIMMANN, cit., p. 103.

³⁵ KATHARINA SCHEJA, “Schutz von Algorithmen in Big Data Anwendungen”, *Computer und Recht*, (2018), pp. 487-488.

Pelas mesmas razões, os modelos de inteligência artificial não são, por si só, objeto de proteção por direito de autor ou patenteáveis³⁶. Também aqui o segredo é a via de proteção mais viável³⁷.

4. Interfaces de programação de aplicações (APIs)

As APIs (interfaces de programação de aplicações) são uma parte essencial do funcionamento dos sistemas informáticos contemporâneos, assegurando a fácil comunicação e interoperabilidade entre diferentes programas de computador. Aquele que desenvolve um programa ou função informática frequentemente disponibiliza uma ou várias APIs, gratuitamente ou em troca de um pagamento, permitindo a terceiros utilizar esse programa ou função. É graças a esta tecnologia que é possível criar uma conta num *site* de um terceiro com um perfil de Facebook ou uma conta Google, pagar com Paypal numa loja *online* ou integrar informação financeira de várias contas bancárias numa única aplicação³⁸.

A 5 de Abril de 2021 o *Supreme Court* norte-americano decidiu a título definitivo um litígio sobre APIs que opôs a Google e a Oracle durante dez anos e que esteve no centro das preocupações da indústria do *software*³⁹. O caso lidava com a reprodução das declarações das APIs de Java no desenvolvimento do Android. Em vez de desenvolver novos métodos, a Google copiou cerca de 11.500 linhas de código de declaração da API de Java (os nomes, parâmetros e estrutura das funções), na altura a linguagem de programação mais utilizada no mundo. O objetivo era garantir que os programadores familiarizados com a programação em Java tinham facilidade em utilizar Android. A Oracle processou a Google por violação de direitos de autor, argumentando que esta cópia violava os direitos de autor da Oracle sobre a plataforma Java, enquanto a Google argumentou que as declarações de uma API não deviam ser passíveis de proteção jusautorais e que a sua utilização constituía *fair use*.

³⁶ KATARINA FOSS-SOLBREKK, “Three routes to protecting AI systems and their algorithms under IP law: The good, the bad and the ugly”, *Journal of Intellectual Property Law & Practice*, Vol. 16, No. 3, (2021), pp. 247-258. De acordo com as *Guidelines* do IEP (Parte G 3.3.1): “Artificial intelligence and machine learning are based on computational models and algorithms for classification, clustering, regression and dimensionality reduction, such as neural networks, genetic algorithms, support vector machines, k-means, kernel regression and discriminant analysis. Such computational models and algorithms are per se of an abstract mathematical nature, irrespective of whether they can be “trained” based on training data”.

³⁷ SIMON APEL/MARKUS KAULARTZ, “Rechtlicher Schutz von Machine Learning-Modellen”, *Recht Digital*, (2020), pp. 24-34.

³⁸ NUNO SOUSA E SILVA, “Ligações perigosas? – Reflexões sobre APIs e Direito de Autor a partir do acórdão Google v. Oracle do Supremo Tribunal dos EUA”, *Revista de Direito Intelectual*, 1 (2022), pp. 213-227.

³⁹ *Google LLC v. Oracle America, Inc.*, 593 U.S. 1 (2021).

O acórdão tranquilizou a generalidade dos programadores e empresas produtoras de *software*, tendo legitimado aquela que tem sido a prática habitual – a livre utilização de *APIs*⁴⁰. No entanto, o Tribunal não tomou posição quanto à possibilidade de proteção das *APIs* ao abrigo do direito norte-americano.

Na União Europeia a questão também não é pacífica⁴¹. A meu ver, a exclusão das interfaces do direito de autor para *software* parece resultar do considerando II da Diretiva 2009/24⁴². Este entendimento é reforçado pela decisão do Tribunal de Justiça no acórdão *Sony* onde se lê que “um elemento desse programa através do qual os utilizadores exploram as funcionalidades do referido programa, não constitui uma forma de expressão de um programa de computador”⁴³. Pode afirmar-se que uma *API* também constitui um elemento do programa através do qual os utilizadores exploram as suas funcionalidades. Por último, a proteção por direito de autor comum exigiria originalidade das interfaces, a qual tenderá a ser afastada em virtude da respetiva (estrita) funcionalidade⁴⁴.

5. Criações geradas por IA

Atualmente muitos sistemas de IA produzem resultados que seriam indiscutivelmente passíveis de proteção por direitos de Propriedade Intelectual se fossem criados por seres humanos. Os exemplos são infindáveis.

⁴⁰ Assim, a propósito da decisão em primeira instância, PETER S. MENELL, “Rise of the API Copyright Dead?: An Updated Epitaph for Copyright Protection of Network and Functional Features of Computer Software”, *Harvard Journal of Law & Technology*, vol. 31, (2018), p. 386: “the ruling validated what was considered a best practice”. Também MARK LEMLEY/PAMELA SAMUELSON, “Interfaces and Interoperability After Google v. Oracle”, *Texas Law Review* 100, (2021), p. 2: “...important victory for software developers and for an open internet”.

⁴¹ RICHARD ARNOLD, “Copyright in software: functionality”, in: *Research Handbook on Intellectual Property and Digital Technologies* (Ed. TANYA APLIN), Edward Elgar, Cheltenham, 2020, p. 41.

⁴² Também assim *Cloud Computing Law* (ed. CHRISTOPHER MILLARD), Oxford University Press, Oxford, 2021, p. 165. Com uma posição intermédia PAMELA SAMUELSON, “The Past, Present and Future of Software Copyright Interoperability Rules in the European Union and United States”, *European Intellectual Property Review*, (2010), p. 232: “The Software Directive does not categorically exclude interfaces from the scope of copyright’s protection, although it anticipates that interfaces may be among the unprotectable elements of programs as ideas or principles”. Em sentido contrário, cfr. Opinião do AG YVES BOT, no caso C-406/10, *SAS Institute* (EU:C:2011:787), §81: “Parece-me que a Diretiva 91/250 não exclui as interfaces da proteção pelos direitos de autor”.

⁴³ C-159/23, *Sony*, §35.

⁴⁴ Nesse sentido BENTLY/YIN-HARN, *cit.*, p. 245. Foi também esse o entendimento dos tribunais ingleses nas decisões *SAS Institute Inc v World Programming Ltd* [2010] EWHC 1829 (Ch) de 23 de julho de 2010 e *SAS Institute Inc v World Programming Ltd* [2013] EWHC 69 (Ch), de 25 de janeiro de 2013.

Para enquadrar este tema dos resultados do processamento (*outputs*) temos de fazer distinções a dois níveis: primeiro é importante distinguir criações autonomamente geradas por IA de criações feitas com assistência de IA⁴⁵; segundo, a relevância de uma presença humana no processo gerador de um bem imaterial difere consoante os regimes jurídicos da Propriedade Intelectual.

Por enquanto parece claro que, não havendo “direitos sem sujeito” a IA não poderá ser titular de qualquer direito de propriedade intelectual⁴⁶. Esta abordagem foi confirmada pelo Instituto Europeu de Patentes no caso DABUS⁴⁷. Assim, em relação a criações imateriais geradas por ou através de IA a dicotomia será apenas entre a possibilidade de proteção por direitos exclusivos atribuídos a pessoas singulares ou coletivas, ou a respetiva inclusão no domínio público.

No caso das invenções obtidas com recurso a Inteligência Artificial em que a atividade inventiva (“o processo cognitivo”) não é totalmente desenvolvida por um ser humano, mas é despoletada por este, o problema não é tão premente⁴⁸. Do ponto de vista do direito de patentes, o processo é irrelevante, importando apenas o resultado⁴⁹. O atual sistema de patentes pressupõe uma intervenção humana, mas não exige que o processo cognitivo seja humano, a intervenção pode residir apenas em iniciar o processo inventivo. Além disso, por imposição internacional, as patentes devem estar disponíveis para “todos os campos da tecnologia” (art. 27º/1 TRIPS) e independentemente da forma como foram obtidas. A questão residirá em saber se o resultado (a invenção) reúne os requisitos de patenteabilidade, em especial saber se tem altura inventiva.

O que se concluiu em relação às patentes será aplicável, nos sistemas jurídicos que os prevejam, aos modelos de utilidade. Quanto a desenhos ou modelos, será irrelevante para a respetiva proteção o meio de obtenção da criação; desde que o desenho ou modelo seja novo e possua caráter singular, será suscetível de proteção. De igual modo, os sinais distintivos podem ser gerados automática e autonomamente por sistemas de IA (*v.g.*, Looka ou Logoi.com) sem que esse

⁴⁵ TITO RENDAS, “Da suscetibilidade de proteção jusautor de obras geradas por sistemas de inteligência artificial”, in: *Católica Talks: Direito e Tecnologia* (Coord. ELSA VAZ DE SEQUEIRA), Universidade Católica Editora, Lisboa, 2021, pp. 129-144.

⁴⁶ NUNO SOUSA E SILVA, “Direito e Robótica: Uma primeira aproximação”, *Revista da Ordem dos Advogados*, (2017), pp. 485-551.

⁴⁷ Decisão de 21 de Dezembro de 2021, J 0008/20, DABUS (EP:BA:2021:J000820.20211221). Sobre o tema cfr. JOÃO DE OLIVEIRA GERALDES, “Sobre novos desafios do direito industrial: patentes, Inteligência Artificial e o caso Dabus”, *Revista de Direito Comercial*, (2022), pp. 1941-1996.

⁴⁸ DARIA KIM, “AI-Generated Inventions’: time to get the record straight?”, *GRUR-Int*, (2020), pp. 443-456.

⁴⁹ PETER BLOK, “The inventor’s new tool: artificial intelligence – how does it fit in the European patent system?”, *European Intellectual Property Review*, (2017), pp. 70-72.

facto afete a sua possibilidade de registo. Assim, conclui-se que, em relação ao Direito Industrial, a circunstância de o bem imaterial ter sido gerado por ou com recurso a sistemas de IA não afeta a suscetibilidade de proteção⁵⁰.

Ao contrário do que acontece no Reino Unido, na Índia ou na Irlanda, que reconhecem direitos de autor para “computer-generated works”, em Portugal e no Brasil, as criações geradas por computador, quando não possam ser imputadas a uma ou mais pessoas, não são protegidas⁵¹. A obra protegida é definida como a expressão de um ser humano, de cariz criativo, isto é, apresentando originalidade. Esta posição é reafirmada pela jurisprudência do Tribunal de Justiça, que consistentemente exige que haja uma “criação intelectual do próprio autor” para a proteção jusautoral⁵².

A autoria pressupõe um controlo do processo criativo, fundamento para uma imputação subjetiva do resultado a uma ou mais pessoas. Não é, porém, necessário que o autor tenha um controlo sobre todas as variáveis que geram a obra, exigindo-se apenas um mínimo de controlo. Assim, as obras chamadas de arte aleatória podem ser protegidas, sendo o seu autor aquele que controla minimamente o resultado e a respetiva apresentação⁵³. A criação ocorre na disposição do sistema gerador e na possibilidade de selecionar, compor/editar, aceitar ou rejeitar os resultados aleatoriamente produzidos. Nesse sentido, a atividade criativa precede e sucede ao acto de geração do som, cor ou imagem – mas a apresentação do resultado como “seu”, a adesão pessoal ao que a álea produziu basta para estarmos perante uma criação daquela pessoa, o autor⁵⁴.

Naturalmente, o criador (ser humano) pode servir-se de meios informáticos (como CAD, um processador de texto ou um sintetizador) para criar e, nesse caso, a criação continuará a ser-lhe imputada, adquirindo o ser humano que utiliza o programa, na sua qualidade de criador, o direito de autor. Em contrapartida,

⁵⁰ NUNO SOUSA E SILVA, “Inteligência Artificial e Propriedade Intelectual: está tudo bem?”, in: AAVV, *I Congresso Inteligência Artificial e Direito*, Almedina, Coimbra, 2023, pp. 201-220.

⁵¹ JOSÉ ALBERTO VIEIRA, “Obras geradas por computador e direito de autor”, in: AAVV, *Direito da Sociedade da Informação*, vol. II, Coimbra Editora, Coimbra, 2001, pp. 132-133.

⁵² Cfr. desde C-5/08, *Infopaq* (EU:C:2009:465) até C-227/23, *Kwantum* (EU:C:2024:914).

⁵³ NUNO SOUSA E SILVA, “Comentário à intervenção e texto de Tito Rendas “Da suscetibilidade de proteção jusautoral de obras geradas por sistemas de inteligência artificial””, in: AAVV, *Católica Talks: Direito e Tecnologia*, Universidade Católica Editora, Lisboa, 2021, pp. 146-153.

⁵⁴ Isto aproxima-se do que diz o Tribunal de Justiça em relação à fotografia: “No estádio da fase preparatória, o autor poderá escolher o pano de fundo, a pose da pessoa a fotografar ou a iluminação. No momento em que tira o retrato fotográfico, poderá escolher o enquadramento, o ângulo de que deve ser tirado ou ainda a atmosfera criada. Por último, durante a revelação, o autor poderá escolher a técnica que deseja adoptar de entre as diversas existentes, ou ainda, se for caso disso, utilizar aplicações informáticas. Através dessas diferentes escolhas, o autor de um retrato fotográfico pode, assim, imprimir o seu «toque pessoal» à obra criada” (C-145/10, *Painer*, EU:C:2011:798, §91-92).

sempre que se produza um resultado que não possa ser imputado ao controlo de pelo menos um ser humano, então não estaremos perante uma criação de um autor⁵⁵. Logo, esse resultado, ainda que artística e/ou monetariamente muito valioso, será irrelevante do ponto de vista jusautorais⁵⁶.

Ou seja, a distinção essencial a fazer é entre criações autónomas de sistemas de Inteligência Artificial, isto é, aquelas em que não há contributo humano relevante, e criações em que os sistemas funcionam como ferramentas ao serviço de um criador. Se é verdade que muitos destes sistemas conseguem produzir criações de forma totalmente autónoma, aquilo que se verifica é que muitos desses sistemas (DALL-E, Midjourney, etc.) tendem a ser utilizados como ferramentas. Os criadores tornam-se especialistas em *prompts*, que nalguns casos funcionam quase como uma linguagem de programação, ou convocam outras formas de interação com estas ferramentas, manipulam, revêm ou validam os resultados autonomamente produzidos, mantendo (algum) controlo sobre o processo criativo⁵⁷. Se o sistema de IA gozasse de personalidade jurídica, esse processo criativo seria equivalente ao de criação conjunta ou coautoria. Nesse sentido, uma grande parte destas situações (provavelmente as mais relevantes do ponto de vista económico) continuam a dever ser qualificadas como criações humanas e, como tal, suscetíveis de proteção por direito de autor.

Já os resultados autonomamente gerados, sem intervenção humana relevante, integram o domínio público, não gozando de qualquer proteção⁵⁸.

Esta distinção, válida, nem sempre será fácil de aplicar. Se nos casos de colaboração humana se exige um contributo expressivo para acesso à coautoria, já não é evidente que o mesmo possa ser transposto para o contexto da IA generativa. No exemplo de RAHMATIAN referido *supra* (nota 15), o pai de Mozart limita-se a fazer-lhe sugestões no plano das ideias – utilização de música turca, o recurso a mudanças de andamento e de tom –, o que não o qualifica como coautor. Ora, a utilização dos sistemas de IA generativa passa em grande medida por formular sugestões e pedidos essencialmente no plano das ideias, através dos designados

⁵⁵ JANE C. GINSBURG, “People not machines: authorship and what it means in the Berne Convention”, *IIC*, (2018), p. 133: “If the human intervention in producing these outputs does not exceed requesting the computer to generate a literary, artistic or musical composition of a particular style or genre, one may properly consider these works to be “computer-generated” because the human users do not contribute sufficient “intellectual creation” to meet minimum standards of authorship under the Berne Convention”.

⁵⁶ BERNT HUGENHOLTZ/JOÃO PEDRO QUINTAIS, “Copyright and artificial creation: Does EU Copyright Law protect AI-assisted output?”, *IIC*, (2021), pp. 1195-1196.

⁵⁷ BERNT HUGENHOLTZ/JOÃO PEDRO QUINTAIS, *cit.*, pp. 1203-1205.

⁵⁸ PEDRO PERDIGÃO DA LANA, *A autoria das obras autonomamente geradas por Inteligência Artificial e o domínio do público*, Almedina, Coimbra, 2023.

*prompts*⁵⁹. Será que devemos olhar para essas instruções dirigidas a um computador como uma outra forma de código informático?⁶⁰

6. Prompts e AI Slop

A emergência dos sistemas de inteligência artificial e dos modelos de linguagem de grande escala (*Large Language Models*) introduz uma nova complexidade. Os *prompts* – instruções em linguagem natural dirigidas a sistemas de IA – representam uma forma híbrida de programação que desafia as categorias tradicionais. Não se trata de código no sentido clássico, mas de instruções estruturadas que dirigem o comportamento de sistemas computacionais complexos, muitas vezes expressas em português ou inglês. Em contrapartida, continuam a ser comandos em texto para que um sistema informático realize as suas funções.

Este novo paradigma gera interrogações sobre a natureza da “programação”. Se esta tradicionalmente requeria conhecimento de linguagens formais específicas (C++, Python, Java), os sistemas de IA contemporâneos podem ser “programados” ou “controlados” através de linguagem natural⁶¹. Esta abordagem representa aquilo que alguns denominam de “programação por *vibes*” (*vibe coding*), em que a precisão técnica tradicional é substituída por instruções mais vagas e intuitivas. A mudança levanta questões sobre a titularidade de código automaticamente gerado e, simultaneamente, facilita a criação de aplicações fraudulentas que imitam interfaces conhecidas (nomeadamente para efeitos de *phishing*).

Mas os *prompts* não são todos iguais do ponto de vista da complexidade e organização. Um *prompt* simples como “Bulldogue alemão” para gerar uma imagem difere substancialmente de instruções elaboradas que incluem especificações detalhadas sobre estilo, técnica, iluminação, composição e outros parâmetros técnicos⁶². Estes *prompts* complexos podem envolver centenas de palavras cuidado-

⁵⁹ MARK LEMLEY, “How generative AI turns copyright upside down”, *Science & Technology Law Review*, (2024), pp. 21-44.

⁶⁰ NUNO SOUSA E SILVA, “Prompts as code?”, in *Kluwer Copyright Blog*, 5 de Novembro de 2024.

⁶¹ Curiosamente a definição da lei brasileira admite explicitamente que o programa de computador seja expresso em linguagem natural. Admite-se que o legislador se referia ao código fonte, mas nada impede uma interpretação atualista.

⁶² Um exemplo: “A hyper-detailed cyberpunk geisha assassin with traditional white kabuki makeup mixed with neon circuit tattoos, standing on a rain-soaked Tokyo rooftop at midnight, cherry blossom petals swirling in the wind illuminated by holographic advertisements, wearing a flowing silk kimono that seamlessly blends into digital pixel fragments, holding a katana with a blade made of pure light, background featuring towering skyscrapers with flying cars and massive LED billboards reflecting in puddles, atmospheric fog with volumetric lighting, shot with a 85mm lens with shallow depth of field, in

samente escolhidas e estruturadas, refletindo conhecimento especializado sobre arte, fotografia, *design* ou outros domínios criativos, para além de constituírem uma forma de controlar o sistema de IA. Fala-se mesmo em engenharia de *prompts* (*prompt engineering*) como uma disciplina emergente de conhecimento humano.

Do ponto de vista jusautorais, a questão central é determinar quando um *prompt* transcende a mera funcionalidade e atinge o patamar de expressão criativa protegível. Seguindo a jurisprudência estabelecida, a proteção por direito de autor exige que a criação constitua uma “criação intelectual do seu autor”. *Prompts* elaborados que demonstrem escolhas criativas específicas, estruturação original e conhecimento especializado podem, em princípio, cumprir este requisito.

Admitindo que alguns *prompts* são passíveis de proteção jusautorais, levanta-se desde logo uma questão de qualificação: devem ser enquadrados como programas de computador, constituindo uma nova linguagem de programação ou constituem obras literárias “clássicas”? Na mesma linha, também não é evidente qual a consequência da atribuição desta tutela e o âmbito de proteção de um *prompt*. Impede-se a sua reutilização (reprodução e comunicação ao público) como *input* de um model e o *output* deve ser encarado como obra derivada e ilícita? Devemos, em alternativa tratar a utilização de um *prompt* para gerar um resultado como uma mera mudança de formato (da mesma forma que olhamos para a relação entre um edifício e um projeto de arquitetura ou uma pauta e uma canção gravada)? A questão é particularmente difícil porque o mesmo *prompt* utilizado em sistemas diferentes ou até no mesmo sistema, não gera garantidamente o mesmo resultado.

Mas as interrogações não se colocam apenas quanto ao *input* do sistema. Há, justificadamente, um receio com o impacto que as criações geradas por IA têm na cultura. Fala-se em *AI Slop*, um termo recente que se refere a conteúdo de baixa qualidade, genérico ou indesejado gerado por sistemas de IA. Este conceito está para o *browser* como o *spam* está para o email⁶³. O custo (esforço necessário) de criar textos, imagens e vídeos tornou-se extremamente reduzido. Agora é possível escrever dezenas de livros e produzir milhares de imagens em poucas horas. Isto tem gerado uma preocupação com a visível degradação de conteúdos na Internet com a proliferação destas criações.

the style of Blade Runner 2049 meets traditional Japanese woodblock prints by Hokusai, mixed with the digital art style of Simon Stålenhag and the cinematography of Roger Deakins, color palette of electric blues, hot magentas, and warm amber highlights, dramatic chiaroscuro lighting with rim lighting on the character, 8K resolution, photorealistic rendering with subtle film grain, --ar 16:9 --v 6 --style raw --stylize 750 --chaos 25 --weird 50 --quality 2 --seed 12345”.

⁶³ ALEX HERN/DAN MILMO, “Spam, junk ... slop? The latest wave of AI behind the ‘zombie internet’”, *The Guardian*, 19 de Maio de 2024.

Há que ponderar se não faz sentido adoptar uma exigência reforçada de originalidade/contribuição para estas criações, indo para lá do “contributo mínimo” típico da Propriedade Intelectual⁶⁴.

Poder-se-á dizer que este não é sequer um tema de Direito Intelectual. Creio que a existência e proliferação deste tipo de conteúdos não será afetado pela existência ou não de tutela jusautoral e, paradoxalmente, arredá-los de proteção jusautoral pode mesmo contribuir para a sua disseminação. Em todo o caso, estou convencido que se tornará rapidamente um problema merecedor da atenção dos juristas.

7. O direito *sui generis* do fabricante de bases de dados

Nos anos 90 a União Europeia entendeu ser útil proteger o investimento realizado na criação de bases de dados, esperando com isso incentivar e proteger as empresas europeias que investiam na criação de produtos de dados. Assim, a Diretiva 96/9/CE do Parlamento Europeu e do Conselho, de 11 de Março de 1996, relativa à proteção jurídica das bases de dados, estabeleceu um regime de proteção duplo (direitos de autor e um direito *sui generis* do fabricante de bases de dados) para as bases de dados, tendo a mesma sido transposta, em Portugal, pelo Decreto-Lei nº 122/2000, de 4 de Julho.

Este direito *sui generis* não conhece paralelo no Brasil ou nos EUA e está sujeito a uma reserva de reciprocidade (art. 11º Diretiva 96/9 do e art. 2º/2 do DL 122/2000)⁶⁵. Ou seja, só pode ser titular deste direito o nacional/empresa de um país que tenha um regime equivalente a este, o que significa que empresas norte-americanas, chinesas ou brasileiras não gozarão de proteção.

Base de dados é definida como “a coletânea de obras, dados ou outros elementos independentes, dispostos de modo sistemático ou metódico e susceptíveis de acesso individual por meios eletrónicos ou outros” (art. 1º/2 do DL 122/2000). Trata-se de um conceito com “...um alcance amplo, livre de considerações de ordem formal, técnica ou material”⁶⁶. Este conceito pode incluir,

⁶⁴ DENIS BORGES BARBOSA/RODRIGO SOUTO MAIOR/CAROLINA TINOCO RAMOS, *O contributo mínimo na propriedade intelectual: atividade inventiva, originalidade, distinguibilidade e margem mínima*, Editora Lumen Juris, Rio de Janeiro, 2010.

⁶⁵ Chamou-se-lhe direito *sui generis* porque não se pretendeu que ele fosse abrangido pelas convenções internacionais em matéria de direito de autor ou direitos conexos, evitando a regra de tratamento nacional/não discriminação em função da nacionalidade.

⁶⁶ C-444/02, *OPAP* (EU:C:2004:697), §20. O TJ confirmou esta interpretação ampla ao decidir que também “os dados geográficos que são extraídos de uma carta topográfica por um terceiro para produzir e

nomeadamente, listas telefónicas, antologias poéticas, dicionários, *websites*, anúncios num jornal, coleção de *links*, horários de comboio, resultados desportivos, programação televisiva ou dados meteorológicos.

Em contrapartida, o mesmo é objeto de uma delimitação negativa: por um lado, o *software* que suporta a base de dados (como *MySQL*, *MongoDB* ou *MariaDB*) está expressamente excluído pelo art. 1º/4 do DL 122/2000; por outro lado, o considerando 17 da Diretiva 96/9 explicita que “a fixação de uma obra audiovisual, cinematográfica, literária ou musical, como tal, não é abrangida pelo âmbito de aplicação da presente diretiva” (ou seja, não basta a mera recolha de materiais)⁶⁷.

O direito do fabricante de base dados é atribuído ao produtor de base de dados que tenha feito um investimento substancial em termos qualitativos ou quantitativos na “obtenção, verificação ou apresentação do conteúdo de uma base de dados” (art. 12º/2 do DL 122/2000).

O objetivo do direito do fabricante de bases de dados é proteger o investimento, conferindo ao produtor o direito de proibir a extração e a reutilização da totalidade ou de uma parte substancial (aferida do ponto de vista qualitativo, ou seja, a parte mais importante em termos de investimento, ou quantitativo, isto é, tendo em conta a totalidade dos dados⁶⁸) do seu conteúdo. Por isso mesmo, extrair apenas alguma informação, que constitua a utilização típica de quem consulta uma base de dados, não constitui uma infração do direito *sui generis*. Isso é explicitamente mencionado nos direitos do utilizador legítimo, consagrados no art. 14º do DL 122/2000. O direito *sui generis* é atribuído àquele que realize o investimento, mesmo que não realize diretamente qualquer outra atividade que não seja pagar.

O Tribunal de Justiça produziu já jurisprudência abundante sobre o conceito de investimento substancial. Nos quatro julgamentos referidos coletivamente como os *fixtures cases*, emitidos todos no mesmo dia 9 de Novembro de 2004,⁶⁹ o Tribunal esclareceu que o âmbito do direito do fabricante das bases de dados

comercializar outro mapa mantém, depois da sua extração, um valor informativo suficiente para poderem ser qualificados de «elementos independentes» de uma «base de dados» na aceção da referida disposição» (C-490/14, *Verlag Esterbauer* (EU:C:2015:735), §28-29).

⁶⁷ JOSÉ ALBERTO VIEIRA, *Direito de Autor...*, cit., p. 586.

⁶⁸ C-203/02, *BHB v William Hill* (EU:C:2004:695), §71: “O conceito de parte substancial, avaliada em termos qualitativos, do conteúdo da base de dados refere-se à importância do investimento ligado à obtenção, à verificação ou à apresentação do conteúdo do objeto do ato de extração e/ou de reutilização, independentemente da questão de saber se este objeto representa uma parte quantitativamente substancial do conteúdo geral da base de dados protegida. Uma parte quantitativamente negligenciável do conteúdo de uma base de dados pode, com efeito, representar, em termos de obtenção, de verificação ou de apresentação, um importante investimento humano, técnico ou financeiro”.

⁶⁹ C-203/02, *BHB v William Hill*; C-46/02, *Oy Veikkaus* (EU:C:2004:694); C-338/02, *Svenska Spel* (EU:C:2004:696); C-444/02, *OPAP*.

não é tão extenso como se pensava inicialmente⁷⁰. Segundo esta jurisprudência o único investimento relevante para aceder à proteção é aquele que é dirigido especificamente à produção da base de dados. Não deve ser tido em conta aquele investimento que, apesar de ser relacionado com a atividade em que a base de dados surge (v.g., a organização de uma corrida, de viagens de comboio ou de jogos de futebol), não diz diretamente respeito à obtenção, verificação ou apresentação desses dados⁷¹. Isto significa que não se protegem quaisquer bases de dados, mas apenas os produtos de informação, aquelas coletâneas de informação em relação às quais houve um investimento específico e substancial.

Este direito tem uma duração de 15 anos (art. 16º do DL 122/2000), mas pode ser infinitamente renovado com cada investimento substancial que seja feito na atualização ou verificação dos elementos que compõem a base de dados (art. 17º do DL 122/2000).

Com as evoluções recentes no contexto das ciências da computação, a sua aplicabilidade pode expandir-se, designadamente aos dados sintéticos e aos pesos de redes neuronais artificiais.

8. Dados sintéticos

Os dados sintéticos são dados artificiais, gerados automaticamente para imitar dados reais, reproduzindo as suas características estatísticas⁷². Embora sejam criados automaticamente por *software*, o seu desenvolvimento implica investimento substancial em infraestrutura computacional, conhecimentos técnicos, e processos de validação da qualidade dos dados gerados. A meu ver este investimento na geração dos chamados dados sintéticos, pode sustentar a respetiva proteção pelo direito do fabricante de bases de dados, que não tem requisitos de contribuição humana para lá do investimento⁷³. Em contrapartida, os dados sinté-

⁷⁰ MATTHIAS LEISTNER, “The protection of databases”, in: *Research Handbook on the Future of Copyright* (Ed. ESTELLE DERCLAYE), Edward Elgar, Cheltenham, 2009, pp. 429-449.

⁷¹ C-203/02, *BHB v William Hill*, §30-42. Para uma confirmação desta abordagem cfr. C-604/10, *Football Dataco (Yahoo)* (EU:C:2012:115), §28 (sublinhando que esta conclusão não afecta a possibilidade de a base de dados em causa gozar de direitos de autor).

⁷² PETER LEE, “Synthetic Data and the Future of AI”, *Cornell Law Review*, 110, (2024), pp. 1-74. No processo de geração destes dados coloca-se frequentemente o problema de licitude de utilização não autorizada de *inputs* para treino (cfr. KALPANA TYAGI, “Synthetic Data, Data Protection and Copyright in an era of generative AI”, *JIPITEC*, 16 (2025) pp. 186 e ss.).

⁷³ BERNT HUGENHOLTZ, “Implementing the European Database Directive”, in: *Intellectual Property and Information Law, Essays in Honour of Herman Cohen Jehoram* (Ed. JAN J.C. KABEL/GERARD J.H.M. MOM), Wolters Kluwer, Alphen an den Rijn, 1998, pp. 183-200.

ticos não podem gozar de direito de autor visto não revestirem a originalidade exigida para o efeito e, frequentemente, não terem autor humano⁷⁴.

Importa igualmente distinguir o *dataset* sintético (protegível) do algoritmo gerador (que, como vimos, no máximo pode ser um segredo de negócio, sem prejuízo de se proteger o código que o implemente). A proteção do direito *sui generis* incide sobre a coleção estruturada de dados sintéticos, não sobre o respetivo método de produção. Por outro lado, a proteção incide sobre a base de dados como um todo e não sobre cada um dos dados sintéticos individualmente considerados.

9. Pesos de redes neuronais

Ao contrário dos programas de computador clássicos escritos por programadores, muitos dos modelos de Inteligência Artificial atualmente utilizados baseiam-se em vastas redes neuronais artificiais treinadas com quantidades enormes de dados. Ninguém escreve ou codifica individualmente estes modelos; eles são gerados através de um processo automatizado de treino. Preparados os dados e definida a arquitetura (as características e particularidades de um modelo definidas antes do treino são chamadas hiperparâmetros), os computadores funcionam durante muito tempo e com custos elevados para adquirir conhecimento “por si próprios” (fala-se a este propósito de *software 2.0*.⁷⁵).

O resultado final, isto é, o modelo, consiste em dois ficheiros – um ficheiro simples que estabelece o funcionamento do modelo (a arquitetura) e um ficheiro muito maior de parâmetros ou pesos (expressos como números de vírgula flutuante). Os pesos constituem uma expressão matemática da ligação entre os neurónios artificiais que compõem a rede. Como refere MARTIN ANDERSON, “na aprendizagem automática, os pesos são tudo – o ‘ouro’ que emerge após semanas ou mesmo meses de treino de um sistema”⁷⁶.

Do ponto de vista da proteção por direito de autor, os pesos dos modelos de IA enfrentam obstáculos significativos. Primeiro, são simples valores numéricos, desprovidos de qualquer expressão criativa. Segundo, não possuem um autor humano, uma vez que são determinados automaticamente através de algoritmos

⁷⁴ PETER LEE, *cit.*, pp. 45-46.

⁷⁵ HAO-YUN CHEN, “Copyright protection for software 2.0.: Rethink the justification of software protection under Copyright Law”, in: *Artificial Intelligence and Intellectual Property* (Ed. JYH-AN LEE/RETO HILTY/KUNG-CHUNG LIU), Oxford University Press, Oxford, 2021, pp. 323-340.

⁷⁶ MARTIN ANDERSON, “Weights in machine learning”, *Metaphysic*, 12 de Janeiro de 2023.

de otimização. Terceiro, a sua natureza é exclusivamente funcional. Assim, parece seguro concluir que não são passíveis de proteção jusautorai⁷⁷.

Em contrapartida, nada parece obstar à sua proteção como base de dados⁷⁸. Os pesos de um modelo de IA podem, em princípio, cumprir os requisitos estabelecidos pela Diretiva 96/9/CE para qualificação como base de dados. Constituem uma coleção de elementos independentes (cada peso individual), dispostos de modo sistemático (segundo a arquitetura da rede neuronal) e suscetíveis de acesso individual por meios eletrónicos. O treino de modelos de IA contemporâneos envolve um investimento substancial: recursos computacionais muito exigentes, conhecimentos técnicos especializados, e processos de classificação, curadoria e preparação de dados que podem custar centenas de milhões de euros e ocupar dezenas de pessoas durante meses.

O facto de não existirem requisitos de autoria humana para a concessão do direito *sui generis* torna esta via particularmente atrativa para a proteção de modelos de IA. Ao contrário do direito de autor, que exige originalidade, o direito do fabricante de bases de dados basta-se com a realização de investimento.

10. Conclusão

Como resulta desta exposição alguns dos novos artefactos digitais não têm um enquadramento óbvio. Mesmo assim, há algumas conclusões preliminares que podemos enunciar.

O regime especial de proteção do *software*, estabelecido há mais de 30 anos começa a mostrar sinais de inadequação face às realidades do “Software 2.0” em que o código não é escrito diretamente por humanos, mas emerge de processos automatizados de treino.

A dicotomia ideia/expressão, pedra angular do direito de autor, é posta à prova de forma particular quando é aplicada a algoritmos, modelos de IA e *prompts*. MARK LEMLEY sugere mesmo que a IA generativa inverte a lógica desta ideia: o autor fornece a ideia e o sistema de IA é que realiza as tarefas “clássicamente autorais”⁷⁹.

⁷⁷ HAO-YUN CHEN, *cit.*, pp. 323-340; PETER SLOWINSKI, *cit.*, p. 355; BEGOÑA GONZALEZ OTERO, “Machine learning models under the Copyright microscope: Is EU copyright fit for purpose?”, *GRUR-Int*, (2021), pp. 1043-1055.

⁷⁸ NUNO SOUSA E SILVA, “Are AI models’ weights protected databases?”, in *Kluwer Copyright Blog*, 18 de Janeiro de 2024.

⁷⁹ MARK LEMLEY, “How generative AI turns Copyright upside down”, *Science & Technology Law Review*, (2024), p. 28: “The result is that, increasingly, the things humans contribute in a collaboration with

Apesar de não ter sido pensado especificamente para este efeito, o direito *sui generis* do fabricante de bases de dados tem potencial para proteger alguns destes novos artefactos digitais, nomeadamente dados sintéticos e pesos de redes neurais. A ausência de requisitos de autoria humana e o objetivo de proteção do investimento tornam este regime mais adaptável. Em contrapartida, este regime não conhece paralelo fora da União Europeia e está sujeito a uma regra de reciprocidade, pelo que a sua utilização pode causar uma divergência.

A emergência de novos paradigmas como *vibe coding* e *prompt engineering* sugere que poderemos estar perante novas formas de expressão criativa no contexto do *software* que podem obrigar a repensar o “contributo humano” indispensável à autoria. A questão coloca-se também quanto a saber como enquadrar os *prompts*, quando estes revistam originalidade: estão como a pauta para a obra musical (atraindo proteção jusautoral), como o código fonte para o código objeto (instruções que são convertidas para o funcionamento de uma máquina), ou como instruções genéricas dadas por um orientador a um orientado (que não geram coautoria)?

Estas questões ilustram a tensão fundamental, mas típica, entre proteger o investimento e a inovação, por um lado, e garantir os princípios estabelecidos do Direito de Autor, designadamente a adequada remuneração da atividade criativa e a liberdade de ideias. Não parece ser possível proteger a generalidade destes artefactos digitais sem prescindir de princípios fundamentais do direito de autor ou criar um regime novo⁸⁰. Não vejo razão para tal.

generative AI will be ideas and high-level concepts. AI will contribute the expression. That turns copyright law on its head”.

⁸⁰ Em diálogo com o Claude Sonnet 4 (LLM da Anthropic) sobre este tema (carreguei este ficheiro e pedi “faz uma análise crítica seguida de propostas inovadoras que possam resolver as questões enunciadas neste artigo. Sê criativo e preciso.”), sugeriu-me a criação de um artigo novo com este teor: “Artigo 1º-A (novo) “Expressões Algorítmicas Complexas” 1. São protegidas como programas de computador as expressões algorítmicas que, não obstante serem funcionais, demonstrem: a) Escolhas técnicas não impostas pela função pretendida; b) Estruturação original não predeterminada por standards técnicos; c) Capacidade generativa que transcenda mera execução determinística. 2. A proteção abrange exclusivamente a expressão específica, não a função ou método subjacente”. A proposta é criativa, mas apresenta vários problemas. Quanto ao sentido crítico, por enquanto, ainda tenho mais fé nos seres humanos do que nestes sistemas.