



UNIVERSIDADE CATÓLICA PORTUGUESA

Big Data e Propriedade Intelectual: *a*
(des)proteção das grandes bases de dados

Diniz Caseiro Patrão

Mestrado em Direito

Faculdade de Direito | Escola do Porto

2022

UNIVERSIDADE CATÓLICA PORTUGUESA

Big Data e Propriedade Intelectual: *a*
(des)proteção das grandes bases de dados

Diniz Caseiro Patrão

Orientador: Professora Doutora Maria Victória Rocha

Mestrado em Direito

Faculdade de Direito | Escola do Porto

2022

*“Para ser grande, sê inteiro: nada
Teu exagera ou exclui.
Sê todo em cada coisa. Põe quanto és
No mínimo que fazes.
Assim em cada lago a lua toda
Brilha, porque alta vive.”*

(Ricardo Reis)

Agradecimentos

Em primeiro lugar, aos meus pais e irmã, por todo o amor, carinho, paciência e apoio incondicional,

Em especial, à minha Margarida, que me inspira a fazer mais e melhor,

À Avó Lurdes, pela ternura,

Aos meus amigos, pela alegria e amizade,

À Professora Doutora Maria Victória Rocha, pela disponibilidade, partilha de bibliografia, correções e sugestões, que foram imprescindíveis para a concretização deste estudo,

A todos aqueles, que de uma maneira ou de outra, me ajudaram a completar este ciclo.

RESUMO

O crescente desenvolvimento do setor tecnológico tem provocado um importante crescimento nos mercados de Big Data. Estes, por sua vez, enquanto tecnologia disruptiva, originam sérios desafios jurídicos, em especial relativos ao Direito de Propriedade Intelectual. Com o presente estudo, procuramos apurar até que ponto é que as bases de dados decorrentes de Big Data se encontram (des)protegidas pela legislação vigente. Em particular iremos avaliar a aplicabilidade da Diretiva relativa à proteção das bases de dados, em contexto de Big Data

PALAVRAS-CHAVE: Big Data; Direito de Autor; Direito Sui Generis; Base de Dados

ABSTRACT

The growing development of the technology sector has caused a significant growth in the Big Data markets. As, a disruptive technology, Big Data poses serious legal challenges, especially regarding Intellectual Property Law. With this study, we seek to understand to what extent the databases arising from Big Data are (un)protected by current legislation. In particular, we will assess the applicability of the Directive on the protection of databases in the context of Big Data.

KEY WORDS: Big Data; Copyright; Sui Generis Law; Database

ÍNDICE

Introdução	9
1. A Era dos Big Data	11
2. O conceito de Big Data	13
2.1 <i>Modus operandi</i> da tecnologia de Big Data	17
3. Big Data e Propriedade Intelectual	18
4. Transformações na concepção tradicional das Bases de dados	19
5. A Tutela Jurídica das Bases de Dados	21
5.1. A definição de base de dados à luz da Diretiva	22
6. A dupla proteção Europeia	25
6.1. A proteção de bases de dados numa perspetiva de direitos de autor	26
6.2 O Direito Sui Generis	29
6.2.1 Titularidade	33
6.2.2 Proteção	34
6.2.3 Exceções	36
6.2.4 Duração	36
7. Poderão as bases de dados em contexto de Big Data ser protegidos ao abrigo do Direito Sui Generis?	37
8. A Diretiva 96/9/CE na era dos Big Data	41
9. O Futuro	43
Conclusão:	45
Bibliografia:	47

LISTA DE SIGLAS E ABREVIATURAS

al. -Alínea

art .- Artigo

arts. - Artigos

CDAC- Código dos Direitos de Autor e Direitos Conexos

CE - Comissão Europeia

Cfr.- Conferira

Dir.- Diretiva

DL- Decreto-Lei

Ed. -Edição

Et al. -E outros

UE - União Europeia

Ex. -Exemplo

IA - Inteligência Artificial

IoT- Internet of things

NoSQL – Not Only Structured Query Language

OCDE- Organização para a Cooperação e Desenvolvimento Económico

p. - Página

p. ex.- Por exemplo

Par. - Parágrafo

pp. – Páginas

Proc.-Processo

RDBMS – Relational Database Management System

SQL -Structured Query Language

TIC - Tecnologias da Informação e Comunicação

TJUE - Tribunal de Justiça da União Europeia

V. -Ver

Vol. -Volume

Introdução

Nos últimos anos, temos assistido a uma enorme (r)evolução tecnológica. Em consequência, a eclosão das Tecnologias da Informação e Comunicação (TIC) e a Internet - que, por sua vez, permitiram que ocorresse um aumento da digitalização¹- conduziram a uma revolução económica e social em todas as áreas². De facto, vivemos em plena quarta Revolução Industrial (ou indústria 4.0)³ – e já de olhos postos numa quinta – tendo-se alterado a economia mundial e a forma como vivemos e nos relacionamos, através de inovações tecnológicas.

Na base – e, simultaneamente, em consequência – de toda esta inovação encontram-se os dados. Estes têm sido objeto de um desenvolvimento sem precedentes, tendo aumentado, tanto em volume, como em valor. E assim surgiu o fenómeno de Big Data.

De facto, a sociedade viu-se submersa em dados, e rapidamente se concluiu que a sua análise e coleta, os transformava num valioso ativo. Desta forma, dada a sua importância e valor, os dados são considerados como o novo petróleo da era digital⁴, tendo, definitivamente, impulsionado a inovação. Contudo, apesar do seu uso se poder considerar muito promissor, não é isento de desafios.

Assim, não sendo novidade que o direito é constantemente desafiado pelas tecnologias emergentes, a nossa atenção deverá, agora, recair sobre a questão particular de saber se o regime vigente relativo à proteção de bases de dados, regulado pela Diretiva 96/9/CE de 11 de março, será adequado para garantir proteção às bases de dados no âmbito de Big Data.

Impõe-se a questão: poderá um regime concebido numa fase onde prevalecia o método analógico ser aplicável agora que avanços tecnológicos e a revolução digital atingiram patamares nunca antes vistos? Conseguirá o Direito acompanhar a mesma tendência?

¹ A Digitalização poderá ser entendida como o processo pelo qual uma imagem ou sinal analógico é transformado em código digital.

² CANNATACI *et al*, 2020, p.1.

³ Christoph J. Bartodzie, afirma que a Indústria 4.0 é marcada pela integração de tecnologias inovadoras, tais como a Internet das Coisas, Big Data e a Computação em Nuvem, nos processos de fabricação e logística, ou seja, a “integração técnica de sistemas Cyber-Físicos nos processos de manufaturação e logística. BARTODZIEJ, 2017, p. 2.

⁴ The economist, 2016.

Para tanto, partiremos de uma análise e exposição do conceito de Big Data – procurando desmistificar este conceito – quanto à sua definição, funcionamento e tecnologias que abrange. O que, porém, se mostrará ímprobo, dada a ambiguidade deste conceito.

Em seguida, o foco incidirá na eventual proteção conferida às bases de dados, em especial, na supramencionada Diretiva. Como tal, será analisada a dupla proteção conferida pela mesma, os Direitos de Autor e o Direito Especial do Fabricante de Bases de Dados, que se apresenta como um direito “Sui Generis”.

Por último, o foco estará, ainda, em analisar se a Diretiva relativa à proteção das bases de dados se encontra apta a responder aos problemas levantados pelos Big Data.

1. A Era dos Big Data

É indubitável que vivemos, hoje, numa sociedade profundamente marcada por um alvoroço tecnológico. Os avanços tecnológicos que, num passado recente, se cingiam à pura ficção científica, encontram-se hoje, e mais do que nunca, na ordem do dia⁵.

Ao longo dos anos, assistimos a uma constante disrupção nos padrões normais da sociedade. E essas mudanças ocorrem a par e passo com as transformações tecnológicas que conduzem a novas maneiras de pensar e ver o mundo.

Com a implementação de tecnologias tão avançadas em vastas áreas – tais como a Internet das Coisas (IoT)⁶, Big Data, Robótica e Inteligência Artificial⁷ –, não só os indivíduos se encontram conectados entre si, como também com objetos utilizados no dia-a-dia. E, assim, tem diminuído ao longo do tempo o número de indivíduos que não se encontram conectados a qualquer aparelho eletrônico, seja por smartphones, smartwatches, smartcars, tablets ou computadores portáteis⁸.

Em suma, a internet faz parte do dia-a-dia da maioria das pessoas, fazendo com que haja uma interação global⁹.

Nos últimos anos, a realidade digital desenvolveu-se exponencialmente. Estudos recentes revelam que a recente pandemia de COVID-19 foi um “grande acelerador”¹⁰ no que respeita à tendência global para adotar tecnologias emergentes. O que, efetivamente, levou a transformações no estilo de vida, nos padrões de trabalho e nas estratégias de

⁵ CÔNSUL, 2019, p. 183.

⁶ Não existe ainda nenhum consenso quanto à definição de IoT, contudo, a definição que nos parece mais adequada é a que caracteriza a IoT como “uma rede aberta e abrangente de objetos inteligentes que têm a capacidade de se auto-organizar, compartilhar informações, dados e recursos, reagindo e agindo diante de situações e mudanças no ambiente”. MADAKAM, *et al*, 2015; SHARMA & SHAMKUWAR, 2019

⁷ Sobre Inteligência Artificial, veja p. ex. BODEN, 2018, pp. 1-7.

⁸ A este respeito o Bundeskartellamt, juntamente com a Autorité de la Concurrence, salientam, que os consumidores, utilizam cada vez mais a Internet para atividades que até então não justificavam a sua utilização. BIDARRA, 2018, p. 19; Bundeskartellamt e L’Autorité de la Concurrence, “Competition Law and Data”, 2016, p. 11. disponível em: https://www.bundeskartellamt.de/SharedDocs/Publikation/DE/Berichte/Big%20Data%20Papier.pdf?__blob=publicationFile&v=2.

⁹ Estudos realizados apontam que no ano de 2021, com uma população mundial, estimada de 7,8 biliões, aproximadamente 4,93 biliões da população mundial têm acesso e usam a internet com frequência. O que significa que 63.2% da população mundial utiliza a internet. disponível em - <https://www.broadbandsearch.net/blog/internet-statistics>.

¹⁰ Uma pesquisa realizada pela empresa de consultoria empresarial, McKinsey & Company veio concluir que durante a pandemia COVID-19, os consumidores mudaram drasticamente para os canais online, sendo que as empresas e indústrias viram-se obrigadas a responder a estas mudanças, adotando medidas as que provocaram a uma aceleração na adoção de tecnologias digitais em vários anos.

Disponível em: <https://www.mckinsey.com/business-functions/strategy-and-corporate-finance/our-insights/how-covid-19-has-pushed-companies-over-the-technology-tipping-point-and-transformed-business-forever>.

negócios dos indivíduos e das empresas, revelando-se um catalisador no que respeita à crescente digitalização¹¹.

A verdade é que, como consequência da crescente aplicação de tecnologias disruptivas como a Internet das Coisas, Inteligência Artificial e Computação em Nuvem¹², temos assistido a um grande aumento de dados gerados e armazenados. Isto poderá dever-se ao facto de qualquer acontecimento poder ser convertido em dados.¹³ Estamos assim perante o fenómeno da datificação¹⁴. Portanto, nos dias que correm, é praticamente inconcebível que não estejam a ser gerados, a todo o tempo, vastos volumes de dados –das mais variadas fontes –, levando certos autores a declarar que os dados são o novo petróleo da economia digital¹⁵.

A tendência de aumento do tamanho, número e variedade de dados recolhidos é uma certeza, principalmente numa altura em que surgem tecnologias como 5G¹⁶ e “Metaverso”¹⁷ (sendo de prever que num futuro próximo as pessoas “vivam” online)¹⁸.

Recuando a uma fase – não muito distante – em que prevalecia o método analógico, a criação, coleta, armazenamento e, sobretudo, a análise de dados constituía um procedimento dispendioso

Por outro lado, atualmente, a coleta de dados é considerada ubíqua, sendo também viabilizada pelo constante uso de aparelhos conectados à Internet (no âmbito da IoT), o que possibilitou a integração de componentes digitais nas ferramentas analógicas do dia-a-dia. Esta circunstância de haver uma maior, mais abrangente e mais democrática disponibilidade de tecnologias e meios, tem como consequência menores custos de armazenamento e a evolução dos próprios instrumentos – nomeadamente no que respeita à, cada vez maior, capacidade de análise, armazenamento e processamento de dados.

Com efeito, essa explosão reúne o potencial de criar oportunidades nas mais variadas áreas, designadamente no comércio, uma vez que a possibilidade de cruzar dados

¹¹ Cfr. AMANKWAH *et al*, 2021, p. 608.

¹² MASSENO, 2019, p.3.

¹³ CÔNSUL, 2019.

¹⁴ SOLDATOV, 2020, p.234.

¹⁵ Aponta neste sentido Mateja Durovic e Franciszek Lech, e em sentido inverso, Joe Canatacci, Valeria Falce e Oreste Polliciono afirmam que contrariamente ao petróleo, os dados não perdem o seu valor depois de utilizados. O que significa que a utilização destes por parte de uma empresa não impede outra de recorrer exatamente aos mesmos, nem diminui o seu valor. DOROVIC & LECH, 2020; CANNATACCI *et al*, 2020.

¹⁶ Cfr. RASHID & RAZAK, 2019; RUBINFELD & GAL, 2017.

¹⁷ Cfr. LEE *et al.*, 2021.

¹⁸ Estima-se que a quantidade total de dados criados, capturados, copiados e consumidos globalmente atinja, até 2025, 175 zettabytes. - https://seedscientific.com/how-much-data-is-created-every-day/?fbclid=IwAR1BXNr4CLdyzTrgoh1YjnVGzVI7OrWnSKiT7L3e4TA5-2scX3BNq_EtdxY

através de algoritmos de análises – que permitem a obtenção de informações de mercado provenientes, diretamente, dos consumidores, – está a revolucionar o *modus operandi* das empresas. Também outros setores, tal como a agricultura¹⁹, o setor bancário²⁰, o ensino²¹, a saúde²², poderão beneficiar desta circunstância.

Os dados são, hoje, os responsáveis por sustentar o ecossistema digital²³, havendo até quem afirme que são a sua unidade básica e fundamental, assim como os átomos são a unidade básica da matéria²⁴.

É, por isso, seguro afirmar que nos encontramos em plena transição da era da informação para a era dos Big Data²⁵.

2. O conceito de Big Data

Nos dias que correm, torna-se cada vez mais comum depararmo-nos com o conceito de Big Data. Imediatamente damos conta de que a sua importância é já amplamente reconhecida e incontestável. Contudo, o mesmo não sucede quando o assunto é a sua definição²⁶. Big Data é um conceito que se reveste, ainda, de alguma ambiguidade e é motivo de divergências doutrinárias, levantando ainda mais perguntas do que respostas acerca dos desafios que poderá gerar no futuro.

Face a esta falta de clareza em torno do conceito, a Comissão Europeia (CE) procurou oferecer uma definição de Big Data²⁷, falhando no seu propósito, diga-se. De facto, tal definição não se mostra adequada, uma vez que limita o âmbito de Big Data, apenas, ao seu volume de dados:

¹⁹ SYKUTA, 2016.

²⁰ LEAL, 2020.

²¹ BAIG, et al, 2020. Os autores concluíram que o fenómeno Big Data está a tornar-se uma parte essencial no campo educacional.

²² HONG et al., 2018.

²³ Cfr. COMISSÃO EUROPEIA, Comunicação da Comissão ao Parlamento Europeu, ao Conselho, ao Comité Económico e Social Europeu e ao Comité das Regiões – Uma estratégia europeia para os dados COM(2020) 66 final Bruxelas, 2020, p. 3, disponível em <https://eur-lex.europa.eu/legal-content/PT/TXT/PDF/?uri=CELEX:52020DC0066&from=EN>

²⁴ PADINHA, 2020.

²⁵ PALKA & LIPPI, 2021, p. 115.

²⁶ Conforme salienta Ana Alves Leal, no ordenamento jurídico português não existe uma definição legal de Big Data. Encontra-se apenas no Regulamento Geral da Proteção de Dados, a menção – sem definição a «operações de tratamento de grande escala» ou em larga escala. LEAL, 2020 p.201.

²⁷ Foram avançadas outras definições nomeadamente pelo Grupo de Trabalho do Artigo 29º; Grupo de Trabalho Internacional sobre Proteção de Dados nas Telecomunicações; Autoridade Europeia para a Proteção de Dados; e Relatório Gartner. Cfr.SLOOT & SCHENDEL, 2016.

“O termo «megadados» designa grandes quantidades de dados de diferentes tipos produzidos a partir de vários tipos de fontes, nomeadamente pessoas, máquinas e sensores”²⁸.

Posto isto, é importante referir que o conceito de Big Data não deve ser visto somente como um vasto conjunto de dados, na medida em que este deve ser também caracterizado pela abrangência de diferentes componentes²⁹. Partindo deste pressuposto, será de concluir que este conceito, se refere, por um lado, a conjuntos de dados numa escala massiva, - estruturados e não estruturados³⁰ – (caso em que diremos “os” Big Data). Por outro lado, compreende um conjunto de tecnologias³¹ e processos que permitem a conversão em dados de muitos conceitos nunca antes quantificados, convertendo esses fluxos de dados num conhecimento altamente específico,³² nomeadamente, a recolha, o armazenamento e a análise dos dados (caso em que diremos “a” Big Data). Assim, conforme apontam Mateja Durovic e Franciszek Lech, é importante notar que Big Data se refere “a todo o círculo de vida dos dados”³³.

Como já se deixou dito, a definição de Big Data é vista como um ponto de discórdia, havendo inúmeras e distintas definições. Nesse sentido, pode ler-se o pensamento de Romeu Cónsul, que faz uma análise comparativa das definições existentes, criadas por diversos autores. Daí se conclui que, pelo menos, a grande maioria das definições parece concordar que Big Data se trata de um sistema ou conjunto de recursos tecnológicos cujo objetivo é a coleta, armazenamento e processamento de dados em larga escala, por meio do uso de diferentes técnicas computacionais, como o uso de Prospecção de dados³⁴,

²⁸ COMISSÃO EUROPEIA – “The EU data protection reform and big data.” Disponível em: <https://publications.europa.eu/en/publication-detail/-/publication/51fc3ba6-e601-11e7-9749-01aa75ed71a1/language-en/format-PDF>

²⁹ CHEN, *et al* 2014.

³⁰ Cfr. PRAVEEN & CHANDRA 2020.

³¹ “A tecnologia é o hardware utilizado para peneirar, classificar e analisar grandes quantidades de dados em milésimos de segundos. O processo consiste em analisar os dados em padrões, e aplicar os padrões à análise preditiva para depois serem utilizados em novos dados.” – DOROVIC & LECH, 2020 p.156.

³² COSTA, 2021, p. 38.

³³ DOROVIC & LECH, 2020 p.156.

³⁴ A prospecção de dados consiste no processo de obtenção de informações de alta qualidade, de materiais de texto e de bases de dados recorrendo a software, com o fim de encontrar padrões consistentes, que permitem detetar relacionamentos sistemáticos, e por sua vez novos subconjuntos de dados. Cfr. “Big Data and Data Mining: The Role Data Mining Plays in Big Data”. Disponível em: https://online.maryville.edu/blog/big-data-and-data-mining-the-role-data-mining-plays-in-big-data/?fbclid=IwAR34VKM9aJyMMhHkCTOcPfdjMAn7L2vulmb-d_5yCB7OqBT1U86z99u6Ow0

Algoritmos e Inteligência Artificial³⁵. Não existindo, porém, uma definição consensual do conceito de Big Data.

Não obstante, existem três propriedades definidoras de Big Data que podem facilitar a decomposição deste termo, **V**olume, **V**elocidade e **V**ariiedade. Além disso, uma parte da Doutrina defende a existência de uma quarta que, embora não essencial, é cada vez mais típica a **V**eracidade, e uma quinta, derivada de todas as outras, o **V**alor³⁶.

a) Volume:

Em primeiro lugar, o **volume** de dados³⁷ refere-se à dimensão dos conjuntos de dados a analisar e processar. Nos dias de hoje, esses conjuntos são frequentemente maiores que terabytes ou petabytes, chegando mesmo aos zettabytes³⁸.

A possível explicação para o crescimento do volume de dados é, certamente eclosão da computação em nuvem, o advento da banda larga, a disseminação de dispositivos eletrônicos inteligentes e a adesão em massa a redes sociais³⁹, que contribuem largamente para esse aumento, uma vez que os consumidores fornecem ativa e passivamente uma abundante quantidade de dados sobre si mesmos, tornando-se assim verdadeiras máquinas de gerar dados⁴⁰.

Simultaneamente, verificamos uma diminuição dos custos associados à recolha e armazenamento de grandes volumes de dados.

Com efeito, este grande volume de dados requer tecnologias de processamento e armazenamento avançadas, já que, pelas suas proporções colossais, se torna impossível recorrer unicamente aos instrumentos tradicionais.

³⁵COMISSÃO EUROPEIA “The EU Data Protection Reform and Big Data”, 2016, p. 2, disponível em: <https://publications.europa.eu/en/publication-detail/-/publication/51fc3ba6-e601-11e7-9749-01aa75ed71a1/language>

³⁶ Tal como acontece com a definição de big data, também a doutrina diverge relativamente ao número dos V's. Ao modelo de 3 V's alguns autores acrescentam mais dois, havendo até autores que defendem a existência de nove V's, Volume, Velocidade, Veracidade, Valor, Variiedade, Validade, Volatilidade e Visualização Owais & Hussein, 2016.

³⁷ Cumpre, contudo, salientar que as definições de volume são relativas e variam de acordo com o tempo e o tipo de dados. É impossível definir um limite específico, a partir do qual se passa a falar de volumes de Big Data, visto que, no futuro a capacidade de armazenamento irá aumentar, permitindo capturar ainda maiores volumes. GANDONI & HAIDER, 2015 p.138.

³⁸ Um zettabyte é uma unidade de medida de informação que equivale a 1.000.000.000.000 de gigabytes.

³⁹ MASSENO, 2019.

⁴⁰ MCAFEE & BRYNJOLFSSON, 2012.

b) Velocidade:

Paralelamente ao crescimento exponencial do volume de dados, aumentou, também, a velocidade de geração e disponibilização de dados⁴¹, impondo, assim, o desafio de processar os dados rapidamente e, mais importante, utilizá-los a um ritmo mais rápido do que nunca⁴².

Em termos mais simples, a velocidade em Big Data consiste na rapidez de coleta, análise e utilização de um grande volume de dados. A celeridade das diversas operações de tratamento tem aumentado, chegando a ocorrer, praticamente em tempo real, o que conseqüentemente, permite a tomada de decisões de maneira instantânea⁴³.

c) Variedade:

Com a coleta de dados fornecidos por sensores, máquinas e indivíduos⁴⁴, têm-se gerado vastos volumes de dados, sem que seja possível existir um controle sobre o formato dos mesmos. A **variedade** dos dados diz respeito à diversidade de fontes e aos diferentes tipos de dados: estruturados, não estruturados e semiestruturados⁴⁵. Tradicionalmente, os dados tinham uma natureza estruturada. Nos dias de hoje, porém, os dados assumem as mais diversas formas. E essa diversidade constitui uma das razões pelas quais os dados são criados a uma velocidade surpreendente, revolucionando, desse modo, os meios tradicionais de conservação, extração e análise dos mesmos.

⁴¹ Segundo estudos recentes, estima-se que, no ano de 2021, foram realizadas 5,700,000 pesquisas por minuto na Google, foram compartilhadas 65,000 fotografias por minuto no Instagram, o que vem demonstrar que grandes volumes de dados são gerados e disponibilizados a uma enorme velocidade. –

<https://www.statista.com/statistics/195140/new-user-generated-content-uploaded-by-users-per-minute/>

⁴² BOUTIN & CLEMENS, 2017.

⁴³ Cfr. <https://www.oracle.com/big-data/what-is-big-data/>

⁴⁴ Owais & Hussein, 2016 pp.254-255.

⁴⁵ Os dados podem ser classificados como, dados estruturados, não estruturados e semiestruturados. Os dados estruturados são aqueles que tem um padrão e um formato pré-definido antes de serem colocados na base de dados, são exemplos comuns os números, frases fáceis de registrar e codificar. Os dados não estruturados, são dados armazenados no seu formato bruto, sem serem processados, são exemplos comuns o histórico da web, ou os dados de redes sociais. Por sua vez os dados semiestruturados constituem uma mistura dos dois tipos de dados já explicados, são exemplos os arquivos de vídeo e os arquivos de áudio. ZHANG & CHEN, 2015 p.4.

d) Veracidade:

No contexto de Big Data, a **veracidade** está relacionada com a precisão e qualidade dos dados coletados, sendo uma característica essencial para a sua operacionalização. A veracidade encontra suporte, não apenas na precisão dos dados, mas também na confiabilidade da sua fonte, na forma como foram coletados e no contexto em que são analisados. Além disso, inclui a habilidade de assegurar que a multiplicidade de fontes e dados, quando agrupados, farão sentido e proporcionarão informações de alta qualidade.

e) Valor:

O Big Data é frequentemente caracterizado como tendo um baixo valor relativamente ao seu enorme volume de dados. A extração do valor de um grande volume de dados depende do seu tratamento, mediante a coleta e análise dos mesmos em alta velocidade. Será de afirmar que mais valioso do que o próprio volume de dados captados são as próprias informações que podem ser extraídas desse mesmo volume de dados⁴⁶.

O valor dos dados está intrinsecamente ligado ao valor socioeconómico que é alcançado através da utilização de Big Data. E é o potencial valor económico e social que motiva a concentração e o tratamento dos dados nesta escala⁴⁷.

De acordo com a OCDE, a característica “valor” do Big Data é tanto causa como consequência do aumento de volume, variedade e velocidade⁴⁸.

2.1 *Modus operandi* da tecnologia de Big Data:

A Big Data engloba os processos de coleta, processamento, filtragem e análise de extensos conjuntos de dados para que as organizações possam utilizá-los para desenvolver e produzir melhores produtos. Estes processos dependem, em grande medida, de instrumentos de Inteligência Artificial (IA) que sejam capazes de extrair informações de vastos conjuntos de dados. A conjugação destas duas áreas, IA e Big Data,

⁴⁶ ARRUDA, 2018, p 22.

⁴⁷ RUBINFELD & GAL, 2017.

⁴⁸ OCDE. Big data: “Bringing competition policy to the digital era”.

permite a obtenção de resultados incríveis, já que possibilitam a previsão de tendências futuras nos mais variados setores económicos⁴⁹.

Numa primeira fase, os dados são coletados das mais variadas fontes, capturados de forma direta ou indireta, de indivíduos, entidades e sensores⁵⁰. Depois de coletados, os dados são armazenados em servidores internos ou em serviços de computação em nuvem.

Numa segunda fase, os dados coletados e armazenados são processados e analisados em alta velocidade. Essa análise é feita, geralmente, por meio de prospeção de dados, algoritmos inteligentes e softwares, normalmente assentes em técnicas de Machine Learning⁵¹. Dessa operação extrai-se um resultado que servirá o propósito de auxiliar na tomada de decisões⁵².

O tratamento dos Big Data tem como especificidade o facto de lidar com vastos conjuntos de dados, geralmente apresentados em diferentes tipos, formatos e suportes. Perante isto, torna-se evidente a necessidade de filtrar esses dados, sem que se perca o valor que os mesmos apresentam. Este processo é comumente denominado por Big Data Analytics.⁵³ E estes “filtros”, ou modelos de análise, consistem geralmente em: análise descritiva, análise preditiva⁵⁴ e análise prescritiva⁵⁵.

3. Big data e Propriedade Intelectual

Atualmente, enfrentamos os efeitos de um aumento, sem precedentes, na disponibilidade e utilidade dos dados. Os Big Data, têm um papel primordial no que toca

⁴⁹ Cfr. BURDESE, 2020.

⁵⁰ Podem também ser indiretamente adquiridos no mercado por meio de terceiros intermediários, como os chamados data brokers.

⁵¹ Machine Learning é um ramo da Inteligência artificial e da ciência da computação. Os sistemas baseados em machine learning tem a grande de conseguirem extrair padrões a partir de quantidades massivas de dados, o que lhes permite melhorar o seu desempenho numa tarefa específica. Em geral, podemos classificar os usos da aprendizagem de máquinas em dois tipos gerais de tarefas: previsão e automatização. Cfr. SURDEN, 2014.

⁵² Essas informações podem ter diversas aplicações, como a melhoria de processos produtivos e da organização industrial.

⁵³ O Big Data Analytics, em termos simples, consiste no estudo de uma enorme quantidade de dados, com o propósito de retirar padrões de comportamento, descobrir correlações desconhecidas, e avaliaras tendências do mercado de forma a descobrir as preferências dos consumidores. – “Big Data Analites o que é e porque é importante”. – Disponível em: <https://www.jasminsoftware.pt/blog/big-data-analytics-importante/>

⁵⁴ Disponível em: <https://www.aquare.la/analise-descritiva-preditiva-prescritiva-e-cenarizacao/>

⁵⁵ T.CRUIZ, 2017, p. 11.

à inovação, crescimento e competitividade europeus⁵⁶. Contudo, não obstante, a sua crescente importância e relevância, é consensual que estes carecem de uma regulamentação adequada, designadamente a nível europeu.

Toda esta ambiguidade em torno do tema será suscetível de provocar uma desvantagem competitiva. O que já se tem verificado nos dias de hoje, sendo um claro exemplo disso o mercado das bases de dados.

O nosso foco recai, no desafio colocado pelos Big Data à legislação que incide sobre os direitos de Propriedade Intelectual das bases de dados, uma vez que este fenómeno tem sido particularmente problemático na perspetiva dos proprietários das Bases de Dados.

Na União Europeia, a proteção das bases de dados encontra-se plasmada na Diretiva 96/9/CE de 11 de março de 1996 relativa à proteção jurídica das bases de dados⁵⁷.

Cada vez mais surgem dúvidas quanto à utilidade dos instrumentos contidos nesta Diretiva, havendo quem defenda a necessidade de repensar este ato legislativo à luz do fenómeno de Big Data⁵⁸.

4. Transformações na concepção tradicional das Bases de dados:

Como temos vindo a constatar, com a crescente globalização a permitir a interligação de pessoas em todo o mundo, e a Internet das Coisas, por sua vez, a permitir a interligação de máquinas e dispositivos, a criação e o fluxo de dados aumentaram extraordinariamente. Com este exponencial crescimento do volume de dados, verificamos que, como já se disse, os sistemas tradicionais de coleta, tratamento e análise de dados se mostraram incapazes de responder às exigências impostas pelos grandes volumes de dados. Em consequência, desenvolveram-se novas arquiteturas e tecnologias de base de dados⁵⁹.

As tecnologias digitais vieram simplificar o desenvolvimento de bases de dados, as quais permitem, hoje, a digitalização de uma enorme quantidade de materiais, poupando, assim, quantidades extraordinárias de tempo, espaço físico e recursos financeiros⁶⁰.

Importa notar que uma base de dados, em sentido técnico, diz respeito a um conjunto de arquivos relacionados entre si, com registos sobre pessoas, lugares ou coisas. São

⁵⁶ PAIS, 2020, p. 16.

⁵⁷ Diretiva 96/9/CE de 11 de março de 1996 relativa à proteção jurídica das bases de dados.

⁵⁸Cfr. LEISTNER, 2018; ZEITLIN, 2018; ALEJANDRE, 2010.

⁵⁹ CHEN et al., 2014, pp.184-185.

⁶⁰ CORREIA, 2018.

coleções organizadas de dados que se relacionam de forma a criar alguma informação e dar mais eficiência durante uma pesquisa ou estudo científico⁶¹.

Cumprido, então, analisar as diferenças entre as bases de dados tradicionais e as bases de dados atuais, as quais, para efeitos do presente trabalho, serão denominadas bases de dados em contexto ou decorrentes de Big Data.

As bases de dados tradicionais, ou relacionais, utilizam o Structured Query Language (SQL) como linguagem standard e armazenam os dados de acordo com as suas relações em tabelas constituídas por linhas e colunas. Estes sistemas de bases de dados relacionais⁶², estiveram em voga durante muitos anos, constituindo-se como poderosos mecanismos de armazenamento e consulta de dados estruturados, com garantia de consistência⁶³.

Contudo, como já foi observado, o fluxo de dados aumentou exponencialmente nos últimos anos, impondo sérias exigências relativamente ao seu processamento, exigindo uma maior velocidade, esquemas mais flexíveis e bases de dados distribuídas⁶⁴.

Face a este panorama de crescimento, foram surgindo também problemas, na medida em que as bases de dados tradicionais, se mostraram incapazes de operar em contexto de Big Data. Como consequência direta da necessidade de operar em contextos de dados massivos, foram desenvolvidas novas classes de bases de dados, como é o caso do Not Only SQL, doravante NoSQL⁶⁵, capaz de lidar com vastas quantidades de dados distribuídos⁶⁶.

Contrariamente às bases de dados tradicionais (ou RDBMS), as do tipo NoSQL, são bases de dados de código aberto, o que significa que estas não possuem um esquema. Logo, são flexíveis e compostas por uma variada gama de arquiteturas e tecnologias que permitem solucionar problemas de desempenho e escalabilidade, – o que se revela especialmente útil num ambiente de Big Data que envolve dados não estruturados e semiestruturados.

⁶¹ Cfr. https://pt.wikipedia.org/wiki/Banco_de_dados.

⁶² NOYES, K 2015.

⁶³ Cfr. WINGERATH et al 2017.

⁶⁴ CORREIA, p.29.

⁶⁵ Revela-se importante clarificar que NoSQL não é uma base de dados, nem um tipo de base de dados, é sim um termo genérico para referir todas as bases de dados que não seguem os modelos relacionais tradicionais e que tentam resolver os problemas de escalabilidade e disponibilidade contra a atomicidade ou consistência.

⁶⁶ Cfr. ZHANG & CHEN, 2014.

Importa, ainda, salientar que as bases de dados do tipo NoSQL não vieram substituir as bases de dados SQL, mas sim, complementá-las⁶⁷.

5. A Tutela Jurídica das Bases de Dados:

Os Big Data, são muitas vezes descritos como o combustível da economia digital. Como tal, revestem-se de uma crescente importância no que respeita à inovação e à competitividade europeia, desempenhando um papel primordial na construção de um mercado único digitalizado e funcional. Nesse sentido, é imperativo que as suas bases de dados sejam protegidas pela legislação europeia e nacional⁶⁸.

Como já se deixou dito, o que caracteriza os Big Data, é a capacidade de extração e aproveitamento de informações a partir de dados estruturados, semiestruturados e não estruturados, as quais não poderiam, certamente, ser obtidas sem um enorme volume e variedade de dados. Assim, como seria de prever, estes dados foram aumentando, não só em volume, como em valor.

Já no final do século passado, as bases de dados eram consideradas bens valiosos que implicavam, porém, avultados investimentos para a sua criação. Nos dias de hoje, isso já não se verifica, porquanto, com a constante evolução tecnológica, o custo de armazenamento de dados tornou-se inferior ao custo da eliminação destes mesmos dados⁶⁹.

No entanto, como em muitas áreas, a evolução é acompanhada de alguns riscos. Para além dos elevados esforços investidos na sua criação, as bases de dados, apresentavam a desvantagem de ser facilmente copiadas, suscitando riscos acrescidos para os seus fabricantes.

Perante este panorama de insegurança, provocado pela insuficiente proteção concedida às bases de dados, não havia qualquer estímulo ao investimento nas mesmas, o que, por sua vez, não promovia o crescimento do mercado⁷⁰. Por outro lado, havia uma fragmentação no que respeita à sua regulação nos Estados-Membros, o que colocava grandes entraves ao funcionamento do mercado interno⁷¹.

⁶⁷ Cfr. BATHLA et al. 2018.

⁶⁸ PAIS, 2020, p.15.

⁶⁹ PILA & TORREMANS, 2020.

⁷⁰ Considerandos 1 e 11 da Diretiva 96/9/CE.

⁷¹ Considerandos 2 e 3 da Diretiva 96/9/CE.

Neste sentido, na União Europeia, face à ausência de uma proteção adequada, e à necessidade de uma uniformização da legislação interna dos países comunitários quanto ao regime da proteção de bases de dados, o Parlamento Europeu e o Conselho da União Europeia adotaram a Diretiva 96/9/CE de 11 de março de 1996, relativa à proteção jurídica das bases de dados (doravante designada por Diretiva), tendo sido transposta para o ordenamento jurídico português pelo DL n.º122/2000, de 4/7.

5.1. A definição de base de dados à luz da Diretiva:

A definição de base de dados encontra-se plasmada no n.º 2 do artigo 1.º da Diretiva, que define base de dados como “uma coletânea de obras, dados ou outros elementos independentes, dispostos de modo sistemático ou metódico e suscetíveis de acesso individual por meios eletrónicos ou outros”, ficando, desde logo, excluídas dos limites da Diretiva todas as compilações que não se enquadrem nesta definição.

Posto isto, é importante esclarecer que o conceito de base de dados delimitado pela Diretiva consubstancia-se como requisito de aplicação da mesma, não tendo capacidade de determinar, *per se*, a proteção quer pelo direito de autor, quer pelo direito *Sui Generis*⁷².

Por fim, pela análise deste artigo, podemos concluir que devem ser preenchidos quatro requisitos cumulativos para se considerar que estamos perante uma base de dados⁷³.

a) Bases de dados sob qualquer forma

A Diretiva, começa por elucidar que a sua aplicabilidade é abrangente, na medida em que será aplicável a qualquer forma de base de dados⁷⁴. O que vem, aliás, reforçado nos considerandos.

⁷² ALEJANDRE, 2010, p. 29 .

⁷³ VENÂNCIO, 2012, p. 612.

⁷⁴ No n.º1 do artigo 1.º da proposta da Diretiva, Base de dados era definida como “-, uma compilação de obras ou materiais dispostos, armazenados e a que se tem acesso por meios eletrónicos, incluindo o material eletrónico necessário à exploração da base de dados, como o thesaurus, os índices, o sistema de acesso à informação e sua apresentação ; a expressão não abrange os programas de computador utilizados na criação ou exploração da base de dados.” o que excluía todos os outros tipos de bases de dados, o que, por consequência, limitava o surgimento de novos tipos de bases de dados não eletrónicas Disponível em: <https://eur-lex.europa.eu/legal-content/PT/TXT/PDF/?uri=CELEX:51992PC0024&from=EN>

O considerando 13 determina que a Diretiva abrange a proteção das compilações de coleções "que incluam processos eletrônicos, eletromagnéticos ou eletro-óticos ou processos análogos". E o considerando 14, por sua vez, veio "alargar a proteção conferida pela presente às bases de dados não eletrônicas". Da conjugação do artigo primeiro com os mencionados considerandos, podemos afirmar que, a proteção irá recair, não só nas bases de dados informatizadas, como também nas bases de dados não informatizadas⁷⁵ – tais como listas de clientes ou listas telefônicas em suporte de papel.

Do que foi dito, podemos retirar como observação geral que a Diretiva tem propositadamente um vasto âmbito de aplicação, que deve ser tido sempre em conta na interpretação do seu âmbito de aplicação⁷⁶.

b) Bases de dados que são coleções de obras, dados ou outros elementos independentes

Em segundo lugar, a Diretiva refere-se ao acervo que pode constituir uma base de dados. Salienta, ainda, que a mesma não se dirige apenas às coletâneas de obras, mas também a quaisquer outros dados ou elementos independentes⁷⁷.

Dentro destas duas categorias, as "obras"⁷⁸ devem abranger conteúdos dotados de um certo grau de originalidade, por forma a estarem em conformidade com a legislação europeia relativa aos direitos de autor. Por sua vez, o conceito de "dados ou outros elementos independentes" mostra-se como um conceito vago, sobretudo, tendo em conta a exigência de independência⁷⁹.

A questão da independência, na prática, significa que os conteúdos armazenados em bases de dados são separados e não interagem entre si. Ou seja, nem a fixação de uma obra, nem a própria obra podem ser consideradas individualmente como bases de dados para efeitos da Diretiva, já que as partes em que podem ser decompostas não são independentes umas das outras e não podem ser acessíveis individualmente⁸⁰. Esta regra

⁷⁵Cfr. https://itlaw.fandom.com/wiki/EU_Database_Directive?fbclid=IwAR38W7Y6adwUAmvZ_XspE7pIP2i1s9YLxWpFbpJyOsPqWzPVjP-MYD7--r8

⁷⁶ VENÂNCIO, 2012.

⁷⁷ Os quais podem ser entendidos como dados de qualquer natureza.

⁷⁸O Art. 1.º, n.º 1, do Código do Direito de Autor e Dos Direitos Conexos (CDAC), retira-se uma noção unitária de obra enquanto "criação intelectual, do domínio literário, científico ou artístico, por qualquer modo exteriorizada". Neste âmbito Cfr. ROCHA, 2011.

⁷⁹ A Advogada-Geral Stix-Hackl considera que o requisito de dados independentes ou outro material significa que o conteúdo da base de dados não pode ser ligado ou deve pelo menos ser capaz de ser separado sem perder o seu conteúdo informativo.

⁸⁰ PILA & TORREMANNS, 2019, p. 488.

da não interação, foi confirmada pelo TJUE no Processo C-444/02 Fixtures Marketing Ltd contra OPAP⁸¹ que descreveu estes elementos independentes como aqueles "elementos que são separáveis uns dos outros sem afetar o valor do seu conteúdo informativo, literário, artístico, musical ou outro".

c) Organizadas de forma sistemática ou metódica

Um terceiro requisito a considerar, é o de que os dados devem estar dispostos, na base de dados, de forma "sistemática ou metódica". Melhor dizendo, os dados devem ser organizados de alguma forma lógica ou racional⁸².

A Diretiva não define o que deve ser entendido como uma disposição sistemática ou metódica. No entanto, o TJUE⁸³ confirmou que uma disposição sistemática ou metódica pressupõe a existência de um método de classificação específico para permitir a recuperação de qualquer material independente contido na base de dados. O que implica, desde logo, a exclusão do âmbito da definição de bases de dados as acumulações de dados completamente não estruturadas, isto é, meros fluxos de dados recolhidos indiscriminadamente.

Enquanto houver um meio técnico ou outro meio⁸⁴, tal como um índice, ou um método particular de classificação, que permita a recuperação de uma obra ou outro elemento, os requisitos serão atendidos⁸⁵.

d) Individualmente acessíveis por meios eletrónicos ou outros

Por último, analisámos o requisito da acessibilidade individual. Muitos autores questionaram se existe alguma diferença entre o requisito da independência e o da acessibilidade individual. Somos levados a crer que o significado de acessibilidade individual seria que os elementos devem ser recuperáveis separadamente. No entanto, o significado do requisito não é claro. Provavelmente, a sua inclusão pelo legislador teve o

⁸¹ Proc C-444/02, Fixtures Marketing, par. 29.

⁸² O considerando 21 veio excluir a necessidade de "que essas matérias tenham sido fisicamente armazenadas de modo organizado;".

⁸³ Proc. C-444/02 Fixtures Marketing.

⁸⁴ Tal como um índice ou um método particular de classificação, lógico, conveniente e ordenado.

⁸⁵ ELFERING, 2019, p.35; DERCLAYE, 2008, pp. 65-67.

objetivo de garantir claramente que os dados coletados de forma aleatória não se pudessem qualificar como uma base de dados⁸⁶.

A Diretiva aponta, ainda, um requisito de delimitação negativa. Dada a proliferação das novas tecnologias, o artigo 1º, mais especificamente no nº3, acrescenta que a mesma “não é aplicável a programas de computador utilizados no fabrico ou no funcionamento de bases de dados acessíveis por meios eletrónicos.” Estes têm necessariamente que ser demarcados das bases de dados e são (potencialmente) cobertos pela Diretiva relativa à proteção jurídica dos programas de computador, conforme vem especificado no considerando 23⁸⁷.

Contudo, centrando a nossa atenção no considerando 20 da Diretiva, o mesmo especifica que “a proteção prevista na presente Diretiva pode ser igualmente aplicável aos elementos necessários ao funcionamento ou à consulta de certas bases de dados, como o thesaurus⁸⁸ e os sistemas de indexação”, o que se mostra contraditório.

Comparando os artigos e considerandos acima, pode concluir-se que há margem para a interpretação. Assim, a proteção será concedida consoante o caso específico possa ser qualificado como "elemento necessário para a operação ou consulta" ou como "programa informático utilizado no fabrico ou operação", sendo que a delimitação de uma ou outra categoria é uma questão-chave.

6. A dupla proteção Europeia

No sentido de atingir os objetivos enunciados, a Diretiva veio consagrar um regime de proteção de bases de dados duplo⁸⁹, cumulando a proteção pelo direito de autor e o inovador direito Sui Generis do fabricante das bases de dados, (doravante designado por direito Sui Generis).

⁸⁶ A autora Estelle Derclaye considera que este requisito “é provavelmente redundante e, portanto, desnecessário.” DERCLAYE, 2008, p.66.

⁸⁷ A “base de dados” não deve aplicar-se aos programas de computador utilizados no fabrico ou no funcionamento de uma base de dados, que são protegidos pela Diretiva Diretiva 2009/24/CE de 23 de abril de 2009 relativa à proteção jurídica dos programas de computador transposta para o ordenamento jurídico português pelo Decreto-Lei n.º 252/94, de 20 de outubro.

⁸⁸ Em ambiente de base de dados, um thesaurus consiste num nome alternativo para uma tabela, visão, seqüência ou outro objeto de esquema. São usados principalmente para facilitar o acesso dos usuários a objetos da base de dados - [https://en.wikipedia.org/wiki/Synonym_\(database\)](https://en.wikipedia.org/wiki/Synonym_(database))

⁸⁹ Cfr. VENÂNCIO, 2012.

Enquanto os direitos de autor visam proteger a criação intelectual através da atribuição de um direito de autor sobre a forma de expressão (protegem a estrutura das bases de dados e, conseqüentemente, a seleção ou disposições do seu conteúdo) da base – desde que aquela forma de expressão seja original⁹⁰ – o direito Sui Generis, por sua vez, protege as bases de dados independentemente da sua originalidade, desde que tenha existido um investimento substancial na obtenção, verificação ou apresentação do conteúdo.

Os diferentes direitos concedidos pelos dois artigos conferem uma proteção sobreposta em que, tanto os direitos de autor como a proteção Sui Generis podem ser aplicáveis independentemente da existência do outro – não são dependentes ou pré-requerentes⁹¹.

O Direito Sui Generis foi pensado no sentido de atribuir algum tipo de proteção às bases de dados que não preenchiam o requisito de proteção dos direitos de autor. A enorme dificuldade no preenchimento do requisito da originalidade acabava por excluir um grande número de bases de dados, como se demonstrará. O que, por seu turno, se refletia num desincentivo por parte dos seus fabricantes, gerando uma enorme desvantagem no mercado europeu em relação a outros mercados⁹².

6.1. A proteção de bases de dados numa perspectiva de direitos de autor

A proteção das bases de dados ao abrigo dos Direitos de Autor encontra-se no capítulo II da Diretiva, sob a epígrafe “Direito de Autor”.

Da análise do nº1 do artigo 3º, será possível constatar que a elegibilidade para proteção é determinada pela "razão da seleção ou disposição das matérias". O que automaticamente nos afasta do conteúdo das bases de dados, e nos remete para a estruturas das mesmas⁹³. Assim, protege-se apenas a forma como estes dados são organizados dentro da base de dados, como podem ser visualizados e como podem ser

⁹⁰ MARQUES, 2007, p. 847.

⁹¹ VENÂNCIO, 2012.

⁹² Como o mercado dos Estados Unidos.

⁹³ O que por sua vez vem confirmado no considerando 15 da Diretiva ao esclarecer que o objeto do direito de autor “incide sobre a estrutura da base”.

accedidos e pesquisados.⁹⁴ As obras, os dados e outros elementos que a constituem⁹⁵ e não se encontrem protegidos não serão abrangidos pelos direitos de autor através da sua mera inclusão na base de dados. No mesmo sentido, qualquer direito pré-existente sobre o conteúdo da mesma não será afetado. Os direitos de autor numa base de dados limitam-se, assim, à seleção e disposição ou estrutura dos materiais que nela estão contidos⁹⁶.

Os considerandos da Diretiva esclarecem, ainda, que “não devem aplicar-se outros critérios que não o da originalidade, na aceção da criação intelectual, para determinar se a base de dados é suscetível de proteção pelo direito de autor, e que, em especial, não deverão intervir critérios estéticos ou qualitativos”⁹⁷.

Concluimos, desde logo, que a mera justaposição de elementos, sem qualquer reflexo da originalidade do seu autor, não atinge o perímetro dos direitos de autor delineado na Diretiva, a mesma reserva a sua aplicação a bases de dados cuja originalidade seja suscetível de imputação subjetiva. Ficando, assim, excluídas deste âmbito, por exemplo as bases de dados de jurisprudência e as bases de dados para fins estatísticos.

No processo *Football Dataco*⁹⁸, o TJUE concluiu que o “trabalho considerável e os conhecimentos necessários para a criação de tal base de dados não podem, por si só, justificar a proteção se não exprimirem qualquer originalidade na seleção ou disposição dos dados que contém”.

Assim, a originalidade – ponto chave de aplicação do direito de autor – opera como linha de fronteira, no acesso a este direito. A sua concreta configuração exprime simultaneamente opções de política legislativa⁹⁹, divergindo em grande medida nas tradições de copyright, que se centra na proteção do investimento (what is worth copying is worth protecting) e de *droit d’auteur*, assente na pessoa do criador e nos direitos de autor como direitos naturais do homem¹⁰⁰.

Este é um conceito envolto em alguma ambiguidade na medida em que, em regra, a legislação sobre Direito de Autor, ou não refere o conceito, ou limita-se a referi-lo sem indicar o seu conteúdo. Revelando, portanto, a necessidade de uma densificação normativa. Face às dificuldades na sua definição, tanto a Doutrina como a Jurisprudência,

⁹⁴ VENÂNCIO, 2012, p.622

⁹⁵ Pedro Dias Venâncio afirma que os metadados (theaurus e sistemas de indexação), se revestidos de originalidade na sua criação também poderão ser protegidos pelos direitos de autor VENÂNCIO, 2012, p.146.

⁹⁶ PILA & TORREMANS, 2019, p. 489.

⁹⁷ Considerando 16 da Diretiva 96/9/CE.

⁹⁸ Proc C-604/10 *Football Dataco*.

⁹⁹ SILVA, 2013, p.1365.

¹⁰⁰ Cfr. ROCHA, 2008.

têm delimitado o conceito de originalidade, variando em função da tradição acolhida em cada ordenamento jurídico.

Desse modo, e tendo por base a Jurisprudência europeia, concluímos que será protegida por direitos de autor uma base de dados em que algum juízo intelectual do próprio autor tenha contribuído para a seleção ou disposição das matérias, estando aí – na estrutura – manifestadas as “escolhas livres e criativas deste último”¹⁰¹.

Em síntese, parece resultar, ainda que de forma implícita, que a “criação intelectual específica do respetivo autor”, à qual o artigo 3º nº 1 da Diretiva se refere está reservada apenas aos humanos.

É fundamental sublinhar que a Diretiva foi adotada numa época em que vigorava o analógico, em que a quantidade disponível de dados não atingia, nem de perto, os valores que atingem nos dias de hoje.

Parece-nos que esta abordagem, nos dias de hoje, impossibilita a atribuição de direitos de autor sobre bases de dados. Nas bases de dados tradicionais, é possível, embora muito difícil, que um indivíduo, despendendo de vários esforços intelectuais, consiga criar, de facto, uma base de dados criativa e, como tal, suscetível de proteção de direitos de autor. Numa época marcada pelos Big Data, a abordagem adotada negará a proteção abrangida pelos Direitos de Autor¹⁰².

Como foi analisado anteriormente, os Big Data caracterizam-se dessa forma pelo seu volume, velocidade e variedade e que as bases de dados no seu contexto são tipicamente não estruturadas e não relacionais. Assim, pela sua natureza, dificilmente são selecionadas ou organizadas de forma a satisfazerem suficientemente o limiar da originalidade, uma vez que, além do mais, organizam os seus materiais de forma standard, ou seja, por ordem alfabética ou outra forma padrão. O que não revela qualquer originalidade ¹⁰³.

Voltando a atenção para os 3 V’s em Big Data, – volume, variedade, velocidade, – torna-se claro que os Big Data tendem a operar de forma automatizada, sendo regulados por algoritmos pré-determinados. Desde logo, em contexto de Big Data, as bases de dados

¹⁰¹ Cfr. Acórdão Painer; Renckhoff; Cofemel.

¹⁰² Cfr. PILA & TORREMANS, 2019.

¹⁰³ Tendemos a concordar com Justine Pila e Paul Torremans, quando é afirmado que esta abordagem deve ser aplaudida, porque a alternativa teria envolvido a concessão de múltiplos direitos de autor a estruturas padrão e comuns, e um tal múltiplo de direitos exclusivos poderia ter asfixiado a concorrência nesta área. - PILA & TORREMANS, 2019, p.489.

são criadas automaticamente, sem o envolvimento de um autor, o que inviabiliza a atribuição da criação de uma base de dados a um determinado autor humano.

Posto isto torna-se impossível a proteção ao abrigo dos Direitos de Autor constante na Diretiva.

6.2 O Direito Sui Generis

Fruto da ausência de uma adequada proteção e da dificuldade na aplicação dos Direitos de autor, surge o Direito especial do fabricante Sui Generis.

O artigo 7.º da Diretiva define-o como o direito que o fabricante de uma base de dados tem de “proibir a extração e/ou reutilização da totalidade ou de uma parte substancial, avaliada qualitativa ou quantitativamente, do conteúdo desta, quando a obtenção, verificação ou apresentação desse conteúdo representem um investimento substancial do ponto de vista qualitativo ou quantitativo”¹⁰⁴.

Este prevê um direito de propriedade que é concedido ao fabricante de uma base de dados, com vista a proteger o seu investimento, incluindo, desde logo, as bases de dados não criativas uma vez que o único requisito é um investimento substancial.

Contrariamente ao que foi analisado na proteção concedida pelos Direitos de Autor, a qual incide sobre a estrutura da base de dados sem afetar o seu conteúdo, a proteção concedida pelo direito Sui Generis recai especificamente sobre o conteúdo unitário da base de dados, sem interferir com os dados ou outros elementos existentes¹⁰⁵.

Da análise do artigo 7º, depreende-se que o principal elemento suscetível de desencadear a proteção ao abrigo do direito Sui Generis, é a existência de um investimento substancial na obtenção, verificação e apresentação da base de dados, ficando o ónus da prova desse investimento a cargo do próprio criador da base de dados.

Desde logo, surge a questão de saber o que poderá ser entendido como investimento. A resposta a esta questão é fornecida pela própria Diretiva nos seus considerandos¹⁰⁶, que preveem um espectro amplo de investimentos, que tanto podem resultar da elevada

¹⁰⁴ O artigo 12.º n.º 1 do DL n.º 122/2000 consagra-o da seguinte forma, “quando a obtenção, verificação ou apresentação do conteúdo de uma base de dados represente um investimento substancial do ponto de vista qualitativo ou quantitativo, o seu fabricante goza do direito de autorizar ou proibir a extração e/ou a reutilização da totalidade ou de uma parte substancial, avaliada qualitativa ou quantitativamente, do seu conteúdo”.

¹⁰⁵ Considerandos 6, 8, 18 da Diretiva 96/9/CE

¹⁰⁶ Considerando 40 da Diretiva 96/9/CE.

natureza “qualitativa,” ou seja, o investimento baseado nos esforços técnicos e profissionais, como da sua grandeza “quantitativa,” ou investimento financeiro¹⁰⁷. Alguns autores referem que teria sido suficiente referir-se à grandeza quantitativa, dado que qualquer utilização de recursos técnicos e humanos pode ser traduzida em termos económicos¹⁰⁸.

Em geral, qualquer tipo de recurso – económico ou não – será tido em conta, desde que o fabricante possa provar a sua existência¹⁰⁹.

Apresentado o conceito de “investimento”, cumpre, agora, analisar o conceito “substancial” que, tal como o requisito da originalidade no âmbito dos Direitos de Autor, está envolto em opiniões distintas sem qualquer consenso aparente. A própria Diretiva não revelou qualquer esforço no sentido de desmitificar este conceito, tarefa essa que deixa para a Doutrina e para a Jurisprudência¹¹⁰.

Numa tentativa de aclarar este conceito, Estelle Derclaye¹¹¹ sugeriu que o carácter substancial deverá ser aplicado no caso concreto, ou seja, avaliado proporcionalmente aos meios e recursos do fabricante da base de dados.

Esta, seria no nosso entender, a avaliação mais adequada a realizar aquando da aferição da substancialidade do investimento. Se esta avaliação fosse realizada abstratamente, sem qualquer consideração pelos recursos da entidade fabricante, haveria um claro desequilíbrio na atribuição deste direito, uma vez que, provocaria uma desproteção dos pequenos produtores – que eventualmente empregaram todos os seus esforços e que, do seu ponto de vista, eram substanciais, mas que de uma perspetiva global não poderiam ser qualificados como tal.

O TJUE também não determinou qual o nível que um investimento deverá atingir para ser considerado substancial. Contudo, os tribunais nacionais têm-se debruçado sobre esta questão, sendo possível afirmar que a falta de uma delimitação do nível de substancialidade é suscetível de gerar insegurança jurídica. Existem casos em que é clara

¹⁰⁷ DERCLAYE, 2008, p.73.

¹⁰⁸ ALEJANDRE, 2010, p. 37.

¹⁰⁹ Assim, a substancialidade, qualitativa ou quantitativa, do investimento necessário à construção da base de dados apresenta-se não só como pressuposto da proteção dada pelo direito especial do fabricante, mas também como limite desta proteção.

¹¹⁰ O considerando 19 da Diretiva é bastante útil para determinar, a contrário, o que deve ser entendido neste sentido por substancial. Se o investimento feito ao compilar fixações de espetáculos musicais em formato CD não for um investimento substancial, já temos uma ideia do que aos olhos do legislador europeu não merece a proteção do direito Sui Generis e, portanto, podemos prever o que o legislador europeu pode considerar ser um investimento que, por ser suficientemente substancial, é digno de proteção.

¹¹¹ DERCLAYE, 2008.

a existência de um investimento substancial. No entanto, a interpretação dada a estes conceitos por parte do Tribunal tem variado de caso para caso¹¹².

Apesar das diferentes abordagens jurisprudenciais e doutrinárias, a exigência de "substancialidade" continua a ser um conceito jurídico indeterminado¹¹³.

Ainda na análise da atribuição do direito Sui Generis, os esforços técnicos e/ou financeiros, cumprido o requisito da substancialidade, devem incidir na obtenção verificação ou apresentação, sendo estes atos não cumulativos, na medida em que qualquer um torna o investimento elegível para proteção¹¹⁴.

A) Obtenção

O investimento na obtenção deve ser entendido como aquele que é feito nos recursos dedicados à pesquisa de dados existentes e à sua posterior compilação na base de dados.

No entanto, a subjetividade de interpretação deste conceito faz emergir abundante discórdia no seio doutrinário e jurisprudencial.

O TJUE, no Processo C-203/02¹¹⁵, decidiu que a obtenção deverá ser interpretada como referente aos recursos utilizados na procura de materiais independentes e existentes. Por outras palavras, o TJUE veio delimitar o âmbito do conceito de investimento na obtenção e cingindo-o aos dados já existentes, ficando excluindo desta forma o investimento realizado na criação dos dados.

Conforme salienta Remédio Marques, “o que motivou a proposta da Diretiva comunitária de 1996, terá sido a intenção do legislador promover o armazenamento e o processamento dos dados ou informações preexistentes e não promover a criação de novos dados, elementos ou materiais”¹¹⁶.

Desta leitura, levanta-se, desde logo, a problemática advinda das situações em que a criação da informação coincide com a criação da base de dados, também designados de base de dados de fonte única.

No entendimento do TJUE, o direito Sui Generis só pode proteger as bases de dados que são o objeto principal da atividade dos seus fabricantes, e não aquelas que foram

¹¹² Por exemplo, o Supremo Tribunal Federal Alemão, considerou que o requisito seria objetivamente cumprido, se não fossem necessários gastos insignificantes para criar o banco de dados. BGH Processo ZR196/08 – Zweite Zahnarztmeinung.

¹¹³ DERCLAYE, 2005, p. 7.

¹¹⁴ ELFERING, 2019, p.37.

¹¹⁵ Proc. C-203/02 The British Horseracing Board.

¹¹⁶ MARQUES, 2007, p.875.

criadas automaticamente como produtos secundários derivados da atividade principal do seu criador.

Com esta decisão, houve, aparentemente, por parte do TJUE, uma clara aproximação à teoria do Spin-Off, segundo a qual os subprodutos não são cobertos pelo direito.

Contudo, no caso *Fotball Dataco*¹¹⁷, o TJUE introduziu uma exceção, que se revelou favorável para os fabricantes de bases de dados cujos dados não são provenientes da sua atividade principal. Deste modo, veio determinar que, nas situações em que a criação de uma base de dados esteja ligada a uma atividade principal e em que o criador da base de dados também cria o conteúdo, a proteção não ficará automaticamente excluída, desde que, seja provada a existência de um investimento substancial realizado de forma independente na obtenção, verificação ou apresentação do conteúdo¹¹⁸.

Esta decisão, apesar de provocar um alargamento no âmbito de proteção concedido pela Diretiva, não será isenta de problemas, na medida em que será sempre difícil a tarefa de distinção entre o investimento feito na criação dos dados como consequência do exercício da sua atividade principal e o investimento que a pessoa pode posteriormente fazer na obtenção dos dados¹¹⁹.

B) Verificação

O investimento na **verificação** pode ser entendido com o investimento decorrente da necessidade fornecer ao utilizador informações precisas e atualizadas.

De acordo com o TJUE, este diz respeito aos "recursos utilizados, com vista a assegurar a fiabilidade da informação contida nessa base de dados, para controlar a exatidão dos materiais recolhidos aquando da criação da base de dados e durante o seu funcionamento"¹²⁰.

Assim, poderá consistir no investimento realizado na correção de registos duplicados, a eliminação de erros e na atualização do conteúdo, no sentido de assegurar a fiabilidade dos dados tanto no momento da recolha como continuamente durante o funcionamento da base de dados¹²¹.

¹¹⁷ V. Proc. C-604/10, *Football Dataco*.

¹¹⁸ Cfr. *The British Horseracing; Fixtures Marketing Ltd e OPAP; Fixtures Marketing Ltd e Svenska Spel AB*.

¹¹⁹ Cfr. ALEJANDRE p.43.

¹²⁰ V. Proc. C-203/02 *British Horseracing*, par. 37.

¹²¹ Cfr. ALEJANDRE.

Mais uma vez, o criador da base de dados deve provar que o investimento, nos meios utilizados é revestido de substancialidade com vista à obtenção da proteção Sui Generis. É igualmente necessário repetir que os custos de verificação durante a criação não podem ser incluídos na avaliação da substancialidade do investimento. Consequentemente, a proteção Sui Generis só pode surgir através da verificação de um conjunto de dados existentes.

C) Apresentação

O investimento na apresentação trata-se, essencialmente, de um investimento na estruturação da base de dados. Na perspetiva do TJUE, um investimento será considerado na apresentação quando diga respeito "aos recursos dedicados a conferir à base de dados, recursos dedicados à disposição sistemática ou metódica dos dados contidos na base de dados e à organização da sua acessibilidade individual"¹²².

Este investimento refere-se especificamente à forma como os dados serão apresentados. Assim, a estruturação de dados que não se encontravam estruturados e o investimento no sistema de pesquisa ou acesso a itens individuais poderão ser considerados como um investimento na verificação¹²³.

6.2.1 Titularidade

Nas situações em que, efetivamente, se esteja perante um investimento substancial, qualitativo e/ou quantitativo, na obtenção, verificação ou apresentação, o artigo 7º, nº1 prevê que a titularidade da proteção conferida pelo direito Sui Generis será atribuída ao fabricante da base de dados.

Na lógica do Artigo 7º, - em conjugação com o considerando 41 - o criador da base de dados é a pessoa que toma a iniciativa de obter, verificar, ou apresentar o conteúdo da desta, assumindo, desse modo o risco de investir nessa obtenção, verificação ou

¹²² Cfr. Fixtures Marketing Ltd e OPAP; Fixtures Marketing Ltd e Oy Veikkaus Ab.

¹²³ Nos casos Fixtures Marketing Ltd e OPAP, e Fixtures Marketing Ltd e Oy Veikkaus Ab, o TJUE decidiu que as listas de fixações de jogos de futebol não eram investimentos independentes na apresentação do conteúdo, uma vez que a apresentação em si não exigia qualquer esforço particular que pudesse ser considerado independentes para a criação das listas.

apresentação. Significa isto que os subempreiteiros não serão considerados fabricantes, nem aqueles que trabalhem por encomenda, por conta de outrem¹²⁴.

A Diretiva não faz qualquer menção quanto à natureza do fabricante. Nesse sentido, consideramos que o fabricante da base de dados poderá ser, tanto uma pessoa singular, como uma pessoa coletiva.

Da mesma forma, não foi dada qualquer atenção à questão da cotitularidade. Presumimos, assim, que esta é possível. Logo, nos casos em que duas ou mais pessoas assumam a iniciativa e o risco do investimento na criação de uma base de dados, efetivamente poderá haver copropriedade¹²⁵.

6.2.2 Proteção

O titular do direito Sui Generis terá o poder exclusivo de “proibir¹²⁶ a extração e/ou a reutilização da totalidade ou de uma parte substancial” da base de dados, conforme vem previsto no artigo 7º, nº1 da Diretiva.

A proibição de extração e reutilização de partes substanciais do conteúdo não abrange apenas os atos dos concorrentes, mas qualquer ato que cause “prejuízo significativo, avaliado qualitativa ou quantitativamente, ao investimento”¹²⁷.

Mais uma vez, a Diretiva não fornece nenhuma definição do conceito “substancial”. Neste âmbito, no que diz respeito à parte substancial do conteúdo, o TJUE tem decidido que deve ser apreciado em relação ao volume do conteúdo total da base de dados¹²⁸.

Quanto aos termos qualitativos e quantitativos devem, de acordo com o TJUE, “reportar-se ao investimento ligado à constituição da base de dados e ao prejuízo causado ao investimento pelo ato de extração e/ou de utilização relativo a esta parte”¹²⁹. Recentemente, o AG Spuznar, no caso CV Online Latvia, veio determinar que o principal

¹²⁴ VENÂNCIO, 2014.

¹²⁵ Cfr. ELFERING, 2019 p.39.

¹²⁶ Conforme salienta Pedro Venâncio, em Portugal difere da formulação comunitária, na medida em que “o legislador nacional deu ao conteúdo do direito especial do fabricante de base de dados uma dupla faculdade positiva e negativa – autorizar ou proibir” O Autor afirma que o direito comunitário por sua vez atribui-lhe uma formulação negativa, no entanto, analisando o artigo 7º nº3 da Diretiva verificamos que, “o direito previsto no nº1 pode ser transferido, cedido ou objeto de licenças contratuais.” O que tal como o direito português, não descurando da visão protecionista da atividade económica inerente à produção de bases de dados, concede um direito de fruir da base de dados criada em exclusão de todos os demais. VENÂNCIO, 2014, p.29.

¹²⁷ Considerando 42 da Diretiva 96/9/CE.

¹²⁸ Processo C-203/02 The British Horseracing par.70.

¹²⁹ Ver proc. C-203/02 The British Horseracing, par. 69

critério para equilibrar os interesses legítimos em jogo deve ser o risco potencial para o investimento substancial do fabricante da base de dados¹³⁰.

A Diretiva, prevê nas alíneas a) e b) do nº2 do artigo 7º, uma definição de Extração, caracterizando-a como “a transferência permanente ou temporária da totalidade ou de uma parte substancial do conteúdo de uma base de dados para outro suporte, seja porque meio ou porque forma for”. E bem assim, prevê uma definição de Reutilização como sendo “qualquer forma de pôr à disposição do público a totalidade ou uma parte substancial do conteúdo da base através da distribuição de cópias, aluguer, transmissão em linha ou sob qualquer outra forma”. Sendo que, a expressão “qualquer forma”, consiste numa formulação indeterminada abrange, não só os métodos expressamente mencionados no artigo, como também outros que possam existir.

Contudo, o TJUE havia já esclarecido que estes conceitos devem ser interpretados como referindo-se a qualquer ato que consista, respetivamente, na apropriação e na colocação à disposição do público sem o consentimento da pessoa que constituiu a base de dados, privando assim esta última dos rendimentos que lhe permitam amortizar o custo desse investimento¹³¹.

Por outro lado, de acordo com o artigo 8º, o fabricante da base de dados não pode impedir a extração ou reutilização pelo utilizador legítimo de partes não substanciais do conteúdo da base de dados desde que essa extração ou reutilização não implique, pela sua natureza repetida ou sistemática, atos contrários à exploração da base de dados ou cause um prejuízo injustificado aos legítimos interesses do fabricante da base de dados conforme expressa o nº 5 do artigo 7º da Diretiva.

De igual modo, a partir do momento em que uma base de dados é tornada pública, o direito Sui Generis não pode ser utilizado para impedir os utilizadores de consultarem a base de dados. Este poderá, no entanto, reservar os direitos de acesso exclusivos a um grupo específico. A autorização de um utilizador para fazer uma consulta não impede, além disso, o fabricante da base de dados de proibir o utilizador de fazer uma extração ou reutilização¹³².

Desde logo, será presumível que só existam “utilizadores legítimos”, após a base de dados ser posta à disposição pelo fabricante ou pelo titular do direito Sui Generis. Contudo, a Diretiva não oferece a definição do conceito de “utilizador legítimo”.

¹³⁰ Proc C-762/19 CV-Online Latvia

¹³¹ C-203/02 The British Horseracing Board par.51

¹³² Ver Proc.C-203/02 The British Horseracing, par.54.

6.2.3 Exceções

As exceções a este direito encontram-se previstas no artigo 9º, o qual determina que, o utilizador legítimo tem direito a praticar sobre a Base de Dados protegida por Direito Especial do Fabricante: “todos os atos inerentes à utilização obtida; extração para uso privado do conteúdo de uma base de dados não eletrónica; extração para fins didáticos ou científicos; extração ou reutilização para fins de segurança pública, ou para efeitos de um processo administrativo ou judicial”.

Neste âmbito, cumpre de igual forma salientar a exceção prevista nos artigos 3º e 4º da Diretiva (EU) 2019/970 do Parlamento Europeu e do Conselho de 17 de abril de 2019,¹³³ que veio resolver a problemática da prospeção de dados, na sua interferência com os Direitos de Autor e com bases de dados protegidas pelo direito Sui Generis uma vez que violavam os direitos exclusivos¹³⁴.

A Diretiva consagra, no seu artigo 3º um direito de utilização livre e obrigatória para fins de investigação científica, permitindo a extração, reprodução, adaptação e conservação de materiais a que tenham acesso legal para prospeção de dados, por parte de organismos de investigação e por instituições responsáveis pelo património cultural.

O Artigo 4º, por seu turno, consagra uma exceção para as reproduções e as extrações de obras e de outro material protegido legalmente acessíveis para fins de prospeção e textos de dados, desde que essa utilização não tenha sido expressamente reservada pelos respetivos titulares de direitos de forma adequada.

6.2.4 Duração

Relativamente à sua duração, o direito Sui Generis existe por um período de 15 anos. O artigo 10º da Diretiva acrescenta que a proteção estará disponível a partir da data da conclusão da elaboração da base de dados. No entanto, o n.º 3 do artigo 10º contém um aditamento significativo às regras sobre o prazo de proteção.

¹³³ Diretiva (EU) 2019/970 do Parlamento Europeu e do Conselho de 17 de abril de 2019 relativa aos direitos de autor e direitos conexos no mercado único digital e que altera as Diretivas 96/9/CE e 2011/29/CE.

¹³⁴ SILVA 2020 p.251.

De acordo com este artigo, qualquer modificação substancial, avaliada quantitativa ou qualitativamente, do conteúdo de uma base de dados, incluindo quaisquer modificações substanciais resultantes da acumulação de aditamentos ou alterações sucessivas que levem a considerar que se trata de um novo investimento substancial, permitirá atribuir à base resultante desse investimento um período de proteção próprio. Ou seja, qualquer esforço e investimento sustentado para manter a base de dados atualizada conduzirá automaticamente, a uma proteção permanente através da sempre renovada base de dados¹³⁵.

7. Poderão as bases de dados em contexto de Big Data ser protegidos ao abrigo do Direito Sui Generis?

A tarefa primordial no que concerne à aplicabilidade, quer do direito Sui Generis, quer dos Direitos de Autor, consiste no enquadramento das compilações na definição de base de dados, fornecida pelo artigo 1º da Diretiva.

Como tal, ao abordar a potencial proteção das Bases de dados em contexto de Big Data ao abrigo da Diretiva, será imprescindível atender aos requisitos da mesma.

Nesse sentido, para que as Bases de dados, nesse contexto, se enquadrem no perímetro da Diretiva e, como tal, possam beneficiar do direito Sui Generis, estas devem consistir numa coletânea de obras, dados ou outros elementos independentes, e estes elementos terão que se encontrar dispostos de modo sistemático ou metódico, devendo, ainda, ser suscetíveis de acesso individual por meios eletrónicos ou outros.

Não restam dúvidas que a Diretiva abrange todas as formas de bases de dados, sejam bases de dados eletrónicas ou base de dados analógicas. Desta forma, será irrelevante que as bases de dados decorrentes de Big Data se demarquem das bases de dados tradicionais.¹³⁶ De facto, o artigo 1º da Diretiva, revela a intenção de dotá-la de uma maior abrangência no que concerne à sua aplicação, sendo esta livre de considerações de natureza formal, técnica ou material¹³⁷. Deste modo, permite que as Bases de dados em contexto de Big Data se enquadrem na definição fornecida pela Diretiva¹³⁸.

¹³⁵ Neste sentido PILA & TORREMANS, 2020

¹³⁶ Neste sentido, ZEITLIN, 2018, p.23.

¹³⁷ ELFERING, 2019, p.33.

¹³⁸ A AG- Stix-Hackl, no seu parecer relativo ao proc. C-46/02 Fixtures Marketing, no par. 37 afirmou que “são abrangidas quer bases de dados estáticas quer bases de dados dinâmicas”.

Relativamente aos restantes requisitos, começando pelo requisito da independência dos elementos, parece-nos que o mesmo não constitui um problema na medida em que os Big Data, por definição são constituídos por dados massivos e independentes uns dos outros¹³⁹.

Relativamente ao acervo que constitui a base de dados, determina a Diretiva que a base de dados deverá ser composta por “obras,” “dados” e “outros materiais independentes” aos quais é exigido valor informativo, que devem ser separáveis uns dos outros sem perder esse valor.

Em contextos de Big Data, os dados podem assumir as mais variadas formas, isto é, dados estruturados, semiestruturados e não estruturados, sendo que, este último tipo é, na sua maioria, produzido por máquinas, ficando assim excluídos do domínio das obras. Contudo será, sem dúvida alguma, subsumível nos “outros elementos independentes”¹⁴⁰.

No que concerne ao valor informativo, em ambiente de dados estruturados e semiestruturados, será inegável a existência desse valor. A problemática recai nos dados não estruturados, ou dados em bruto, que para adquirirem valor informativo estão dependentes da sua análise.

Contudo, o TJUE veio elucidar que basta que exista um potencial valor informativo. Assim, é nosso entendimento que existe margem para interpretação, na medida em que os dados não estruturados, enquanto dados brutos, não possuem um verdadeiro valor informativo. Não obstante, revestem-se de um potencial valor informativo, dado que, após a sua análise, será efetivamente possível retirar valor dos mesmos.

Por outro lado, vale a pena assinalar que o valor informativo não deverá ser avaliado do ponto de vista do utilizador da base de dados, mas sim do ponto de vista de um terceiro interessado no material extraído, isto é, os dados que não tenham valor informativo para uma empresa, poderão ter valor informativo para outra¹⁴¹.

Segundo a Diretiva, para além dos requisitos já mencionados, para que estejamos perante uma base de dados, os dados devem estar dispostos de forma “sistemática ou metódica”. Por outras palavras, um elemento interno da base de dados deve organizar a informação de forma sistemática e metódica, podendo esta assumir qualquer forma padrão, nomeadamente, a organização por ordem alfabética.

¹³⁹ ZEITLIN p.21.

¹⁴⁰ Artigo 1º nº2 da Diretiva 96/9/CE.

¹⁴¹ Neste sentido, Luís Silva Rodrigues afirma que “o valor e utilidade da informação dependem do contexto em que a informação é utilizada por determinado utilizador” RODRIGUES, 2000, p.27.

Este revela-se um ponto problemático, já que uma base de dados constituída apenas por dados em bruto, não pode, à partida, ser considerada uma base de dados.¹⁴²

No entanto, somos da opinião que as bases de dados automatizadas poderão ser apresentadas numa multiplicidade de combinações diferentes utilizando os devidos meios técnicos, preenchendo, desta forma, o requisito da organização sistemática ou metódica¹⁴³.

Assim, verificamos que a existência de uma base de dados deverá ser analisada individualmente caso a caso, não concordamos com a adoção de uma visão generalista no qual se exclui automaticamente do conceito de bases de dados as bases de dados em contexto de Big Data.

Para que se possa beneficiar do Direito Sui Generis constante na Diretiva, é essencial a existência de um investimento substancial, recaindo o ónus da prova deste investimento sobre o criador da base de dados.

Na recente discussão sobre a relevância do direito Sui Generis para a economia de dados europeia, constatou-se que na maioria das situações de Big Data o requisito do investimento substancial não será preenchido¹⁴⁴.

Esta suposição é fundada no entendimento do TJUE relativamente aos casos *The British Horseracing* e *Fixture Marketing*¹⁴⁵, no qual o Tribunal decidiu que o investimento substancial na obtenção deve ser interpretado como referindo-se aos recursos utilizados para procurar os materiais independentes existentes e para os recolher na base de dados. Ou seja, decidiu-se no sentido que apenas o investimento feito para obter dados já existentes, e não o investimento feito para os criar, será considerado para efeitos da sua proteção pelo direito Sui Generis.

Esta decisão levou muitos autores¹⁴⁶ a afirmar que em contextos de Big Data, os dados gerados por sensores ou máquinas, de todos os tipos serão excluídos do direito Sui Generis, na medida em que tais investimentos teriam que ser considerados como realizados na criação dos dados. Ora, do ponto de vista destes autores, a maior parte dos dados gerados no mundo, (aproximadamente 80%)¹⁴⁷ ficariam excluídos deste âmbito.

¹⁴² DREXL p.12; PILA & TORREMANS 2020, p.488.

¹⁴³ Neste sentido LA DIEGA, 2018; LEINSTER, 2018.

¹⁴⁴ Comissão Europeia – “Study in support of the evaluation of Directive 96/9/EC on the legal protection of databases”, 2018.

¹⁴⁵ Proc. C-203/02 *The British Horseracing*.

¹⁴⁶ Cfr. GERVAISE, 2019.

¹⁴⁷ Cfr. <https://www.justit.co.uk/insight/rise-of-the-machines-data-analytics-in-2021-and-beyond/>

De facto, concordamos que a tarefa de traçar a linha que distingue a criação da obtenção dos dados poderá ser de elevada dificuldade, especialmente no caso de dispositivos conectados¹⁴⁸.

Contudo, é nosso entender que deverá ser valorizado o argumento que os dispositivos conectados coletam dados através da observação dos mesmos. Um claro exemplo desta observação poderá consistir na medição - realizada por um sensor - da densidade de tráfego ou de dados meteorológicos. Esta situação vem, por sua vez colocar dúvidas quanto à classificação destes dados como criados ou obtidos.

No seguimento deste pensamento, o Supremo Tribunal Federal Alemão debruçou-se sobre esta questão no caso *Autobahnmaut*.¹⁴⁹ Neste caso, o tribunal considerou que a empresa Toll Collect, cuja função seria a cobrança de portagens pela utilização das autoestradas, recorrendo a sensores e terminais, era detentora de um direito *Sui Generis* sobre a base de dados dinâmica utilizada para a faturação.

O tribunal afirmou que os dados registados pelos sensores e terminais não foram criados pela Toll Collect, mas sim obtidos, uma vez que estes já existiam independentemente do investimento realizado pelo fabricante da base de dados.

Por outro lado, o TJUE levantou a problemática advinda das situações em que a criação da informação coincide com a criação da base de dados. Contudo, o tribunal afastando-se dessa problemática, veio concluir que, se a criação de uma base de dados estiver ligada a uma atividade principal, onde o criador da base de dados também cria o conteúdo, a proteção não fica necessariamente excluída. No entanto, salienta que o próprio investimento deve ser feito de forma independente na obtenção, verificação ou apresentação do conteúdo, ou seja, distingue-se da elaboração do conteúdo¹⁵⁰.

No cenário de dados gerados por sensores ou máquinas, não nos parece possível identificar qualquer investimento substancial, quer na recolha, quer na verificação dos dados. Salvo, nas situações em que se considere que o investimento realizado nos próprios sensores é, efetivamente, substancial.

Assim, é nosso entender que a proteção poderá, efetivamente, ser encontrada no investimento substancial realizado na apresentação dos dados, na medida em que, estes são dados são tipicamente brutos - recolhidos de forma desorganizada- e, como tal, necessitam de ser organizados de forma metódica ou sistemática, o que por sua vez irá

¹⁴⁸ DREXL, 2018 p.71 p. 71; ELFERING, 2019.

¹⁴⁹ BGH, GRUR 2010, 1004 – Caso I ZR 47/08, 25 de março de 2010, *Autobahnmaut*.

¹⁵⁰ Cfr. *The British Horseracing; Fixtures Marketing Ltd e OPAP*.

implicar um investimento substancial no desenvolvimento do algoritmo e do próprio sistema que possibilita essa organização.

Apesar de concordarmos que, mediante condições muito específicas, um reduzido número de bases de dados decorrentes de Big Data poderão ser protegidas pelo direito Sui Generis, admitimos que a esmagadora maioria deste tipo de bases de dados ficariam excluídas do âmbito de proteção.

Pre vemos que será cada vez mais difícil satisfazer os requisitos necessários à obtenção desse direito, principalmente num contexto de economia de dados, já que os processos de obtenção, verificação e apresentação dos dados serão cada vez mais automatizados.

8. A Diretiva 96/9/CE na era dos Big Data.

Como alguns autores, e a própria Comissão Europeia, assinalam¹⁵¹ a Diretiva reveste-se de conceitos vagos que dificultam sua aplicação.

Desde logo, o conceito de investimento substancial não se encontra bem delimitado, o que torna difícil a sua aplicação, e por consequência, poderá levar a decisões completamente diferentes por parte dos tribunais na UE. Este, sendo o requisito fundamental à aplicabilidade do Direito Sui Generis, mostra-se particularmente ambíguo em contexto de Big Data. O facto de os dados serem gerados e armazenados, de forma constante, implicam custos muito elevados, podendo provocar um aumento no limiar do investimento substancial.

Por outro lado, no caso *Fixtures Marketing*, o TJUE interpretou de forma restritiva que, a proteção Sui Generis se limitava aos produtores “primários” de bases de dados. Com esta decisão excluiu do âmbito de proteção da Diretiva as bases de dados secundárias, em relação à atividade principal, por considerar que estaremos perante um investimento na “criação” e não na “obtenção” dos dados. Como vimos, existe uma exceção caso se consiga comprovar que a posterior obtenção, verificação e apresentação do conteúdo implicou um investimento substancial. Esta exceção, possibilita que um criador de uma base de dados seja, simultaneamente, o criador dos dados nela contidos, o que, por sua vez, é suscetível de dar origem a situações de monopólios económicos, que

¹⁵¹ Neste sentido CÒNSUL,2019; LEINSTER, 2018.

podem resultar num comportamento anti concorrencial e conseqüentemente bloquear o livre acesso à informação¹⁵².

Outro aspeto problemático prende-se com a duração do direito Sui Generis.

O artigo 10º, nº3 da Diretiva determina, especificamente, que o prazo inicial de 15 anos poderá ser facilmente prolongado cada vez que é feito um novo investimento substancial na base de dados. Em contexto de Big Data, esta disposição levanta muitas questões, já que, sendo as bases de dados na sua maioria automatizadas a informação utilizada por esta tecnologia está constantemente a ser atualizada o que, preenchido o requisito do investimento substancial, levará à atribuição de um direito de exclusivo perpétuo¹⁵³.

Finalmente, uma outra questão controversa, resulta da decisão do TJUE no processo C-30/14 processo Ryanair v PR Aviation¹⁵⁴.

Neste processo, o TJUE determinou que uma base de dados que satisfaça a definição constante no artigo 1º da Diretiva, mas que não possa ser protegida quer pelos Direitos de Autor, quer pelo Direito Sui Generis, afastaria a aplicação dos artigos 8º e 15º da Diretiva,- que determinam que o fabricante de uma base de dados posta à disposição do publico não pode impedir o utilizador legítimo dessa base de dados de extrair ou reutilizar partes insubstanciais do seu conteúdo. Se o fabricante de bases de dados tentar restringir esta situação através de limitações contratuais, essas limitações serão nulas - permitindo assim que o fabricante da base de dados proibisse a extração e reutilização por parte do utilizador legítimo.

O TJUE, com esta decisão, veio abrir caminho à imposição de limitações contratuais¹⁵⁵, o que se revela problemático. A possibilidade reconhecida ao proprietário de uma base de dados, não protegida pela Diretiva, de limitar contratualmente o acesso e utilização da mesma, é suscetível de deixar os criadores de bases de dados não protegidas em melhor posição do que os criadores de bases de dados protegidos pelos direitos concedidos pela Diretiva¹⁵⁶.

Consideramos, portanto, que a atual redação da Diretiva é suscetível de causar insegurança jurídica e, por consequência, não confere uma adequada proteção às bases de dados em contexto de Big Data.

¹⁵² Rethinking Database Rights and Data Ownership in an AI World p.17

¹⁵³ PILA & TORREMANS, 2019 p.491.

¹⁵⁴ Cfr. Processo C-30/14 Ryanair .

¹⁵⁵ MYSKA & HARASTA p.183, 2016.

¹⁵⁶ Neste sentido ZEITLIN, 2018 pp.62-63.

9. O Futuro

Passados 26 anos desde a entrada em vigor da Diretiva relativa à proteção das bases de dados, podemos afirmar que o panorama em que a mesma foi pensada mudou completamente. Partiu de uma era marcadamente analógica para a época do digital, onde os dados assumiram um papel essencial na inovação e no crescimento económico.

Nos dias que correm, a Diretiva está envolta em incerteza jurídica relativamente ao seu âmbito de aplicação. Segundo um estudo levado a cabo pela Comissão Europeia, com o intuito de avaliar o impacto da Diretiva¹⁵⁷, ficou claro que a mesma não acompanhou os desenvolvimentos tecnológicos que revolucionaram a economia de dados.

Nesse estudo foi dada especial atenção à crescente utilização de dados gerados por sensores e máquinas, em ambiente de IoT, sendo apresentadas evidências sobre o impacto de possíveis opções políticas para a revisão da Diretiva, especialmente no que diz respeito ao Direito Sui Generis.

Depois de muita expectativa, a Comissão Europeia publicou a sua Proposta de Regulamento Europeu sobre o acesso equitativo aos dados e a sua utilização (Data Act). Esta proposta, consiste no segundo elemento principal da estratégia europeia de dados apresentada no ano de 2020, que tem como objetivo garantir o papel de liderança da UE numa sociedade baseada em dados.

Como expectável, esta Proposta veio introduzir alterações significativas à Diretiva 96/9/CE relativa à proteção das bases de dados. Contudo, contrariamente ao que era esperado, não procedeu a uma verdadeira e necessária revisão da mesma. A proposta consagra apenas um único artigo que esclarece que o Direito Sui Generis sobre bases de dados, presente no capítulo III da mencionada Diretiva, não será aplicável a bases de dados que incluam dados obtidos ou gerados pelo uso de um dispositivo conectado¹⁵⁸.

Presente no capítulo X da lei de dados, o artigo 35º determina na sua redação que, “com o intuito de não dificultar o exercício do direito de acesso e utilização desses dados pelos utilizadores em conformidade com o artigo 4º do presente regulamento ou do direito de partilha desses dados com terceiros em conformidade com o artigo 5º do presente regulamento, o direito Sui Generis previsto no artigo 7º da Diretiva 96/9/CE não se aplica

¹⁵⁷ Comissão Europeia – “Study in support of the evaluation of Directive 96/9/EC on the legal protection of databases,” 2018.

¹⁵⁸ Cfr. DERCLAYE & HUSOVEC, 2022 p.1

às bases de dados que contenham dados obtidos ou gerados pela utilização de um produto ou serviço conectado”

Desta forma, o artigo 35º determina que quaisquer dados que se enquadrem na definição do artigo 4º da proposta de regulamento, isto é, “dados gerados pela utilização de um produto ou serviço relacionado” não serão protegidos. De forma a obter uma melhor compreensão do artigo devemos relacioná-lo com o considerando 84 da Proposta que determina que o objetivo é evitar que os proprietários de bases de dados, que não beneficiam de proteção (as obtidas ou geradas a partir de produtos conectados) reivindiquem indevidamente direitos sobre essas bases de dados¹⁵⁹.

Da interpretação do artigo 35º, verificamos que a proibição nele constante se aplica também às bases de dados que, atualmente, se qualificam para a proteção do direito Sui Generis, mas que são resultado de dados provenientes de um produto conectado.

Como já se deixou dito, a proteção Sui Generis será efetivamente aplicável a bases de dados em contexto de Big Data, – embora em condições muito específicas e sem conferir uma proteção adequada. Com a atual formulação do artigo 35º, será provocada uma enorme mudança de cenário neste contexto. Como vimos, em contexto de Big Data a maioria dos dados é gerada por máquinas, o que, implica que as bases de dados criadas neste contexto fiquem automaticamente excluídas do âmbito da proteção conferida pelo Direito Sui Generis¹⁶⁰.

Tendo em mente que as bases de dados criadas a partir de sensores ou máquinas são um mercado em crescimento exponencial¹⁶¹, consideramos que, se a proposta for aprovada na sua redação atual, terá um impacto negativo no mercado das bases de dados europeu. A proposta, ao determinar que o Direito Sui generis não se aplica às bases de dados obtidos ou gerados pela utilização de um produto ou serviço relacionado, exclui da proteção um enorme número de bases de dados. Consequentemente, provocará um desincentivo ao investimento no mercado da criação de bases de dados Europeu.

¹⁵⁹ Cfr COSTA, 2022

¹⁶⁰ Cfr. BOND & STEPHENS, 2022

¹⁶¹ CÒNSUL, 2019

Conclusão:

Após o estudo e análise efetuados, tecemos agora algumas considerações finais.

Concluimos, em primeiro lugar, que o crescente desenvolvimento no setor tecnológico tem provocado um considerável crescimento nos mercados de Big Data.

Podemos, também, afirmar que, nos dias de hoje, os Big Data se caracterizam como a figura central da economia digital.

Contudo, apesar de ser um conceito já profundamente enraizado na sociedade, não se encontra adequadamente regulado. Desde logo, apontam-se os modelos de negócio baseados na comercialização de Big Data que levantam sérios dilemas em matéria de *Antitrust* e proteção de dados.¹⁶²

A legislação existente, de facto, não foi capaz de acompanhar a evolução tecnológica. Um claro exemplo disso é a Diretiva 96/9/CE relativa à proteção das bases de dados. Esta, inicialmente introduzida com o intuito de harmonizar a legislação europeia em matéria de base de dados e estimular o crescimento da indústria de dados, deixou de servir o seu propósito, principalmente no que respeita ao incentivo à criação de base de dados.

Apesar de admitirmos que, em circunstâncias muito específicas, as bases de dados decorrentes de Big Data possam ser protegidas ao abrigo da Diretiva, consideramos que esta é suscetível de gerar insegurança jurídica.

A Diretiva foi, efetivamente, projetada e introduzida num período fortemente marcado pelo analógico e, como tal, não se mostra adequada a proteger as bases de dados decorrentes de Big Data. Desde logo, colocam-se entraves no enquadramento e na própria definição de base de dados plasmada no artigo 1º, - que, pela necessidade de uma organização sistemática e metódica, exclui a proteção conferida aos dados brutos, que, por sua vez, representam a maior parte do acervo dos Big Data¹⁶³.

Além do mais, a Diretiva está repleta de conceitos vagos, suscetíveis de diversas interpretações, como sejam: o conceito de originalidade, no que respeita aos Direitos de Autor, e o conceito de investimento substancial, no que respeita ao direito *Sui Generis*.

Relativamente à proteção conferida pelos Direitos de Autor, constatamos a sua imediata exclusão no que concerne às bases de dados decorrentes de Big Data,

¹⁶² Cfr. Caso Cambridge Analytica.

¹⁶³ Estes dados correspondem aproximadamente a 80% dos dados gerados no mundo. Cfr. <https://www.treinaweb.com.br/blog/o-que-sao-dados-estruturados-e-nao-estruturados>.

tipicamente criadas por sistemas dotados de inteligência artificial que organizam os dados atendendo a uma forma padrão. O que, por isso, exclui, tanto a originalidade, como a intervenção humana - típicas dos Direitos de Autor.

Por sua vez, o direito *Sui Generis*, foca a sua atenção no investimento substancial realizado na obtenção, verificação e apresentação dos dados. Estes conceitos têm sido alvo de várias interpretações pelo TJUE. Este veio determinar que o investimento só poderá ser realizado em dados pré-existentes, excluindo deste âmbito as bases de dados decorrentes de Big Data, que recorrem muitas vezes a dispositivos conectados que criam, recolhem e analisam os dados. Contudo, poderá ser encontrada uma lacuna que permita a proteção deste tipo de bases de dados, desde que se comprove a realização de um investimento substancial na apresentação.

Assim, consideramos que a Diretiva, na sua atual redação, não é adequada à proteção de bases de dados decorrentes de Big Data. Este facto já havia sido constatado num estudo realizado pela Comissão Europeia,¹⁶⁴ com o fim único de avaliar a Diretiva e, como tal, era esperada e desejada uma revisão da mesma.

Contudo, a Comissão, teve outros planos e na sua proposta de Regulamento sobre o acesso equitativo aos dados e a sua utilização (Data Act), removeu qualquer possibilidade de conferir proteção às bases de dados decorrentes de Big Data, ao excluir do âmbito de proteção da Diretiva qualquer base de dados constituída por dados provenientes de sensores ou dispositivos conectados.

Reconhecemos que é imprescindível encontrar um equilíbrio entre, por um lado o livre fluxo de dados, permitindo que os mesmos circulem dentro da UE e entre todos os setores, em benefício de todos, e por outro, os interesses dos produtores de bases de dados, uma vez que se trata de um mercado em expansão.

A importância dos Big Data é crescente e, como tal, devem ser realizados esforços no sentido de regular os vários aspetos desta tecnologia do futuro que, salvo melhor opinião, não deve ser desprezada.

¹⁶⁴ Comissão Europeia – “Study in support of the evaluation of Directive 96/9/EC on the legal protection of databases,” 2018

Bibliografia:¹⁶⁵

Artigos e Livros

- **ALEJANDRE, Maria**, “*La regulación europea de la protección de las bases de datos a debate: Reflexiones acerca del futuro de la Directiva 96/9/CE*”, 2010, disponível em: <http://hdl.handle.net/10486/13894>
- **AMANKWA, Joseph, KHAN, Zaheer, KNIGHT, Gary**, “*COVID-19 and digitalization: The great acceleration*”, *Journal of Business Research* vol.36, 2021, pp. 602-611 disponível em: <https://doi.org/10.1016/j.jbusres.2021.08.011>
- **ARRUDA, Vinicius**, “*Big Data, dados pessoais e propriedade intelectual*”: limites e tutela jurídica das bases de dados, 2018, disponível em: <http://tede.mackenzie.br/jspui/handle/tede/3932>
- **BAIG, Maria, SHUIB, Livana, YADEGARIDEHKORDI, Elaheh**, “*Big data in education: a state of art, limitations, and future research directions*”, *International Journal of Educational Technology in Higher Education*, 2020, disponível em: <https://educationaltechnologyjournal.springeropen.com/articles/10.1186/s41239-020-00223-0>
- **BARTODZIE, Christoph J.**, “*The Concept Industry 4.0: An Empirical Analysis of Technologies and Applications in Production Logistics*”, Springer Glaber, 2017
- **BATHA, Gourav, RANI, Rinkle, AGGARWAL, Himanshu**, “*Comparative Study of NoSQL databases for big data storage*”, *International Journal of Engineering & Technology*, 2018, pp. 2- 6, disponível em: https://www.researchgate.net/publication/323726150_Comparative_study_of_NoSQL_databases_for_big_data_storage

¹⁶⁵ A bibliografia que consta desta lista foi citada ao longo do texto, em nota de rodapé segundo o estilo autor-data-página. As obras disponíveis online foram consultadas entre outubro de 2021 e março de 2022.

- **BIDARRA, Ana**, “*O impacto do Big Data no Direito da Concorrência: O caso Facebook/WhatsApp e o Controlo de Concentrações na União Europeia*”, 2018
- **BODEN, Margaret A.**, “*AI: a very short introduction*”, Oxford University Press, 2018;
- **BOUTIN, Xavier, CLEMENS, Georg**, “*Defining 'Big Data' in Antitrust*”, Competition Policy International: Antitrust Chronicle, Volume 1, Number 2, 2017, pp.22-28, disponível em: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2938397
- **BURDESE, Paolo** “*AI-Generated Databases. Do the Creation/obtaining Dichotomy and the Substantial Investment Requirement Exclude the Sui Generis Right Provided for under the EU Database Directive? Reflections and Proposals*”, 2020 disponível em: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3850662
- **CANNATA, Joe, FALCE, Valeria, POLLICINO, Orest**, “*Introduction – new legal challenges of Big Data*”, New Legal Challenges of Big Data, Edward Elgar Publishing, 2020, pp. 1 – 14
- **CÒNSUL, ROMEU I**, “*Capítulo VII: Las nuevas bases de datos. Big data, desestructuración e inteligencia artificial*”, Nuevos desafíos para el Derecho de Autor: robótica, inteligencia artificial, tecnología, Reus Editora, 2019
- **CORREIA, José**, “*Processamento Analítico de Dados em Contexto de Big Data com o Druid*”, 2018, disponível em: <http://hdl.handle.net/1822/59393>
- **COSTA, Inês**, “*A proteção da pessoa na era dos big data: a opacidade do algoritmo e as decisões automatizadas*”, Revista Eletrónica de Direito, 2021, disponível em: <https://cije.up.pt/pt/red/edicoes-anteriores/2021-nordm-1/a-protecao-da-pessoa-na-era-dos-big-data-a-opacidade-do-algoritmo-e-as-decisoes-automatizadas/>
- **CRUZ, Tiago**, “*Data analytics para variedade de dados*”, 2017, disponível em: <http://hdl.handle.net/1822/54536>
- **CUKIER, Keneth, MAYER-Shcoenberger**, “*The rise of big data: How It’s Changing the Way We Think About The World*”, Foreign Affairs Vol. 92, No.3, 2013 pp. 28 - 40

- **DERCLAYE, Estelle**, “*The Legal Protection of Databases: A Comparative Analysis*”, Edward Elgar Publishing, 2008
disponível em: <http://hdl.handle.net/10451/37518>
- **DREXL, Josef**, “*Data Access and Control in The Era of Connected Devices: Study on Behalf of the European Consumer Organisation BEUC*”, 2018, disponível em: https://www.beuc.eu/publications/beuc-x2018121_data_access_and_control_in_the_area_of_connected_devices.pdf
- **DUROVIC, Mateja, LECH, Franciszek**, “*Global perspectives on big data and consumer law*”, New Legal Challenges of Big Data, Edward Elgar Publishing, 2020, pp.154 – 172
- **ELFERING, Stephanie**, “*Unlocking the Right to Data Portability: An Analysis of the interface with the Sui Generis Database Right*,” 2019, disponível em: https://www.researchgate.net/publication/336350859_Unlocking_the_Right_to_Data_Portability_An_Analysis_of_the_Interface_with_the_Sui_Generis_Database_Right
- **GERVAIS, Daniel J.**, “*Exploring the Interfaces Between Big Data and Intellectual Property Law*”, *Journal of Intellectual Property, Information Technology and Electronic Commerce Law*, 2019, disponível em: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3360344
- **Hong, Liang et al**, “*Big Data in Health Care: Applications and Challenges, Data and Information Management*”, 2019, disponível em: https://www.researchgate.net/publication/330724271_Big_Data_in_Health_Care_Applications_and_Challenges
- **LEAL, Ana**, “*Big Data nos setores Bancário e Financeiro*”, *Inteligência Artificial & Direito*, Almedina, 2020, pp. 199 – 213
- **LEE, LINK HANG et al.** “*All One Needs to Know about Metaverse: A Complete Survey on Technological Singularity, Virtual Ecosystem, and Research Agenda*,” 2021 disponível em: <https://arxiv.org/abs/2110.05352>
- **LEINSTER, Matthias**, “*Big Data and the EU Database Directive 96/9/EC: Current Law and Potential for Reform*”, 2018, disponível em: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3245937B

- **MADAKAM, Somayya, TRIPATHI, Siddhart**, “*Internet of Things (IoT): A Literature Review*”, Journal of Computer and Comunnications, 2015, pp. 164-173 disponível em: https://www.researchgate.net/publication/280527542_Internet_of_Things_Io_T_A_Literature_Review
- **MARQUES, Remédio**, “*Bioteχνologias e Propriedade Intelectual: obtenções vegetais, conhecimentos tradicionais, sinais distintivos, bioinformática e bases de dados, direito da concorrência*”, Edições Almedina, 2007
- **MASSENO, Manuel David**, “*Como a União Europeia procura proteger os cidadãos-consumidores em tempos de Big Data*”, Revista Eletrônica do Curso de Direito da UFSM vol. 14, n. 3, 2019, disponível em: <https://doi.org/10.5902/1981369441708>
- **MASSENO, Manuel David**, “*Nas fronteiras da PI: os direitos patrimoniais sobre dados, uma perspectiva Europeia*”, Revista Rede de Direito Digital, Intelectual & Sociedade, 2021 pp. 101 – 114 disponível em https://www.gedai.com.br/wp-content/uploads/2021/05/Revista-RRDDIS_v-1-n-1-2021.pdf
- **MCAFEE, Andrew, BRYNJOLFSSON, ERIK** “*Big Data: The Management Revolution*”, Harvard Business Review, 2012, disponível em <https://hbr.org/2012/10/big-data-the-management-revolution>.
- **NOYES, Katherine**, “How big data is changing the database landscape for good”, 2015, disponível em: <https://www.computerworld.com/article/3003427/how-big-data-is-changing-the-database-landscape-for-good.html>
- **OWAIS, Suhail, HUSSEI, Nada**, “*Extract Five Categories CPIVW from the 9V's Characteristics of the Big Data*”, International Journal of Advanced Computer Science Applications, Vol. 7, No. 3, 2016, pp. 254 - 258 disponível em: https://thesai.org/Downloads/Volume7No3/Paper_37Extract_Five_Categories_CPIVW.pdf
- **PALKA, Prezemyslaw, Lippi, Marco**, “*Big data analytics, online terms of service and privacy policies*”, Research Handbook on Big data Law, Edward Elgar Publishing, 2021, pp. 115 – 134

- **PILA, Justin, TORREMANS, Paul**, “*European Intellectual Property Law*”, Oxford University Press, 2^a ed, 2019
- **PRAVEEN, Shagufta, CHANDRA, Umesh**, “*Influence of structured, Semi-Structured, Unstructured data on various models*”, International Journal of Scientific and Engineering Research, 2020, pp. 67 – 69 disponível em: https://www.researchgate.net/publication/344363081_Influence_of_Structured_Semi-Structured_Unstructured_data_on_various_data_models
- **RASHID, Salman, RAZAK Shukor**, “*Big data Challenges in 5G Networks*”, Eleventh International Conference on Ubiquitous and Future Networks, 2019 disponível em: https://www.researchgate.net/publication/335351537_Big_Data_Challenges_in_5G_Networks
- **ROCHA, Maria Victória**, “*Contributos para a delimitação da originalidade como requisito da proteção da obra pelo Direito de Autor*”, Verbo Jurídico, 2003, disponível em <https://www.verbojuridico.net/doutrina/autor/originalidade.html>
- **RODRIGUES, Luís** “*Arquitecturas dos Sistemas de Informação*”, 2000, disponível em <https://recipp.ipp.pt/bitstream/10400.22/4930/1/Disserta%C3%A7%C3%A3o.PDF>
- **RUBINFELD, Daniel L., GAL, Michal**. “*Access Barriers to Big Data*”, Arizona Law Review 339, 2017, disponível em: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2830586
- **SHARMA, Neha, SHAMKUWAR, SINGH, Inderjit**, “*The History, Present and Future with IoT, Internet of Things and Big Data Analytics for smart Generation*”, 2018, pp. 27–51, disponível em: https://www.researchgate.net/publication/330023744_The_History_Present_and_Future_with_IoT
- **SILVA, NUNO SOUSA**, “*Subsídios para a transposição da Diretiva 2019/790*”, Revista de Direito Intelectual, 2020 pp. 245-272 disponível em: https://www.nsousaesilva.pt/web/images/_Data/Publicacoes-Artigos/NSS_Subsidi%C3%B3dios_para_a_transposic%C3%A7%C3%A3o_da_Diretiva_2019790.pdf

- **SILVA, Nuno Sousa**, “Uma introdução ao Direito de Autor Europeu”, Revista da Ordem dos Advogados, 2013, pp 1331 – 1387, disponível em: <https://portal.oa.pt/upl/%7Bdca5e510-7e6b-403c-9e2c-fd091a1cb5dc%7D.pdf>
- **SLOOT, Bart, SCHENDEL**, “*Ten questions for Future Regulation of Big Data: A comparative and Empirical Legal Study*”, 2016 , disponível em: <http://www.jipitec.eu/issues/jipitec-7-2-2016/4438>
- **SOLDATOV, Oleg**, “*The relationship between freedom of expression and big data*”, New Legal Challenges of Big Data, Edward Elgar Publishing, 2020, pp.232- 256
- **SURDEN, Harry**, “Machine Learning and Law”, 2014, disponível em: <https://scholar.law.colorado.edu/cgi/viewcontent.cgi?article=1088&context=articles>
- **SYKUTA, MICHAEL E.**, “Big data in Agriculture: Property Rights, Privacy and Competition in AG Data Services,” *International Food and Agriculture Management Review Special Issue – Vol. 19 Issue A*, 2016, pp. 58 – 73 disponível em: <https://ageconsearch.umn.edu/record/240696/>
- **VENÂNCIO, Pedro**, “*Notas sobre o regime do Direito Especial do Fabricante de Bases de Dados*”, IV Congresso Internacional de Ciências Jurídico-Empresariais, 2014, pp.22-37, disponível em: <http://hdl.handle.net/10400.8/1449>
- **VENÂNCIO, Pedro, Dias** “*Das compilações às bases de dados enquanto objetos de propriedade intelectual*”, *SCIENTIA IURIDICA*, tomo LXI, 330, 2012, pp 597 – 628
- **WINGERATH, et al**, “*The case For Change Notifications in Pull-Based Databases*”, 2017, disponível em: <https://dl.gi.de/bitstream/handle/20.500.12116/923/paper32.pdf?sequence=1&isAllowed=y>
- **ZEITLIN, Martin**, “*Everything Counts in Large Amounts: Protection of big data under the Database Directive*”, 2018, disponível em: <https://uu.diva-portal.org/smash/get/diva2:1211909/FULLTEXT01.pdf>

- ZHANG, Yanqing, CHEN, Chengshang, “*Business model innovation on big data: a case study on viewing big data as a resource.*”, 2014, disponível em: <https://www.diva-portal.org/smash/get/diva2:822154/FULLTEXT01.pdf>

Legislação:

Nacional

Disponível em: <https://dre.pt/dre/home>

- Código do Direito de Autor e dos Direitos Conexos;
- Código da Propriedade Industrial;
- Decreto-Lei n.º 122/2000, de 4 de julho, que transpõe para a ordem jurídica interna 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;
- Decreto-Lei n.º 252/94, de 20 de outubro, que transpõe para a ordem jurídica interna a Directiva n.º 91/250/CEE, do Conselho, de 14 de maio, relativa ao regime de proteção jurídica dos programas de computador;

União Europeia:

Disponível em: <https://eur-lex.europa.eu/homepage.html>

- Diretiva 96/9/CE do Parlamento Europeu e do Conselho, de 11 de março de 1996, relativa à proteção jurídica das bases de dados;
- Diretiva (UE) 2019/790 do Parlamento Europeu e do Conselho, de 17 de abril de 2019, relativa aos direitos de autor e direitos conexos no mercado único digital e que altera as Diretivas 96/9/CE e 2001/29/CE;
- Diretiva 2009/24/CE do Parlamento Europeu e do Conselho, de 23 de abril de 2009, relativa à proteção jurídica dos programas de computador;

Jurisprudência

TRIBUNAL DE JUSTIÇA

Disponível em: <https://eur-lex.europa.eu/collection/eu-law/eu-case-law.html>

- Acórdão do Tribunal de Justiça de 1 de dezembro de 2011, Painer, C-145/10, ECLI:EU:C:2011:798;
- Acórdão do Tribunal de Justiça de 1 de março de 2012, Football Dataco, C-604/10, ECLI:EU:C:2012:115;
- Acórdão do Tribunal de Justiça de 12 de setembro de 2019, Cofemel, C-683/17, ECLI:EU:C:2019:721;
- Acórdão do Tribunal de Justiça de 15 de janeiro de 2021, Ryanair, C-30/14, ECLI:EU:C:2015:10;
- Acórdão do Tribunal de Justiça de 3 de junho de 2021, Online Latvia, C-762/19, ECLI:EU:C:2021:434;
- Acórdão do Tribunal de Justiça de 7 de agosto de 2018, Renckhoff, C-161/17, ECLI:EU:C:2018:634;
- Acórdão do Tribunal de Justiça de 8 de junho de 2004, Fixtures Marketing, processo C-338/02, ECLI:EU:C:2004:338;
- Acórdão do Tribunal de Justiça de 9 de novembro de 2004, Fixtures Marketing, C444/02, ECLI:EU:C:2004:697;
- Acórdão do Tribunal de Justiça de 9 de novembro de 2004, Fixtures Marketing, C-46/02, ECLI:EU:C:2004:694;
- Acórdão do Tribunal de Justiça de 9 de novembro de 2004, The British Horseracing, C-203/02, ECLI:EU:C:2004:695;

- Conclusão da Advogada-Geral Stix-Hackl 8 de junho de 2004, processo C-46/02, ECLI:EU:C:2004:332;

ALEMANHA

Disponível em: <https://dejure.org/dienste/rechtsprechung?gericht=BGH>

- BGH, GRUR 2010, 1004 – Caso I ZR 47/08, 25 de março de 2010, Autobahnmaut
- BGH Processo ZR196/08 – Zweite Zahnarztmeinung, 1 de dezembro de 2010, Zahnarztmeinung

Documentos de Organizações internacionais e Institutos Nacionais

- Bundeskartellamt e L’Autorité de la Concurrence, “Competition Law and Data”, 2016, disponível em: https://www.bundeskartellamt.de/SharedDocs/Publikation/DE/Berichte/Bi-g%20Data%20Papier.pdf?__blob=publicationFile&v=
- OCDE. Big data: “Bringing competition policy to the digital era”, 2016. disponível em: [https://one.oecd.org/document/DAF/COMP\(2016\)14/en/pdf](https://one.oecd.org/document/DAF/COMP(2016)14/en/pdf).
- Comunicação da Comissão ao Parlamento Europeu, ao Conselho, ao Comité Económico e Social Europeu e ao Comité das Regiões – Uma estratégia europeia para os dados COM(2020) 66 final Bruxelas, 2020, disponível em <https://eur-lex.europa.eu/legal-content/PT/TXT/PDF/?uri=CELEX:52020DC0066&from=EN>
- Comissão Europeia– “The EU data protection reform and big data.” disponível em: <https://publications.europa.eu/en/publication-detail/-/publication/51fc3ba6-e601-11e7-9749-01aa75ed71a1/language-en/format-PDF>
- Comissão Europeia – “Study in support of the evaluation of Directive 96/9/EC on the legal protection of databases,” 2018, disponível em: <https://op.europa.eu/en/publication-detail/-/publication/5e9c7a51-597c-11e8-ab41-01aa75ed71a1>

- Singapore Academy of Law - Rethinking Database Rights and Data Ownership in an AI World, 2020 disponível em: https://www.sal.org.sg/sites/default/files/SAL-LawReform-Pdf/2020-09/2020%20Rethinking%20Database%20Rights%20and%20Data%20Ownership%20in%20an%20AI%20World_ebook_0_1.pdf

ARTIGOS DE JORNAL E DE OPINIÃO

- “Big Data Analytics o que é e porque é importante” – Disponível em: <https://www.jasminsoftware.pt/blog/big-data-analytics-importante/>
- “Media usage in an internet minute as of August 2021”, Statista, 2022 disponível em: <https://www.statista.com/statistics/195140/new-user-generated-content-uploaded-by-users-per-minute/>
- “The world’s most valuable resource is no longer oil, but data”, The economist 2016, Disponível em: <https://www.economist.com/leaders/2017/05/06/the-worlds-most-valuable-resource-is-no-longer-oil-but-data>
- “What is Big Data?”, Oracle, Disponível em: <https://www.oracle.com/big-data/what-is-big-data/>
- BOND, Toby, STEPHENS, Katharine “Why IP lawyers need to pay attention to the EU’s draft Data Act”, 2022 disponível em: <https://www.twobirds.com/en/insights/2022/uk/why-ip-lawyers-need-to-pay-attention-to-the-eus-draft-data-act>
- COSTA, Inês, “Proposta de Regulamento Europeu sobre o acesso equitativo aos dados e a sua utilização (Data Act): um grande passo na transformação da economia dos dados?” 2022, disponível em: <https://adcdadvogados.pt/proposta-de-regulamento-europeu-sobre-o-acesso-equitativo-aos-dados-e-a-sua-utilizacao-data-act-um-grande-passo-na-transformacao-da-economia-dos-dados/?fbclid=IwAR3rbJME37bprpKKQgmkjxMr8LNMq-fnGYhAJ6Ap3-rXoozH6yCzDkJNXcQ>
- EU DATABASE DIRECTIVE, disponível em: https://itlaw.fandom.com/wiki/EU_Database_Directive?fbclid=IwAR38W7Y6adwUAmvZ_XspE7plP2i1s9YLxWpFbpJyOsPqWzPVjP-MYD7--r8

- GUEDES, Marrylene, “O que são dados estruturados e não estruturados?”, disponível em: <https://www.treinaweb.com.br/blog/o-que-sao-dados-estruturados-e-nao-estruturados>
- HOPPEN, Joni, SANTOS, Marcos: disponível em: <https://www.aquare.la/analise-descritiva-preditiva-prescritiva-e-cenarizacao/>
- How Much Data Is Created Every Day? [27 Staggering Stats], 2021, disponível em: https://seedscientific.com/how-much-data-is-created-every-day/?fbclid=IwAR1BXNr4CLdyzTrgoh1YjnVGzVI7OrWnSKiT7L3e4TA5-2scX3BNq_EtdxY
- Key Internet Statistics to Know in 2022 (Including Mobile) Disponível em: <https://www.broadbandsearch.net/blog/internet-statistics>
- LaBerge, et al, How COVID-19 has pushed companies over the technology tipping point—and transformed business forever, Mckinsey &Company, 2020, disponível em: <https://www.mckinsey.com/business-functions/strategy-and-corporate-finance/our-insights/how-covid-19-has-pushed-companies-over-the-technology-tipping-point-and-transformed-business-forever>
- PADINHA, Bruno “O Dilúvio de Dados - Bruno Mateus Padinha”, Jornal Económico, 2020 disponível em: <https://jornaleconomico.pt/noticias/o-diluvio-de-dados-547772>
- Rise of the Machines: Data Analytics in 2021 and Beyond, disponível em: <https://www.justit.co.uk/insight/rise-of-the-machines-data-analytics-in-2021-and-beyond/>
- The EU Data Act: “Another Important Step in the EU Data Policy” <https://www.european-lotteries.org/news/eu-data-act-another-important-step-eu-data-policy>