Atlantic Dialogues on Technology and Regulation

2025

Editors

Henrique Sousa Antunes
Luca Belli
Filipa Urbano Calvão
Yasmin Curzi
Walter Britto Gaspar
Eduardo Magrani
Filipe Medon

UCP PRESS

Organization

Sponsors

Title Atlantic Dialogues on Technology and Regulation | 2025
Editors Henrique Sousa Antunes, Luca Belli, Filipa Urbano Calvão, Yasmin Curzi, Walter Britto Gaspar, Eduardo Magrani, Filipe Medon
Authors A. Betâmio de Almeida, Beatriz Costa, Bianca Kremer Daniel Dore Lage, Fernanda Carvalho Dias de Oliveira Silva Fernando Naegele, Inês Neves, Luca Belli Nicole de Barros Moreira Reis, Rodrigo Ardissom de Souza Rodrigo Gomes, Sílvia de Carvalho Homem Sofia Chang Nogueira, Walter Britto Gaspar
Collection Colóquios

Proofreading Francisco Silva Pereira, Patrícia Feio
Graphic Design Magda M. Coelho
Cover Ana Luísa Bolsa | 4 ELEMENTOS
Date abril 2026
eISBN 9789725411926
DOI https://doi.org/10.34632/9789725411926

Universidade Católica Editora,
Sociedade Unipessoal, Lda.
Palma de Cima 1649-023 Lisboa
Tel. (351) 217 214 020
uceditora@ucp.pt | www.uceditora.ucp.pt

ATLANTIC DIALOGUES on Technology and Regulation | 2025

Editors
Henrique Sousa Antunes
Luca Belli
Filipa Urbano Calvão
Yasmin Curzi
Walter Britto Gaspar
Eduardo Magrani
Filipe Medon

UCP PRESS

Nota Prévia

Com satisfação, damos a conhecer os artigos selecionados para apresentações na 1.ª edição dos «Diálogos Atlânticos sobre Tecnologia e Regulação». Esta iniciativa foi concebida como um evento público anual de debate resultante de uma parceria entre a Faculdade de Direito da Universidade Católica Portuguesa (UCP), através do *Católica Research Centre for the Future of Law*, e a Escola de Direito da Fundação Getulio Vargas (FGV), Rio de Janeiro, através do *Center for Technology and Society*.

Os «Diálogos Atlânticos sobre Tecnologia e Regulação» propõem-se analisar diferentes expressões da evolução tecnológica global, os seus impactos em diferentes sistemas regulatórios nacionais e internacionais, e as potenciais sinergias entre tais sistemas. O evento tem um formato multissetorial, promovendo análises pluridisciplinares de alto nível com base na interação entre representantes da academia, do setor empresarial, do setor público e da sociedade civil.

O primeiro debate teve lugar em Lisboa, nos dias 3 e 4 de abril de 2025, e foi dedicado a quatro temas essenciais para o futuro da tecnologia e da regulação: a Inteligência Artificial, a Proteção de Dados, a Cibersegurança e a Regulação de Plataformas Digitais. Antecedeu a realização do encontro uma chamada de artigos sobre os temas convocados para os «Diálogos».

Esperando que a leitura seja proveitosa para a comunidade interessada, publicam-se os textos escolhidos. A este propósito, cabe agradecer ao Instituto de Conhecimento da Abreu Advogados, na pessoa do Professor Luís Barreto Xavier, seu Presidente, o financiamento da publicação. O próximo volume reunirá os artigos selecionados para a segunda edição (Rio de Janeiro, 28 a 30 de abril de 2026).

Henrique Sousa Antunes, Luca Belli,
Filipa Urbano Calvão, Yasmin Curzi,
Walter Britto Gaspar, Eduardo Magrani,
Filipe Medon

We are pleased to bring together in this volume the articles selected for presentation at the 1st edition of the "Atlantic Dialogues on Technology and Regulation." This initiative was conceived as an annual public debate event resulting from a partnership between the Faculty of Law of the Catholic University of Portugal (UCP), through the Católica Research Centre for the Future of Law, and the Law School of the Getulio Vargas Foundation (FGV), Rio de Janeiro, through the Center for Technology and Society.

The "Atlantic Dialogues on Technology and Regulation" propose to examine different expressions of global technological evolution, their impacts on different national and international regulatory systems, and the potential synergies between such systems. The event has a multisectoral format, promoting high-level multidisciplinary analyses based on the interaction between representatives of academia, the business sector, the public sector and civil society.

The first debate took place in Lisbon, on April 3 and 4, 2025, and was dedicated to four essential topics for the future of technology and regulation: Artificial Intelligence, Data Protection, Cybersecurity and Regulation of Digital Platforms. A call for papers on the topics of the "Dialogues" preceded the meeting.

Hoping that the reading will be useful for the interested community, the chosen texts are published. In this regard, we are grateful to the Knowledge Institute of Abreu Advogados in the person of Professor Luís Barreto Xavier, its President, for financing the publication. The next volume will bring together the articles selected for the second edition (Rio de Janeiro, April 28 to 30 2026).

Henrique Sousa Antunes, Luca Belli,
Filipa Urbano Calvão, Yasmin Curzi,
Walter Britto Gaspar, Eduardo Magrani,
Filipe Medon

Regular a Inteligência Artificial: da necessidade à dificuldade

A. BETÂMIO DE ALMEIDA*

Resumo: Após um breve enquadramento da Inteligência Artificial (IA), sublinhando as respectivas características e dilemas principais, a comunicação refere a necessidade e a dificuldade na regulação das suas aplicações. O autor salienta a importância da IA no futuro da sociedade humana e refere algumas potenciais consequências sociais e individuais da mesma. A Ética e o Direito têm uma função relevante na estruturação dos limites a ter em conta na relação com a IA e na fundamentação de medidas de regulação ou regulamentação na legislação que venha a ser considerada adequada.

Os desenvolvimentos da IA são muito rápidos e impõem um acompanhamento ético e uma preparação que permita aproveitar as capacidades úteis da mesma sem pôr em causa as características e direitos essenciais bem como o percurso futuro da Humanidade. A presente comunicação pretende ser um contributo para a reflexão e a discussão sobre a IA tendo em vista proporcionar um desenvolvimento com equilíbrio responsável e justo.

1. Enquadramento introdutório

A Inteligência Artificial (IA) e os respectivos sistemas digitais de suporte constituem actualmente um modo desafiante do processo tecnológico histórico impulsionado pelo ser humano. Tal como as tecnologias anteriores e outras novas tecnologias, a IA mostra-se propiciadora de benefícios notáveis, inigualáveis, mas também geradora de efeitos potenciais preocupantes para a sociedade.

* Professor Catedrático Emérito (Universidade de Lisboa/IST, Portugal) e membro do Grupo Direito e Inteligência Artificial da Faculdade de Direito (U. Católica de Lisboa).

Por opção do autor, o presente artigo não segue o novo acordo ortográfico.

A antropologia mostra a evolução do ser humano e as sucessivas etapas de reforço das suas capacidades em lidar com o mundo para além do corpo biológico. O braço, a mão e os instrumentos podem ser considerados prolongamentos que proporcionaram ganhos fundamentais. Instrumentos e técnicas desenvolveram-se para defesas essenciais à vida, à sobrevivência, e à ampliação e controlo de forças e energias para uso intensivo na exploração e transformação de recursos naturais e realizações que marcaram o poder do ser humano no nosso planeta (e.g. em A. Leroi-Gourhan,1943-1973[1], e L. Mumford, 1967-1970[2]). Esse poder foi também uma consequência do desenvolvimento do cérebro[3] e das manifestações do que se designa como inteligência humana. A inteligência como o centro humano do conhecimento, da racionalidade e da decisão conducentes aos progressos científicos e técnicos e à civilização associada.

Em resultado dos avanços em áreas de computação, comunicação, automação e controlo, o interesse pela reprodução e aplicação de funções cognitivas semelhantes às humanas desenvolveu-se desde meados do século xx. Em 1956, na conferência de Darmouth (E.U.A.), a designação "inteligência artificial" foi cunhada e desde então adoptada e utilizada não obstante não ser, para alguns especialistas, o termo correcto. Mas o termo vingou até ao presente e influencia a apreciação e a identificação que as pessoas fazem desta área tecnológica. É actualmente uma área que fascina jovens estudantes, empresários e a comunicação social sendo incentivada pelas forças políticas dominantes e, naturalmente, muito publicitada ou divulgada pela comunidade técnica associada à IA[4].

Em 2024, a IA e os sistemas digitais associados constituem uma histórica e vibrante componente da tecnologia com avanços e vantagens

---

[1] A. Leroi–Gourhan, "Évolution et Techniques", Vol. I (L´homme et la matière, 1943/1971) e Vol. II (Milieu et techniques, 1945/1973), Ed. Albin Michel.

[2] Lewis Mumford, "The Myth of the Machine", Vol. I (Technics and Human Development, 1967) e Vol. II (The Pentagon of Power, 1970), Ed. H. Bruce Jovanovich.

[3] Miguel Nicolelis, "O Verdadeiro Criador de Tudo. Como o cérebro humano moldou o Universo tal como o conhecemos" (2021), Ed. Elsinore.

[4] Exemplos em Portugal: Arlindo Oliveira, "Mentes Digitais. A ciência redefinindo a Humanidade", 2017, Ed. IST; e "Inteligência Artificial", 2019, Ed. Fundação Francisco Manuel dos Santos.

em vários sectores. Às aplicações na área da saúde (diagnósticos, prevenção e tratamentos) juntam-se outros domínios proporcionando melhorias de eficiência, mais rapidez, menos esforço na obtenção de resultados e a capacidade de criar "valor" económico no mercado. Com o uso intensivo da IA anunciam-se aumentos significativos de produtividade[5] (do PIB) para os países, o que entusiasma naturalmente as comunidades empresarial e política. A possibilidade da IA para melhorar a gestão de bens públicos é também muito considerada (e.g. os sistemas denominados de "cidades inteligentes").

A IA é também uma ideia que é reconhecida pela capacidade de realizar actos anteriormente só associados a mentes humanas. Os dispositivos, os algoritmos da IA não se distinguem por um modo estético imediato, através de objectos técnicos com uma forma física, corpórea, marcante[6]. Sem um sinal exterior das suas capacidades ou funcionalidades específicas. Na realidade, os dispositivos de IA têm como base conjuntos de funções e relações matemáticas, num suporte digital, que para funcionarem bem necessitam de uma grande quantidade de dados e de uma "aprendizagem" intensiva que pode ser muito consumidora de energia.

Fortemente associada a poderosas empresas tecnológicas, no mercado mundial, e numa agressiva competição geoestratégica, a IA e outras novas tecnologias constituem no presente um foco de atracção potenciador de inovações e transformações anunciadas como decisivas para o futuro da Humanidade. Os "robots" e os sistemas autónomos inteligentes, nomeadamente viaturas, constituem temas que têm popularizado a IA e suscitam análises sobre as suas características e desafios. Salienta-se o entusiasmo extraordinário pela IA generativa, nomeadamente do ChatGPT desde 2022, com capacidade para desafiar o conceito de autor como o conhecemos.

A importância do uso de técnicas, agora tecnologias, na evolução da Humanidade é reconhecida e, não estando em causa a existência e os desenvolvimentos da IA, considera-se que o controlo do respectivo uso é

---

[5] O que pode não se traduzir em aumentos dos rendimentos médios dos cidadãos dos países.

[6] O modo frequente de representar um algoritmo de IA são esquemas simbólicos de arquitecturas de redes neurais, mas sem uma forma externa visível.

necessário. Análises, recomendações e orientações éticas[7] têm sido apresentadas institucionalmente[8]. Formas de regulação têm sido ensaiadas e discutidas tendo em vista uma aplicação eticamente responsável da IA. O AI Act da União Europeia, aprovado em 2024, constitui o exemplo internacional mais conhecido de uma regulamentação estruturada das aplicações da IA e é uma referência a nível internacional.

Uma proposta de Código de Conduta para Organizações que Desenvolvem Sistemas Avançados de IA – Processo Internacional de Hiroshima apareceu numa cimeira do G7 em 2023. O Papa Francisco apresentou em 2024, em discurso realizado no Fórum do G7 em Borgo Egnazia (Itália), os aspectos fundamentais da relação entre a IA e a sociedade humana. O tópico mereceu também uma atenção especial no documento final da reunião do G20 no Brasil (2024) – "Declaração de Líderes do Rio de Janeiro" – salientando a ética na defesa do uso seguro e confiável da IA (pontos 77, 78 e 79 da declaração).

Síntese

Não obstante o reconhecimento dos benefícios, afigura-se ser importante sublinhar a necessidade da regulação da IA, com reflexos necessários e desafios novos na aplicação da Lei[9]. Esta regulação deve ser adaptável à evolução rápida da IA. Em nosso entender, não é pertinente alimentar medos existenciais inspirados por alguma ficção. O que importa é o comportamento humano perante o desenvolvimento tecnológico. Não é dispiciente esta afirmação de Jacques Lafitte em 1932[10]: "Na máquina, em todo o progresso humano, o que é necessário temer somos nós e os nossos abandonos." Neste caso, é "nós" não temermos alguns potenciais

---

[7] Luciano Floridi, "The Ethics of Artificial Intelligence, Principles, Challenges and Opportunities", Ed. Oxford; Mark Coeckelbergh, "AI Ethics", 2020, Ed. MIT Press Essential Knowledge series.

[8] Um exemplo em Portugal: o Livro Branco "Inteligência Artificial (IA). Inquietações Sociais, Propostas Éticas e Orientações Políticas" do Conselho Nacional de Ética para as Ciências da Vida (2024).

[9] H. Sousa Antunes et al. (eds.), "*Multidisciplinary Perspectives on Artificial Intelligence and the Law*", 2024, Ed. Springer-Verlag, Law, Governance and Technology, Series 58.

[10] Jacques Lafitte, "Réflexions sur la Science des Machines" (1932), Ed. Librairie Philosophique J. VRIN (1972).

efeitos a curto e a longo prazos da IA. A afirmação de Lafitte reforça o princípio de não submissão a um determinismo tecnológico.

Tal como tem acontecido ao longo da história das técnicas, a regulamentação e a regulação da IA suscitam dúvidas, resistências e oposições para além da respectiva complexidade objectiva. São dificuldades a ter em conta e que devem ser discutidas e gradualmente superadas. Tendo em conta o modo de desenvolvimento e as aplicações actuais da IA, um cenário limite pode ser considerado: o domínio absoluto de um conjunto de empresas tecnológicas na estruturação da sociedade mundial, impondo os seus produtos e regras, e com a tendência para uma submissão mais ou menos acelerada, mais ou menos completa, aos efeitos dessa imposição.

2. Necessidade de regulação

A Humanidade muda sempre, muito ou pouco, cada vez que mudam instrumentos e sistemas técnicos ou sistemas institucionais. As aplicações em curso e os desenvolvimentos previsíveis da tecnologia digital e de IA configuram uma alteração muito profunda da sociedade humana. É frequente a afirmação que muitos problemas associados à IA são decorrentes de comportamentos humanos antigos. Recorrendo à citação de Lafitte, tal pode ser parcialmente correcto, mas as novas capacidades e a escala (no tempo e no espaço) de propagação dos efeitos constituem uma nova realidade. Convém assinalar que a análise dos impactos da IA mobiliza preocupações diversificadas abrangendo diferentes perspectivas, incluindo as das ciências sociais e humanas. Alguns dos efeitos com relevância ética têm sido identificados[11] e são já sentidos.

Salientam-se, numa dimensão individual:

– As transgressões éticas ou ameaças a direitos fundamentais dos cidadãos e das instituições, como os da privacidade, autonomia, segurança e vulnerabilidade. A recolha, gestão e uso direccionado de dados pessoais (e.g. na orientação de decisões pessoais) é um dos aspectos importantes e marcantes dos sistemas de preparação e gestão de dados

---

[11] M. C. Patrão Neves e A. Betâmio de Almeida, Before and Beyond Artificial Intelligence: Opportunities and Challenges, em "*Multidisciplinary Perspectives on Artificial Intelligence and the Law*", 2024, Ed. Springer-Verlag Law, Governance and Technology, Series 58,

pela IA. O regulamento de protecção de dados em vigor na União Europeia é um contributo na protecção dos cidadãos.

Na dimensão social ou colectiva há outros efeitos potenciais:

– Ultrapassados que foram os limites humanos de força e energia, estas novas tecnologias desafiam as capacidades intelectuais humanas de suporte à decisão e à criação, nomeadamente a artística e a científica. De um modo autónomo ou como apoio intensivo, colaborativo (e.g. na escrita, na tradução e edição de textos), os sistemas de IA podem propiciar a perda de algumas qualidades mentais e conduzir a uma desqualificação intelectual progressiva dos humanos, nomeadamente no domínio da criatividade.

– A externalização da memória e do raciocínio em sistemas algorítmicos tenderá a afectar nos humanos as capacidades próprias de conhecimento e de solução de problemas complexos. Os conhecimentos teóricos podem tornar-se operacionalmente desnecessários (o fim de teorias) e substituídos por uma dependência dos algoritmos IA. No entanto, um modo de utilização adequado da IA poderá também conduzir a um aproveitamento equilibrado das suas capacidades na obtenção de benefícios. Este aspecto é importante a ter em conta nos diferentes graus de ensino e no modo de encarar a IA.

– O incremento da "aceleração do tempo" e as alterações relacionais tendo como modelo as características operacionais da IA, com o consequente impacto psicológico e na individuação dos humanos. A implementação de procedimentos automáticos mais eficientes e rápidos, dispensando progressivamente a intervenção humana (considerada prejudicial ou ineficiente), faz com que os sistemas de IA tendam a impor alterações nos relacionamentos entre pessoas e com as instituições. Um processo que intensifica o da digitalização e da desmaterialização em curso, com repercussões sociais e políticas.

– A criação de envolventes sociais estruturadas por sistemas de IA com uma capacidade de controlo pessoal e social ímpar na história da Humanidade. Mais preocupante que a imitação das capacidades humanas pelas "máquinas" é a gradual tendência para um comportamento humano (condicionado ou mesmo imposto),sincronizado e semelhante ao de "máquinas". O relacionamento directo, natural, entre humanos a ser substituído por um "jogo digital" através de interfaces (ecrãs) exigindo uma relação limitada e assimétrica, decorrente das características

dos modos automáticos de comunicação ou conexão (os contactos telefónicos com instituições, com empresas, sujeitos a opções restritas sem a flexibilidade da comunicação pessoal directa, é um dos exemplos).

– Os impactos no trabalho acompanham a técnica desde a revolução industrial. É frequente a afirmação que a IA tomará conta do que é muito complexo e eliminará actividades penosas ou repetitivas. Os empregos que desaparecem poderão ser substituídos por outros, mas resta saber qual será a relevância social dos novos tipos de actividades e daqueles que permanecerão. O impacto no trabalho acompanha a técnica desde a revolução industrial. É frequente referir (e por vezes desconsiderar) os movimentos de revolta laboral designados por "ludistas" (início do século XIX) como sendo opostos a um progresso inevitável. Estes movimentos de luta e protesto de quem sente o seu emprego ameaçado continuaram e continuarão a ocorrer, como aconteceu no século XX (globalização da economia) e também no século XXI (e.g. a reação à substituição pela IA de actividades na indústria do cinema). Não deixando de ser um aspecto importante e que poderá vir a ser muito significativo no futuro (as previsões não são concordantes), as preocupações da filosofia da técnica e da ética que têm sido manifestadas, há mais de um século, incidem mais sobre outros aspectos.

– Salienta-se, em algumas aplicações de IA, uma superioridade epistémica e opaca que desafia e tende a impedir (ou impede mesmo) a explicabilidade das decisões e a responsabilidade pelos resultados. Este facto pode colocar em causa garantias jurídicas e a aplicação de princípios de Direito. Uma tendência para se considerar a IA como manifestação de Verdade e a induzir uma submissão progressiva (cómoda) à sua utilização. As utilizações da IA em actividades socialmente críticas (e.g. medicina, justiça, ensino, sistemas bancários...) exigem um cuidado especial na protecção de direitos e da individuação autónoma humana. A racionalidade e a eficência não são os únicos factores a ter em conta. Há características humanas, mais analógicas que digitais, que se consideram ser (por enquanto) indispensáveis.

– A introdução progressiva de sistemas de IA e de "robots" autónomos nos dispositivos militares não tem sido objecto de uma regulação reconhecida internacionalmente, mas, a par dos perigos no domínio da cibersegurança, esta actividade pode configurar uma ameaça existencial. É um domínio em grande desenvolvimento, muito opaco ou

secreto, mas que permite desenvolvimentos à margem da apreciação ética. O comando de sistemas críticos de defesa e ataque por dispositivos autónomos conectados (fora de uma intervenção humana), nomeadamente os dispositivos baseados em armas nucleares, constitui uma preocupação.

– Os impactos nas democracias da utilização de técnicas avançadas de IA desafiam valores consolidados e colocam em perigo o funcionamento e a confiança das instituições do Estado. Os desafios à percepção da verdade e da realidade social estão identificados, bem como a manipulação de opiniões e decisões em contexto eleitoral. A utilização intensiva da IA em plataformas digitais (redes sociais) pode "sincronizar" em simultâneo ideias em milhões de pessoas, com imagens ou factos "construídos", e criar movimentos de massa para determinados objectivos. A democracia representativa como é conhecida terá dificuldade em ser defendida ou mantida sob a pressão dos sistemas de IA e a sua utilização intensiva para efeitos políticos. Em qualquer regime, a utilização de sistemas eficientes de vigilância e controlo podem ser uma tentação. Em sistemas autoritários podem (já) ser uma realidade.

Síntese

Os aspectos referidos podem vir a afectar a identidade e função social dos humanos segundo diferentes dimensões, nomeadamente a espiritual. Alterações da humanidade em resultado das aplicações digitais e da IA serão inevitáveis, mas a experiência e o bom senso aconselham a reflexão e uma actuação eficaz. A preocupação com os efeitos da técnica, da tecnologia, tem sido objecto de muitos contributos há mais de um século. Muitas personalidades deram o seu testemunho num acervo notável. Salienta-se, entre muitas outras, a obra de Gunther Anders "A Obsolescência do Homem" (1956 e 1980)[12]. Poderosos interesses económicos estão actualmente em jogo, nomeadamente na aplicação da IA como instrumento de indução de consumo, de aumento de produtividade e crescimento económico. Atendendo à dificuldade actual no controlo democrático das empresas tecnológicas, a regulação baseada

[12] Gunther Anders, "A Obsolescência do Homem", Vol. I (1956) e Vol. II (1980), Ed. Pre-Textos (em espanhol).

em princípios éticos, acompanhada por uma regulamentação jurídica consistente e eficaz, torna-se necessária. A ética constitui um referencial para orientação na identificação do que é aceitável fazer e do que é aceitável aplicar na sociedade humana[13]. Pode-se mesmo considerar que se corre o risco de um verdadeiro "periculum in mora" se a regulação da IA não for encarada, no presente, com determinação.

Muitos dos avanços, perigos ou ameaças suscitados por obras de ficção científica não são de ter em conta na IA actual. Outras tecnologias e ciências da vida podem colocar riscos vitais mais relevantes, como parece poder ocorrer, a título de exemplo, com as designadas "bactérias espelho"[14]. Mas há afirmações interessantes de especialistas como o de I. J. Good em 1965[15]: "É mais provável do que não que, no século XX, uma máquina ultrainteligente seja construída e que seja a última invenção que o homem precisa fazer, uma vez que levará a uma 'explosão inteligente'. Isso vai transformar a sociedade de uma forma inimaginável." Talvez tal ocorra no século XXI, mas tudo leva a supõr que esta revolução (a super Inteligência Geral Artificial, IGA, a antecipação de uma "singularidade") esteja a ocorrer. Será a última?

3. Dificuldade na regulação

A revolução digital e da IA está em curso e as dificuldades presentes na regulação eficaz da IA são muitas e diversas.

– Destaca-se a ideia geral do progresso tecnológico como motor indiscutível da melhoria da condição humana. Um ideia antiga que o progresso "traz trabalho e pão". Este tipo de ideia e a da necessidade de inovação tendem a constituir uma barreira psicológica à análise crítica (ética, social, filosófica...) de novas tecnologias. O princípio da inovação tende a sobrepôr-se ao princípio da precaução. A resistência de produtores dos dispositivos técnicos a uma regulamentação de segurança é histórica. Acidentes trágicos é que impulsionaram muitos dos regulamentos

---

[13] M. C. Patrão Neves e M. Graça Carvalho (eds.), "Novas Tecnologias, Ética Aplicada", 2018, Edições 70.

[14] Adamala et al., Technical Report on "Mirror Bacteria: Feasibility and Risks", December 2024, Stanford University.

[15] Irving John Good, Speculations Concerning the First Ultraintelligent Machine. Advances in Computers, Vol. 6 (1965), Academic Press Inc. New York.

técnicos no século XIX. Uma resistência semelhante à que identificamos com a IA no século XXI e com outras tecnologias no século XX.

– As dicotomias superficiais ou maniqueístas, optimismo-pessimismo ou progressista-conservador, são exemplos deste dispositivo de resistência que poderá ter raízes profundas, antropológicas, mas que se afastam da racionalidade ponderada aplicada às novas condições e de uma prudência temperada pela sabedoria. A análise humana racional não segue dualismos rígidos (do 1-0), e com essa dicotomia o realismo fundamentado tende a ser subvalorizado.

– A defesa da inovação e da investigação tecnológicas é compreensível e suscita o apoio de poderes predominantes, económicos, comerciais e políticos, nacionais e internacionais. Na União Europeia colocam-se nesta balança disputas geoestratégicas e de competição entre potências mundiais que tendem a minimizar a importância da apreciação ética das aplicações. Os princípios éticos podem, assim ser, considerados facilmente como "radicais" e prejudiciais ao avanço da sociedade. Uma afirmação muito usada pelos que se opõem: "a Europa tem excesso de normas e de regulação". No embate entre o desenvolvimento acelerado da tecnoeconomia e as preocupações éticas há uma assimetria objectiva difícil de ser superada. O primeiro apresenta-se como gerador de poder económico enquanto a ética sugere uma prudência associada a outros valores.

– O conceito de disrupção, como atitude vanguardista positiva, benéfica, manifesta-se junto da opinião pública. A potencial desvalorização pelo respeito dos direitos fundamentais do ser humano face a valores de oportunidade instrumental constitui uma dificuldade significativa da regulação baseada na ética. Acresce a rápida adaptação, e até submissão, das pessoas e das instituições a dispositivos novos na procura de um incremento de eficiência, poder, rapidez e comodidade na satisfação de desejos e de concretização de objectivos. Umas vezes (talvez uma caracteristíca inconsciente, antropológica) para evitar uma exclusão social ou comercial sem cuidar de potenciais consequências sociais, a médio ou longo prazo. Esta dinâmica conduz a alterações de valores societais que podem dificultar a aceitação de uma regulação. O caso dos "smartphones" e das redes sociais e da respectiva influência no comportamento dos jovens (e talvez de adultos) é já um exemplo a ter em conta.

– A aplicação da análise e avaliação formal dos riscos da IA tem dificuldades. A quantificação precisa dos danos sociais e pessoais é ainda difícil. As probabilidades de ocorrência perdem significado científico: os eventos estão em curso, os efeitos finais previstos nunca ocorreram no passado e serão actos futuros, isolados e difíceis de serem simulados laboratorialmente. Em rigor, estamos perante ameaças que devem ser reconhecidas no presente, são riscos mal definidos.

– A prevenção de potenciais efeitos considerados prejudiciais poderia ser tentada através de uma auto-regulação na fase de concepção dos dispositivos ou da actuação de conselhos de ética empresariais independentes. O Código de Conduta de Hiroshima poderá ser considerado um exemplo de uma auto-regulação preventiva. Contudo, este tipo de regulação não se tem revelado muito eficaz e a multiplicidade de novos tipos de utilização dos dispositivos a jusante do produtor inicial é uma das possíveis razões.

– O desenvolvimento da IA, tal como o de outras tecnologias, é considerado por algumas linhas de pensamento como o resultado de uma co-evolução técnico-humana. Um tipo de evolução resultante da forte influencia da técnica no ser humano e da continuidade, ao longo do tempo, das relações entre humanos e técnicas. Esta evolução não anularia a evolução biológica, mas seria um novo tipo de evolução híbrida e social muito mais rápida. Esta posição torna mais difícil a auto-regulação: o processo de produção resultaria de uma continuação com utilização e alteração de processos anteriores. Os progressos não resultam de uma posição inicial pura ou de "criacionismo digital" como é designado por Edward A. Lee, um dos mentores da co-evolução[16]. Uma ideia muitas vezes transmitida, talvez inconscientemente, por especialistas ao comentarem uma novidade técnica: "não surgiu agora, é o fruto de uma evolução técnica em curso desde há um tempo...". Uma nova antropologia que pode ter convergências com outras posições filosóficas que compreendem (e podem considerar correcta) uma certa despromoção do ser humano na posição face ao Universo corrigindo, assim, uma posição considerada excessivamente antropocêntrica. Uma co-evolução que atenuaria o poder do humano obrigado a partilhar com os sistemas IA

[16] Edward A. Lee, "The Coevolution. The Intertwined Futures of Humans and Machines", 2020, Ed. The MIT Press, Cambridge, Massachusetts London, England.

possuidores de capacidades de inteligência muito superiores. Esta posição, sem impedir acções de regulação, tenta ou sugere uma parceria quase fraterna.

– As novas tecnologias, em particular a IA, não têm fronteiras terrestres e estão "dominadas" por poderosas empresas que defendem os seus interesses desafiando, por vezes, o poder de governos democráticos ou não. Uma regulação que não abranja todo o espaço de intervenção da IA tem dificuldade em ser rápida e plenamente eficaz. No entanto, há que começar por uma parte e tentar influenciar o todo. É o caso do regulamento, o IA Act, da UE que poderá ser uma referência para outras propostas internacionais de regulação. A aplicação deste novo regulamento exigirá uma organização relativamente complexa e poderá ter de enfrentar dificuldades processuais em algumas situações.

Síntese

A protecção contra as aplicações consideradas socialmente nocivas é, actualmente, o processo de regulação considerado possível. É o processo com mais possibilidade de eficácia, não obstante as diferentes dificuldades referidas anteriormente. Está, contudo, muito dependente da capacidade de garantir o respectivo cumprimento, das alterações tecnológicas rápidas que podem tornar ineficazes alguns dos aspectos regulatórios e exigir uma avaliação e uma adaptação consequente. Exige uma estrutura relativamente complexa de apoio. Do ponto de vista ideal deveria ter um âmbito internacional. O AI Act é, na Europa, como já foi referido, o exemplo recente que se espera possa vir a atingir os objectivos desejados.

4. O dilema dos opostos

uma análise ponderada dos efeitos potencialmente positivos e negativos da IA como uma tecnologia poderosa e tendencialmente universal conduz a um dilema de apreciação: a cada aspecto potencialmente nocivo ou perturbador para a sociedade humana, podemos encontrar, como num espelho, um aspecto excepcionalmente promissor. Exemplos:

– A externalização da memória pode propiciar uma gradual perda dessa capacidade em resultado da não exigência na sua prática, mas

conhecemos a grande vantagem prática no fácil acesso às informações armazenadas em sistemas digitais. É uma ajuda já indispensável no apoio à actividade intelectual e uma forma de proporcionar acesso a muitos conhecimentos que tenderiam a ser ignorados. Não é um tipo de preocupação de agora: Platão (428-347 a.C.), no texto "Fedro", quando Sócrates narra o mito de Teuth no qual o deus egipcío considera que a escrita tornaria os homens mais sábios e de melhor memória. Mas tal eficácia é questionada: "... provocará nas almas o esquecimento do quanto se aprende devido à falta de exercício."

– A transformação do conhecimento com base na análise de dados e produção de resultados poderá alterar a criatividade e a exigência intelectual humanas para explicar fenómenos ou realidades, podendo colocar em causa o conhecimento teórico. Por seu turno poder-se-á atingir resultados de um modo mais rápido e eficaz e "libertar" os humanos do esforço em conhecer uma justificação ou o enquadramento teórico de uma solução. Menos esforço e mais tempo para questionar e aplicar.

– A intervenção de algoritmos de IA falantes em sistemas de comunicação públicos pode ser muito limitativa e até penosa para humanos, mas pode ser uma acção para diminuir encargos ou uma resposta desesperada face a uma eventual carência de recursos humanos.

Há, contudo, a consciência de que algumas práticas consideradas como adquiridas e respeitadas podem correr o risco de desaparecerem com inovações. A regulação terá assim de ser eficaz nos propósitos, mas prudente, para se conciliarem os aspectos positivos da IA com a defesa de direitos do ser humano.

5. Considerações finais

– A IA está em pleno desenvolvimento e será cada vez mais um factor de mudança da sociedade humana prosseguindo e ampliando a transição digital. A capacidade rápida de análise e gestão de dados e de interacção com os humanos coloca desafios e perplexidades. Tal como tem acontecido com todos os outros objectos técnicos, poder-se-ia considerar a IA como um produto técnico com novas funcionalidades específicas, mas ontologicamente distinto dos seres humanos. A função de capacidade de decisão inteligente, de produção de "obras" e de diálogo bem como a ideia de uma inteligência superior vieram perturbar, no caso da

IA, o tipo de relação homem-máquina. As inferioridades humanas em velocidade, deslocação, força ou energia foram aceites e superadas. A tentativa de reprodução e o confronto operacional com as capacidades humanas consideradas superiores e únicas veio causar fascínio e perturbação. Mas a realidade não deve ser esquecida: os humanos são o resultado de uma herança biológica de milhões de anos baseada num processo evolutivo relativamente lento. Os algoritmos de IA actuais são um produto humano, são "artefactos artificiais" com uma natureza não biológica muito diferente daquela que é a dos humanos. O futuro dirá se esta situação irá ser alterada com capacidades acrescidas de acção autónoma intencional e de criação de novos sistemas sem intervenção humana criando, então, uma nova forma de vida.

– A IA e outras novas tecnologias estão a criar um "mundo novo", anunciado como um "admirável mundo novo", lembrando o título da obra muito conhecida de Aldous Huxley, mas convém não esquecer a resistência (tenaz) no passado à prevenção activa contra os efeitos das tecnologias predominantes no ambiente, no clima do planeta. No caso das novas tecnologias, convém sublinhar que o alvo pode ser directamente a Humanidade, o que aconselha uma atenção ética acrescida.

– As propostas ou iniciativas de uma eventual convergência induzida entre humanos e sistemas de IA, nomeadamente a introdução de dispositivos artificiais para aumento de capacidades humanas ou a promoção de alguns desses sistemas com estatutos legais equivalentes aos existentes para humanos ou para instituições criadas e dominadas por estes (e.g. a cidadania ou a personalidade jurídica), devem ser encaradas com muita prudência.

– É conveniente saber aproveitar o melhor possível as novas capacidades da IA para a resolução de problemas complexos, nomeadamente os da saúde e os sociais a nível planetário. Com mais atenção para o desenvolvimento, ao contrário de uma ênfase redutora no crescimento e na eficiência.

– A regulação afigura-se ser, no presente, o modo mais eficaz de controlo de algumas aplicações da IA e para estabelecer "baias" que evitem excessos e danos, mas potenciando os benefícios desta tecnologia. Não obstante as dificuldades práticas referidas, a regulação ainda transmite esperança e alguma confiança no futuro.

Mas é de apontar alguns condicionantes:

– A formação académica deveria ser mais multidisciplinar e integrada. Uma proposta frequente, mas com resultados pouco visíveis. A visão integrada do progresso humano, social e técnico, e dos objectivos prioritários e desejáveis para o futuro beneficiaria de uma formação menos compartimentada.

– Os poderes políticos democráticos deveriam ter os instrumentos de intervenção e de regulação da IA necessários para maximizar os benefícios desta nova tecnologia nomeadamente na resolução de problemas sociais prementes, mas também para controlo das ameaças indesejáveis em diferentes domínios ou aspectos da sociedade.

– Os dispositivos digitais e de IA são aplicados, de um modo cada vez mais geral, em todo o nosso planeta. Uma regulação eficaz deve ser o mais abrangente possível. Esperamos que o AI Act da União Europeia motive outras iniciativas e seja uma referência internacional.

Conclusão: a regulação eficaz da IA é uma necessidade difícil de ser concretizada, mas deve ser defendida como uma manifestação de responsabilidade social e civilizacional.

António Betâmio de Almeida
Lisboa, 15/12/2024

Data Protection Compliance in Messaging Apps in Brazil: Measuring the Effectiveness of the Brazilian General Data Protection Law

LUCA BELLI, WALTER BRITTO GASPAR,
BIANCA KREMER, RODRIGO GOMES,
BEATRIZ COSTA, FERNANDO NAEGELE,
SOFIA CHANG NOGUEIRA,
DANIEL DORE LAGE

Abstract: This article assesses compliance with the Brazilian General Data Protection Law (LGPD) by messaging apps in Brazil. It analyses the privacy policies of the six largest messaging platforms in Brazil in 2023-2024 through the lens of three criteria considered as essential for the implementation of LGPD and any other data protection framework: regulatory effectiveness, privacy policy explainability, and accountability of the entities processing personal data. Interdisciplinary research tools and mixed methodologies — including analytical techniques and bibliographic review — are applied to documentary and legislative sources. The article develops a comparative analysis of the technical and legal aspects of these platforms' privacy policies.

Keywords: Data protection, LGPD, Explainability, Platforms, Social media

1. Introduction

Over the past decade, data protection has emerged as a key priority for policymakers around the world, with 70% of nations having implemented such legislation (Apacible-Bernardo & Fischer, 2024; Greenleaf, 2023). This trend has gained momentum due to multiple high-profile scandals, such as the revelations of former NSA-contractor Edward Snowden,[1] and the Facebook-Cambridge Analytica scandal,[2] both tellingly illustrating the need for sound data protection regimes. In such context, established data protection standards, such as the European Union's General Data Protection Regulation (GDPR), have acquired global prominence, becoming a global benchmark and prompting non-European policymakers to follow the EU lead, adopting similar frameworks.

National jurisdictions, such as Brazil with its General Data Protection Law (*LGPD — Lei Geral de Proteção de Dados*), have implemented comprehensive data protection laws aimed at giving individuals more control over their personal information, while also providing more legal clarity for businesses and other entities processing personal data. This article critically explores how Brazilian regulation influences the structure and content of privacy policies (Policies),[3] affecting trust and clear

---

[1] Hu, Margaret, Taxonomy of the Snowden Disclosures (November 28, 2015). Washington and Lee Law Review, Vol. 72, 2015, Washington & Lee Legal Studies Paper No. 2016-5. https://ssrn.com/abstract=2730245 ; Pohle, Julia and Van Audenhove, Leo, Post-Snowden Internet Policy: Between Public Outrage, Resistance and Policy Change (March 22, 2017). Media and Communication, 5(1), 1-6. https://ssrn.com/abstract=2939503; for a full media coverage of the scandal, see: Ewen Macaskill and Gabriel Dance. NSA Files Decoded: What the revelations mean for you. The Guardian. (1 November 2013). https://www.theguardian.com/world/interactive/2013/nov/01/snowden-nsa-files-surveillance-revelations-decoded#section/1

[2] Kim, Tami and Yemen, Gerry, Facebook, Cambridge Analytica, and the (Uncertain) Future of Online Privacy. Darden Case No. UVA-M-0979. http://dx.doi.org/10.2139/ssrn.3660467; Mukunde, Doreen, Analyzing the Influence of Data-Driven Marketing: Cambridge Analytica's Role in Political Campaign Strategies and Public Perception (April 14, 2022). http://dx.doi.org/10.2139/ssrn.4749836 ; for a full media coverage of the scandal, see: The Guardian. The Cambridge Analytica Files. 2018. https://www.theguardian.com/news/series/cambridge-analytica-files

[3] Although we are aware of the technical difference between the terms "Privacy Notice" (a document intended for data subjects and interested third parties) and "Privacy Policy" (a document intended for the data controller's internal audience), for the purposes of this

understanding by individuals. The results are based on a careful examination of the Policies of the largest messaging apps in Brazil (Kemp, 2024), elucidating the empirical impact of data protection regulations.

As we will illustrate in this paper, transparency[4] and comprehensibility of Policies, although essential for compliance and effectiveness of these regulations, remain uncertain. Clear and understandable Policies are crucial because their function is, precisely, to allow individuals to make informed decisions when providing their data to multiple providers. This article examines the structure and content of Policies through objective standards set by the LGPD and via interpretation based on jurisprudence and technical guidance by the Brazilian Data Protection Authority (ANPD).

1.1. Methodology

This article analyses the Policies[5] of six amongst the most popular messaging platforms in Brazil — WhatsApp, Facebook Messenger, Instagram,[6] Signal, Telegram, and Discord — chosen due to their extensive

---

article, we will use both as synonyms, considering that the "Privacy Policy" denomination is the dominant practice adopted by the stakeholders involved in this study.

[4] See, in this regard, item 2 of the WP29 guide (Article 29 Working Party, 2018): "Transparency is a long established feature of the law of the EU. It is about engendering trust in the processes which affect the citizen by enabling them to understand, and if necessary, challenge those processes. It is also an expression of the principle of fairness in relation to the processing of personal data expressed in Article 8 of the Charter of Fundamental Rights of the European Union. Under the GDPR (Article 5(1)(a)), in addition to the requirements that data must be processed lawfully and fairly, transparency is now included as a fundamental aspect of these principles. Transparency is intrinsically linked to fairness and the new principle of accountability under the GDPR. It also follows from Article 5.2 that the controller must always be able to demonstrate that personal data are processed in a transparent manner in relation to the data subject. Connected to this, the accountability principle requires transparency of processing operations in order that data controllers are able to demonstrate compliance with their obligations under the GDPR."

[5] The analysis was conducted in May 2024. The versions of the Policies analysed were as follows: a) Discord – 15 March 2024; b) Facebook Messenger and Instagram – 27 December 2023; c) Signal – May 2018; d) Telegram – 31 March 2024; e) WhatsApp – Policy dated 4 January 2021 and Brazilian User Notice dated 12 September 2023.

[6] Meta's Privacy Policy covers both Instagram and Facebook Messenger.

user bases.[7] While these platforms are not exclusively messaging apps, they all provide messaging functionalities, either as a core function or as a relevant one, fitting the scope of the research. The evaluation of each Policy was based on a set of essential criteria,[8] distilled from the LGPD provisions as highlighted below (Figure 1).

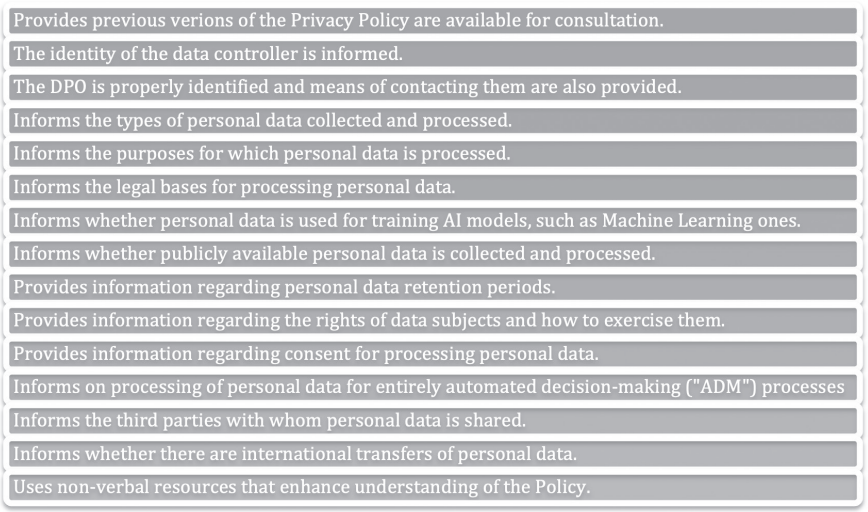

Figure 1. Evaluation criteria of the Privacy Policies subject to analyses.

These criteria have been identified as representing the essential elements of what could be considered as "meaningful transparency," which is the idea that transparency and accountability should not merely imply the disclosure of information but also the auditability of the disclosed

[7] According to DataReportal's social media statistics for Brazil in 2024, there were 144.0 million active identities of social media users in Brazil as of January 2024, which is equivalent to 66.3% of the total population. The most used messaging platforms in Brazil are WhatsApp (93.4%), followed by Instagram (91.2%), Facebook Messenger (60.8%), Telegram (56.5%) and Discord (16.4%). In addition, Panorama Mobile Time 2024 measures the popularity of the main messengers in Brazil based on the percentage of the smartphone base that has each app installed: WhatsApp (98%), Instagram (88%), Facebook Messenger (69%), Telegram (63%) and Signal (13%) (Kemp, 2024; Paiva, 2024).

[8] The criteria were developed and refined based on previous work by Professor Nicolo Zingales and his team. The authors would like to express their gratitude to Professor Zingales.

information, to make sure that legal obligations are duly fulfilled.[9] To achieve meaningful transparency, we posit that sometimes legal requirements must be complemented or interpreted in light of technical standards of best practices.

This is the case of the first criterion, the availability of previous versions of the Policy, which we have considered a key enabler of individuals' capacity to assess how the data controller's processing activities have evolved and, consequently, to what extent the initial 'rules of the game' have changed. Although there is no express legal requirement for it in the LGPD, it is a good practice referenced in at least two ISO/IEC standards[10] and we deemed it as an important criterion to assess transparency, explainability, and accountability of data controllers, and, indirectly, their privacy programmes' effectiveness.

Subsequently, the details provided by the Policy were also evaluated in light of a set of basic requirements set by the LGPD. First, we considered that identification of the data controller and the Data Protection Officer (DPO, referred in the LGPD as *Encarregado*[11]) is critical to allow data subjects to know who is responsible for carrying out decisions

---

[9] Belli, Luca; Curzi, Yasmin; Almeida, Clara; *et al*. Towards Meaningful and Interoperable Transparency for Digital Platforms. Addis Ababa, Ethiopia: United Nations Internet Governance Forum, 2022. https://www.intgovforum.org/en/filedepot_download/57/23886

[10] A critical review of current and historical policies and procedures may be required (e.g., in cases where the client enters into litigation and an investigation by a supervisory authority). The organization must retain copies of its associated privacy policies and procedures for a period as specified in its retention schedule (see 7.4.7). This includes keeping previous versions of these documents when they are updated. ABNT NBR ISO/IEC 27701:2019 (Information Privacy Management System - SGPI).

[11] There remain nuanced distinctions between the Brazilian *Encarregado* and the GDPR's DPO, now refined by two complementary ANPD regulations. In summary: (i) Resolução CD/ANPD n.º 2/2022 exempts "agentes de tratamento de pequeno porte" (*small-scale data processing agents*) from the duty to designate an *Encarregado*, provided they offer a reliable channel of communication with data subjects; should they choose to appoint one, the measure is deemed a best-practice in governance (art. 11, §§ 1.º-2.º). (ii) Resolução CD/ANPD n.º 18/2024, which approves the specific Regulation on the *Encarregado*, reaffirms the general obligation of controllers — and, as a voluntary good-practice, operators — to nominate an *Encarregado*, while specifying guarantees of functional autonomy, resource adequacy and freedom from undue interference that approximate the independence afforded to the GDPR DPO (arts. 3, 6 and 10). Hence, although the GDPR continues to lay down a more explicit compliance-monitoring mandate, the Brazilian framework — fortified by the above

regarding the processing of their data, effectively processing their data, and protecting individuals' data protection rights. According to the LGPD, the *Encarregado* is the natural or legal person responsible for acting as a communication channel between data processing agents, data subjects and the Brazilian Data Protection Authority (*ANPD — Autoridade Nacional de Proteção de Dados*),[12] and for assisting the data controller in data protection compliance (Lopes *et al.*, 2022). Hence, the *Encarregado* plays a fundamental role as regards the construction of both transparency and accountability in personal data protection. Emphasizing this fundamental role, the LGPD establishes that the *Encarregados*'s identity and contact information must be publicly disclosed, clearly and objectively, citing the controller's website as a preferred option.[13] Such information is usually found in the Policy, commonly available on the controller's website, in accordance with the DPO requirements set by Resolution CD/ANPD 18/2024 (ANPD, 2024). Additionally, Resolution CD/ANPD 18/2024 also clarified that "identity" means the full name of the person appointed as the DPO.

The LGPD is also adamant with regard to the need to provide clear and accessible information on the types of personal data (including publicly available data)[14] processed, the processing purposes, and

---

Resolutions — has drawn closer to the European model in safeguarding the *Encarregado*'s autonomy and delineating its advisory and liaison functions. .

[12] *General Data Protection Law*, article 5, VIII (Brasil, 2018). See also ANPD (2024b), which details the Regulation on the performance of the person in charge of the processing of personal data.

[13] LGPD, article 41, §1: "The identity and contact information of the person in charge must be publicly disclosed, clearly and objectively, preferably on the controller's website" (Brasil, 2018).

[14] Before the enactment of the LGPD, Bruno Bioni, following the European position, argued that public data could only be used by data processing agents based on legitimate interest if they were processed according to the reasons for their disclosure (Bioni, 2019, p. 269). However, the Brazilian regulation distanced itself from the European decision after adding paragraph 7 to article 7 of the LGPD. Thus, in theory, there would be no express prohibition on such data being used for new purposes, as long as they are legitimate and specific. Although there is no legal prohibition, João Pedro Ferraz Teixeira points out that "the release of the attribution of new purposes for the processing of personal data publicly accessible and made manifestly public, although it may be considered interesting for processing agents, reduces the degree of protection for data subjects, thus reducing the degree of protection of Brazilian legislation when compared to European legislation" (Teixeira, 2021, p. 190).

information regarding data sharing, such as categories of data shared and sharing purposes. Moreover, data processing agents must clearly identify the legal bases that justify the processing, choosing from the options presented in articles 7[15] or 11[16] of the LGPD. Importantly, this list of legal grounds justifying data processing is exhaustive (CJF, 2022) and clear identification of one of the bases is instrumental for the processing to be considered lawful and transparent.

Information provided by the analysed platforms about the types of personal data processed was also examined in detail. Distinguishing personal data types is complex, especially when defining the legal basis for data processing (de Teffé, 2022, p. 45; Doneda, 2019).[17] Conspicuously, such distinction has legal consequences, as the LGPD is stricter when it comes to special categories of personal data ("sensitive personal data"), to limit its use to the minimum necessary and prevent illicit or illegitimate discrimination (Miragem, 2019).

Similarly, information on data subject rights should be clear, accessible, and the data controller must specify what is the available channel for their exercise. This study focused on the rights listed in article 18 of the LGPD.[18] Moreover, we decided to jointly analyse information regarding disclosures of data sharing[19] and international personal data

---

[15] Legal bases for the processing of personal data.

[16] Legal bases for the processing of sensitive personal data.

[17] Danilo Doneda states that: "The elaboration of this category and the specific disciplines applied to it was not exempt from criticism. One of these criticisms states that it is ultimately impossible to define in advance the effects of the processing of information, whatever its nature. And yet, it is increasingly clear that even data not qualified as sensitive, when subjected to a certain processing activity, can reveal aspects considered sensitive about someone's personality, which can lead to discriminatory practices. It is affirmed, in summary, that a piece of data, in itself, is not dangerous or discriminatory – but the use that is made of it can be so" (Doneda, 2019, p. 142 (ebook)).

[18] These rights include confirming whether there is data processing; accessing data; correcting incomplete or inaccurate data; requesting anonymization, blocking or elimination of unnecessary or unlawfully processed data; data portability; requesting deletion of consent-based data (and being informed of the consequences of withholding or revoking consent); and obtaining information regarding entities with whom data were shared.

[19] According to the legal definition of article 5, XVI of the LGPD, the shared use of data is the *"communication, dissemination, international transfer, interconnection of personal data or shared processing of personal databases by public bodies and entities in compliance with their legal*

transfers, as both stem from controllers' decisions that entail data protection responsibilities after the transfers.

Importantly, this study has not delved into data sharing with public entities, due to distinct legal bases and rules. The analysis focused on transparency and accountability to data subjects, ensuring they understand why their data are shared, what protection measures are taken, and how to exercise rights such as consent revocation and/or opposition. For international transfers, controllers' duty must ensure protection equivalent to that required by the LGPD, even when the data are under another jurisdiction.[20]

Lastly, our analysis also evaluated whether the Policy explicitly mentions the use (or non-use) of artificial intelligence (AI) or machine learning (ML) models in personal data processing and automated decision-making activities. Although the LGPD does not expressly mention AI or ML processes, it establishes principles and best practices to protect data subjects, as well as additional transparency obligations specifically resulting from entirely automated decision-making data processing activities which impacts individuals' interests, according to article 20. These factors are increasingly important, as AI-driven technologies powered by personal data introduce new privacy and personality risks.

Regarding automated decision-making, contrary to the GDPR, the LGPD does not explicitly state that the Policy must contain a provision on the existence of such decisions. However, article 20 LGPD grants data subjects the right to request a review of decisions made solely by automated processes[21] affecting their interests, including decisions intended for profiling. Moreover, paragraph 1 of article 20 also grants individuals' the right to obtain "clear and adequate" information regarding the

---

competences, or between these and private entities, reciprocally, with specific authorization, for one or more processing modalities allowed by these public entities, or between private entities."

[20] The ANPD has recently approved its Resolution for International Data Transfers, setting forth the possibility to use either standard contractual clauses (SCCs), binding corporate rules (BCRs) or specific contractual clauses for international data transfers. However, this obligation will only be enforceable in 2025. Thus, it was left out of the scope of our study.

[21] This topic will not address whether the Policies have any provision on the exercise of the right to request the review of automated decisions or how they comply with this right, but only if there are explicit or generic mentions of automated decisions in the processing of data included in the Policies.

criteria and procedures used for the automated decision. Thus, the Policy must provide information about automated decision-making, enabling data subjects to understand when they are subject to such type of processing, and know when and how to request a review.

To conclude, as an important corollary of meaningful transparency, our analysis also considered whether any visual, audible or other multimedia resources were used to enhance understanding of the respective Policies. To ensure transparency, explainability and accountability, the inclusion of non-verbal elements in the Policies was deemed a key criterion for meeting the principles of transparency and the duty to inform imposed to data controllers.[22]

2. Exploring LGPD Compliance and its Significance

This section will present findings of our analysis of the Privacy Policy of the selected messaging services, following the criteria outlined in the methodology. Subsequently, the descriptive results provided in this section will be critically analysed in the discussion provided in the following section.

Importantly, these empirical results are essential to allow data subjects, the ANPD and the Brazilian society at large to have a clearer picture of how some of the largest platforms operating in the country implement data protection legislation and enable data subjects' right to access information regarding how their personal data are processed. Our findings also offer some evident instances of what can be considered as inadequate transparency measures and are complemented with some of our

[22] In the consumer sphere, which also applies here, the Brazilian Superior Court of Justice (STJ) has already determined: "The right to information aims to ensure the consumer a conscious choice, allowing his expectations in relation to the product or service to be actually achieved, manifesting what has been called informed consent or qualified will. In view of this, the command of article 6, III, of the CDC will only be effectively complied with when the information is provided to the consumer in an appropriate manner, thus understood as that which is simultaneously complete, free and useful, prohibited, in the latter case, the dilution of the effectively relevant communication by the use of loose, redundant or devoid of any use for the consumer" (Resp No. 1.144.840/SP, Rapporteur Minister Nancy Andrighi, judged on 03/20/2012).

subjective interpretations aimed at offering further insight, beyond the mere legally required criteria presented in the tables.

2.1. Availability of Previous Versions of the Policies

The availability of previous versions of data protection policies holds significant relevance in nurturing transparency within messaging platforms. Beyond merely serving as a record of changes, these archives embody principles of transparency, accountability and informational self-determination, thus playing an essential data-subject empowerment function. By preserving historical versions, platforms enable data subjects and regulators to trace the evolution of privacy practices, detect substantive alterations, and assess whether modifications reflect better compliance or potentially adverse shifts in personal data handling.

This retrospective accessibility imposes an implicit constraint on arbitrary or opaque changes to the data controller's Privacy Policy, fostering confidence that platform users' personal data rights will not be diminished without appropriate notice. Moreover, the existence of archives supports audit trails by all interested parties (civil society, data subject and regulators), which are vital in investigations or disputes regarding data processing, thereby enhancing institutional responsibility and the overall integrity of data governance frameworks. Interestingly, all six analysed platforms provide access to previous versions of their Policies through links, and they all state that any Policy content updates will be notified to the impacted individuals.

Table 1. Availability of older versions of Privacy Policy

	WhatsApp	Facebook Messenger	Instagram	Telegram	Signal	Discord
Availability of previous versions of the Policy	Yes	Yes	Yes	Yes	Yes	Yes

Telegram stood out for creating a concise summary of changes for each version, detailing specifically what was added, deleted or modified, improving accountability and transparency for users.

2.2. Controller and Encarregado *Identity and Contact Information*

As mentioned in the introductory section, provision of detailed information concerning the data controller and the *Encarregado*, including their identity and contact information, is another essential transparency mechanism. Particularly, the *Encarregado*'s role is pivotal in bridging the communication gap between data subjects and the organization, tasked with monitoring compliance and handling data protection inquiries or complaints, both from the data subjects and from the ANPD. Indeed, transparent identification and accessibility of the *Encarregado* is instrumental to facilitate effective exercise of data subject rights by providing a designated point of contact for clarifications or remedial actions regarding personal data protection matters.

This transparency not only operationalizes accountability at an institutional level but also affirms the organization's commitment to effectively uphold data protection, through dedicated oversight. In legal terms, appointing, identifying and disclosing the *Encarregado* aligns with regulatory expectations, reinforcing procedural safeguards against non-compliance and trust through verifiable channels of redress.

It is therefore interesting to note that there is a stark difference in how the analysed platforms deal with the disclosure of identity of the data controller and the *Encarregado*. The identity of the data controller is disclosed in all Policies analysed, with each messaging platforms satisfactorily meeting this legal requirement by providing at least the name of the controller in their respective Policies.[23] On the other hand, regarding the *Encarregado*'s identity and contact information, none of the six

[23] As an interesting note, the research found that the Discord controller is different depending on where the data subject is located. Discord is the controller within the European Economic Area and the United Kingdom, while Discord Inc. is the controller in the rest of the world.

platforms (five Policies, considering Facebook and Instagram as one) had included this information at the time of our analyses.[24]

Table 2. Controller and Encarregado Identity and contact information

	WhatsApp	Facebook Messenger	Instagram	Telegram	Signal	Discord
Controller's identity	Yes	Yes	Yes	Yes	Yes	Yes
Encarregado's identity	No	No	No	No	No	No

Despite this, all Policies provided at least one way to contact either their customer services channel, their Privacy team or their DPO, demonstrating varying approaches: (i) an e-mail address (Signal); (ii) an interactive chat interface (Telegram[25]); (iii) online forms (Meta, with a different form for WhatsApp); and (iv) multiple options, such as email and mail addresses (Discord).

2.3. Definition of Data Processing Purposes

Clearly defining the purposes of personal data processing within Privacy Policies emerges as a fundamental transparency element emphasizing not only the principle of purpose limitation, but also of good faith,

[24] We highlight once again that our research was conducted in May 2024. After that, the topic of the *Encarregado* was further developed by Resolution CD/ANPD n. 18/2024, published on July 17, 2024. This Resolution's main subject was the role of the *Encarregado* (DPO) and established once more the requirement to publish their identity and contact details. After that, several platforms analysed in our study made changes in their Policies to address this topic. Moreover, the ANPD also initiated an administrative procedure to investigate the lack of the *Encarregado*'s information against 20 companies, which included Telegram. See: ANPD. (2024c, December 13). *ANPD fiscaliza 20 empresas por falta de Encarregado e canal de comunicação adequado*. Autoridade Nacional de Proteção de Dados. https://www.gov.br/anpd/pt-br/assuntos/noticias/anpd-fiscaliza-20-empresas-por-falta-de-encarregado-e-canal-de-comunicacao

[25] Telegram has a bot, called @LGPDbot, for data subjects with questions about privacy and Telegram's Privacy Policy in Brazil, and also an interface that works as a Voluntary Support, where the data subject can submit a question.

which is a fundamental pillar of Brazilian data protection and civil laws. Such explicitness assists in delineating the lawful grounds upon which personal data are collected and processed, preventing the risk of function creep where data might be repurposed for activities beyond the original scope.

The rationale of this essential requirement is to ensure that data subjects are not left guessing how their information is utilized, thus enhancing informed opt-in and/or consent, whenever applicable, and control. Furthermore, articulating personal data processing purposes with precision supports compliance monitoring by providing clear benchmarks against which actual data use can be audited. This clarity mitigates legal uncertainties and reinforces the legitimacy underpinning data collection practices.

Reassuringly, our results indicate that all Policies provide sufficient information about the purposes of the personal data processing, allowing individuals familiar with the LGPD to at least infer the legal basis for the data processing activities. Additionally, two Policies — WhatsApp and Meta — stood out by offering a detailed correlation, exclusive to Brazilian users, between data categories and their respective processing purposes.

Table 3. Information on processing purposes and legal basis

	WhatsApp	Facebook Messenger	Instagram	Telegram	Signal	Discord
Indicated purpose	Yes*	Yes*	Yes*	Yes*	Yes*	Yes*
Indicated legal basis	No**	No**	No**	No**	No**	No**

* These platforms indicated purpose in a more detailed manner, directly correlating data categories and processing purposes.

** None of the platforms explicitly indicated legal bases contained in articles 7 and 11 of the LGPD; however, in all cases legal bases might be inferred from the wording of the purposes.

As mentioned in the Methodology section, a detailed list of the purposes allows for the inference of the legal basis used. For example,

WhatsApp claims that it processes a certain set of data to "help you find out if your contacts are WhatsApp users when you choose to sync your device's contact list..." (emphasis added) (WhatsApp, 2024). This suggests that the legal basis may be consent, as the user 'chooses' to perform a certain action.

Still, without an explicit statement of the legal basis used, such inferences remain speculative and imprecise at best, thus highlighting the need for clearer and more detailed information on this matter. Furthermore, another preliminary conclusion is that the processing purposes are dispersed throughout the Policies, presented through specific examples correlated to certain data categories. While this structure may enhance readability for users, it can also decrease ineligibility of purposes and, ultimately, create opacity regarding the overall processing activities. Additionally, only two Policies explicitly cite the legal basis of legitimate interests (Telegram and Discor), though its use can be inferred on more occasions in the other Policies (for example, WhatsApp).

2.4. Use of Personal Data for AI/Machine Learning Training

The explicit mention of the use of personal data for artificial intelligence (AI) and machine learning (ML) training in Privacy Policies has gained prominence alongside technological developments affecting data processing. This disclosure addresses the elevated risks associated with automated analytics, where personal data may be leveraged to train AI systems aimed at influencing predictions or decisions that directly affect data subjects' interests.

From a data protection perspective, transparency on AI/ML use is critical because it implicates issues of lawful data processing bases, adherence to data protection principles, such as necessity, and potential impacts on fundamental rights, due to the complexity and potential for unintended consequences of the processing, including discriminatory profiling or reinforcement of biases. Hence, transparency on these elements helps demystify often opaque technological processes, equipping users with knowledge to assess their participation in algorithmic systems and to exercise their rights meaningfully.

Particularly, in the Brazilian context, evaluating whether the platform is deploying AI or ML training is instrumental to ascertain whether LGPD's article 20, which establishes "the right to request a review of decisions made solely based on automated processing of personal data that affect their interests, including decisions intended to define their personal, professional, consumer, and credit profiles, or aspects of their personality," and the consequent controller obligation to "provide, whenever requested, clear and adequate information regarding the criteria and procedures used for automated decision-making," is applicable. In this respect, we assessed whether Policies mentioned the use or non-use of AI or ML systems to process personal data. The results indicate that:

a) Signal does not mention AI or ML at all, making it unclear whether such processing occurs, although its strong data minimisation stance may lead the reader to believe it does not occur;
b) Telegram provides only generic mentions that could indicate an AI or ML type of training;[26]
c) Discord explicitly acknowledges using personal data to train AI in at least two instances;[27]
d) Meta offers the most detailed information, clearly outlining the use of personal data for AI training.[28]

---

[26] For example, Telegram's Policy states that "We may use some aggregate data about how you use Telegram to develop useful features."

[27] (i) data used "to help us train models that proactively detect content that violates our policies"; and ii) "we may also use content posted in larger spaces to help us develop, improve, and power our Services, including features that will help you stay on top of conversations, as well as safety that identifies harmful content on the Services and helps enforce our terms of service and Community Guidelines."

[28] Some examples found in Meta's Policy: "For the purpose of receiving data from public sources: to enhance our AI technologies and to support the research and development of AI products, such as translations, computer vision, content understanding, natural language processing, and tools for people and businesses to create content. [...] To research and innovate for social good: We support research in areas such as artificial intelligence and machine learning to, for example, create COVID-19 prediction models. [...] For sharing purposes between Meta companies: For example, your videos can help train our products to recognize objects, such as trees, or actions, such as a dog chasing a ball. They can also help train tools that allow people and businesses to create content, such as images and videos. This technology is used to help us offer new products or features in the future."

Table 4. Mentions of ML and AI

	WhatsApp	Facebook Messenger	Instagram	Telegram	Signal	Discord
Information on ML / AI	No	Yes	Yes	Yes*	N/A**	Yes

* Mentions purposes of processing personal data that can be interpreted as ML/AI without providing further detail.
** Does not mention ML/AI at all, thus it is impossible to evaluate its effort to inform on it.

It is also important to note that none of the Policies expressly specified which personal data are used for model training, nor did they provide information on whether that processing would constitute an automated decision leading to the person's right to explanation and revision.

2.5. Information on Automated Decision-Making (ADM)

As mentioned in the previous section, disclosure of information relating to automated decision-making (ADM) personal data processing activities, including profiling, plays a fundamental role, reflecting a transparency mandate designed to safeguard individuals from the adverse impacts of such decisions and allowing them to know what is the extent and rationale of the applicable ADM process. In this perspective, article 20 LGPD aims at guaranteeing that data subjects must receive meaningful information about the logic involved, the significance of such processing, and potential consequences of the ADM. Hence, here transparency serves a dual juridical purpose: enabling affected individuals to challenge decisions that impacted their interests and promoting accountability by compelling data controllers to justify automated practices.

Clarifying ADM processes allows users to understand whether and how their data influence decisions without human intervention, addressing concerns about fairness, discrimination, and due process in the digital environment. Therefore, together with the abovementioned disclosure about the usage of AI and ML tools, the use of ADM is deemed as instrumental to trigger the need to respond to data subjects' requests

on information about the criteria and processes used to automate the decision-making process.

Interestingly, all platforms analysed either made generic references to automated decision-making in data processing or did not mention it at all. Only Telegram's Policy contained mixed provisions, with inferred ADM in the descriptions of the contact list analysis and friendship recommendations;[29] and explicit mentions of automated algorithms used to analyse chat messages in the cloud, to avoid spam and phishing.[30]

Table 5. Information regarding ADM

	WhatsApp	Facebook Messenger	Instagram	Telegram	Signal	Discord
Information on ADM	Yes	Yes	Yes	Yes	Yes	Yes

The findings indicate that the platforms studied do not satisfactorily fulfil their information duties regarding ADM. In addition, explicit and clear information on the subject were not available, indicating a significant lack of transparency and clarity in the Policies.

2.6. Information on the Categories of Personal Data Processed

Providing comprehensive information on the categories of personal data collected and processed is an intrinsic component of transparency that enhances the data subject's ability to comprehend the extent and sensitivity of data processing activities. Processing specific categories of personal data, such as sensitive personal data, can lead to adverse effects especially in terms of unjust discriminatory processing, carrying varied implications for a wide range of fundamental rights. By explicitly categorizing the types of personal data, Policies (and, subsequently,

[29] "Our automated algorithms can also utilize anonymized sets of phone numbers to estimate the approximate number of potential contacts that an unregistered phone number may have on Telegram. When you open the 'Invite Friends' interface, we display the resulting stats next to your contacts to give you an idea of who might benefit the most from joining Telegram."

[30] "We may also use automated algorithms to analyse cloud chat messages to prevent spam and phishing."

data controllers) enable users to evaluate potential risks and understand the data footprint created by their use of these messaging services. Clear identification of personal data categories also aids compliance by establishing the scope of personal data processing, thereby facilitating targeted safeguards and reducing ambiguity related to data handling.

Our analysis reveals that all platforms mention the categories of personal data collected and processed by them — such as user-provided data, usage and registration data, device data, metadata, and data provided by third parties — and declare, to some extent, which data are the subject of processing. Signal stands out, since it explicitly indicates the exclusive collection of users' "email, phone number, metadata, name, and photo," which could mean that the app does, in fact, minimise personal data collection for its purposes.

Meanwhile, the other platforms provide extensive lists of data categories in their Policies, with Meta and WhatsApp collecting the most different categories of personal data. Moreover, Meta states that even non-users can have their data processed.[31]

Table 6. Information regarding the categories of personal data processed

	WhatsApp	Facebook Messenger	Instagram	Telegram	Signal	Discord
Disclosure of types of data processed	Yes	Yes	Yes	Yes	Yes	Yes

[31] In Meta's Privacy Policy: *We also collect your contacts' information, such as their name and email address or phone number, if you choose to upload or import it from a device, like by syncing an address book. If you don't use Meta Products, or use them without an account, your information might still be collected*. This is also complemented by a specific link for non-user data processing, available at: https://www.facebook.com/help/637205020878504, that highlights: *Meta Platforms, Inc. ("Meta", "we", "our" or "us") processes your name, mobile phone number and/or your email address if we receive it from our users through the contact uploading or contact syncing feature available on Facebook, Messenger or Instagram ("contact uploading"). We process this information even if you are not a user of Facebook, Messenger or Instagram and/or don't have an account with us (a "non-user").*

In addition, most platforms seem to use broad language when defining the categories of personal data, citing basic and generic examples, giving interpretative leeway to collect or process different personal data attributes within the generic categories (for example: "product usage data") in a manner not always aligned with what is legitimately expected by an average user.

2.7. Processing of Publicly Available Data

The processing of publicly available data, particularly in relation to AI training, introduces distinctive transparency challenges rooted in the complexity of legal bases and user expectations. Public data 'scraped' from external websites, social media, or other sources may nonetheless contain personal information, whose processing clearly falls under data protection regulation. Disclosing how such data are collected and processed addresses the tension between data utilization for technological advancement and respect for data rights.

Hence, transparency regarding this practice informs users about indirect uses of their personal information beyond the platform's direct interactions, elucidating the legal grounds, as well as limits of this processing. It thereby fosters essential trust and compliance by how far data practices go. In this perspective, we adopted this criterion to verify how the platforms approach the topic, whether they indicate its use (or non-use) or even allow the inference that they perform data scraping. Signal and Telegram make no mention of processing publicly accessible data, while Discord, Meta and WhatsApp declare that they collect data from sources other than their services.

Interestingly, WhatsApp states, in a complementary Policy, that it processes data collected from other sources and from the platform's own channels. Meta's Policy, in addition to being the one that most claims to collect this type of data, is the only one that declares different purposes for the collection of publicly accessible data.[32]

---

[32] For example, *"Research and innovate for social good. We use the information we have, such as that from researchers and datasets from publicly available sources, professional groups, and non-profit groups to conduct and support research."*

Table 7. Mention of processing of publicly available data

	WhatsApp	Facebook Messenger	Instagram	Telegram	Signal	Discord
Mention of processing of publicly available data	Yes	Yes	Yes	No	No	Yes

2.8. Data Retention

Explicit explanations of data retention practices constitute a fundamental transparency concern, reflecting the principle of necessity, adequacy and purpose limitation embedded in LGPD's article 6 and most data protection laws. Users' understanding of how long their personal data are held, which categories are affected, and the rationale for retention is vital to assess the proportionality and necessity of continued processing. As such, detailing retention periods and relevant factors that impact such retention assures data subjects that their information is not held indefinitely or without lawful basis, mitigating risks of misuse or unauthorised access over time.

Such disclosures also support data minimisation efforts and provide a temporal framework for rights such as erasure, enabling individuals to gauge when and how their data will be removed, thus reinforcing control and trust in the platform's stewardship of information. Our findings illustrate that, with the exception of Signal, each platform provides at least some information about data retention, informing data subjects about the duration of the processing.

Table 8. Information on data retention

	WhatsApp	Facebook Messenger	Instagram	Telegram	Signal	Discord
Definition of duration of processing	Yes	Yes	Yes	Yes	No	Yes

Signal does not provide any type of explanation in its Policy about the duration of data processing. Telegram, Discord, Meta, and WhatsApp Policies have at least one instance where they provide a specific duration, or a factor that shall define the duration, for a particular data processing. Despite this, some particularities should be highlighted about these provisions.

Discord's Policy has a specific item on retention that mentions that the data will be processed for as long as the controller deems necessary to comply with its legal requirements and the purposes of the processing. In addition, the same item provides a link to a specific Data Retention Policy, where Discord establishes different periods for data deletion, depending on the specific purposes and categories of data collected.[33]

The Policies of Meta, WhatsApp and Telegram also use a general rule establishing that the data will be processed for the time necessary to comply with legal requirements and processing purposes, granting some criteria that will be used to assess this need. Telegram's Policy establishes that all data will be automatically deleted if the data subject does not use the application for 6 months,[34] except for metadata, which will be deleted after 12 months.[35]

It is worth noting that Meta's Policy only establishes specific periods in some cases (e.g., data shared with third parties), lacking a specific document (such as a separate Data Retention page) or comprehensive section on the Policy for the topic. On the other hand, in WhatsApp's Policy, there is a specific deadline for deletion applicable to certain cases, such as: (i) 30 days for undelivered messages; (ii) 120 days for automatic account deletion due to lack of use;[36] and (iii) 90 days for account

---

[33] "We retain personal information until we determine that it is no longer necessary for the processing purposes for which we collected or retain it or for legal compliance. You can learn more about data retention periods in our data retention policy."

[34] "If you stop using Telegram and don't stay online for at least 6 months, your account will be deleted along with all messages, media, contacts, and any other data you have stored in the Telegram cloud."

[35] "To improve the security of your account, as well as to prevent spam, abuse, and other violations of our Terms of Use, we may collect metadata such as your IP address, Telegram devices and apps you have used, username change history, etc. If collected, this metadata can be stored for up to 12 months."

[36] "Check your activity every 30 days and delete your account when your phone hasn't been online for 120 days (in general) to maintain security, limit data retention, and protect your privacy."

deletion, if requested by the data subject.[37] However, it is unclear which categories of data will be deleted with the account, and which will still be processed for other purposes (such as to comply with legal obligations).

Thus, on this criterion, while most Policies checked the requirement, Discord is the data controller that provides the most comprehensive information regarding personal data deletion, not because of the existence of a specific document focused on data retention periods, but because the document contained clear, adequate and specific information. As for the other apps, Telegram also has a satisfactory provision, WhatsApp's Policy has commendable points, despite the need of some clarification, and Meta's Policy requires extra provisions, with some indication of retention deadlines for specific purposes and/or legal grounds. Signal's lack of any information regarding data retention can be considered a serious breach of LGPD obligations, especially compared to the other controllers analysed.

2.9. Data Subjects' Rights

Providing clear and accessible information about data subject rights and practical mechanisms to exercise them is critical to transforming formal LGPD prescriptions into tangible protections. Transparency in this regard encompasses the articulation of data protection rights including access, rectification, erasure, restriction of processing, data portability, the right to object to certain types of processing and the right to revoke consent given for specific processing activities.

Explaining how these rights can be exercised, the procedural steps involved, and any potential limitations ensures that users are not merely informed abstractly but are empowered to assert control over their personal data. This procedural transparency aligns with the LGPD's goal of fostering active data subject participation and promotes

---

[37] *"It may take up to 90 days from the start of the deletion process to delete your WhatsApp information. Copies of your information may also remain after the 90 days in the backup storage we use to recover in the event of a disaster, software error, or other data loss event. Your information will not be available to you on WhatsApp during this time,"* in: 'Why and how we treat your data — WhatsApp' (WhatsApp.com) <https://www.whatsapp.com/legal/brazil-privacy-notice/why-and-how-we-process-data> accessed on April 25, 2024.

a user-centric approach to data governance, enhancing both compliance and trustworthiness.

We found that most platforms indicate the data subjects' rights in their Policies, except for Signal. However, platforms vary in how they explain such rights. While some deal with the subject in a tailored manner (e.g., WhatsApp), others simply repeat legal provisions, without going into further detail about how they operate or how to exercise them on that platform (e.g., Discord).

Table 9. Information on data subjects' rights

	WhatsApp	Facebook Messenger	Instagram	Telegram	Signal	Discord
Explanation of data subjects' rights	Yes	Yes	Yes	Yes	No	Yes

Some highlights include Telegram's interesting use of a 'self-service' chatbot for the exercise of data subjects' rights, a technological solution which can be surprisingly effective, and we have not seen this on any other platform. For some of the rights listed in LGPD article 18, automated exercise is an interesting and perhaps effective choice. However, it is questionable whether users will have their rights fully guaranteed, since the bot can only be accessed by users and provides two functions: (i) access the data collected by the platform, or (ii) contact the controller regarding data privacy.

Discord indicates that options for exercising rights may differ depending on the user's age and location. Meta presented a confusing repetition of the legal text,[38] while rights are only entirely mentioned in a separate document intended for Brazilian users. Thus, in addition

[38] By analysing the Privacy Policy in its printed version, the researchers concluded that the repetition of several passages throughout the document makes it difficult to understand the text and tires the reader in general. This practice can be classified as a dark pattern, specifically overload. For more information on this topic: DIRECTORATE-GENERAL FOR JUSTICE AND CONSUMERS (EUROPEAN COMMISSION) *et al*. Behavioural study on unfair commercial practices in the digital environment: dark patterns and manipulative personalisation : final report. Brussels: Publications Office of the European Union, 2022.

to reading and understanding the general Policy, users would need to access a separate document to obtain a full explanation of their rights. Lastly, WhatsApp's Policy demonstrates an explicit path for exercising rights and, although its wording is not clear enough, the platform chose to develop a specific explanation of some of the rights listed in article 18.

2.10. Withdrawal of Consent and its Consequences

Clarifying the right to withdraw consent and its consequences is relevant within transparency frameworks because consent often forms the basis for personal data processing. Hence, informing data subjects that consent is a dynamic and revocable permission reaffirms the voluntariness and control inherent in data protection law. Furthermore, detailing the consequences of consent withdrawal — such as limitations on service functionality or alternative legal grounds for processing — ensures that individuals can make informed decisions about their ongoing use of a specific service.

Transparency on this point dispels assumptions that consent, once given, is irrevocable and emphasises the responsive nature of data processing activities. However, despite the importance of this dimension, we found that neither Signal nor Telegram provide any kind of explanation on how data subjects can revoke consent. Meanwhile, Discord and Meta's Policies lack extensive explanation on how to revoke consent. Importantly, Discord does provide specific information on revoking consent regarding personal data processed for marketing communication;[39] and Meta repeats the formulation of the LGPD, although lacking more information.

WhatsApp's Policy provides tailored information on how data subjects can revoke their consent through access to in-app privacy settings and device settings. WhatsApp also shows some examples where consent is used, such as to collect location data.[40] Regarding the consequences

---

[39] "You may revoke this consent at any time (usually directly through our Services), but please note that you may not be able to use services or features that require the collection or use of such personal information."

[40] "We only rely on your consent under the LGPD in limited circumstances. For example, when we collect and use information that you allow us to receive through device-based settings when you enable certain features and services, such as access to your GPS location,

of revoking consent, only Discord presents an explanation of what will happen if the data subjects withdraw their consent,[41] while all others fail to provide any type of information on this topic.

Table 10. Information on consent withdrawal

	WhatsApp	Facebook Messenger	Instagram	Telegram	Signal	Discord
How to withdraw consent	Yes	Yes	Yes	No	No	Yes
Conse-quences of withdrawal	No	No	No	No	No	Yes

Thus, the preliminary conclusion of the research is that none of the Policies fully complied with the LGPD according to this criterion. While Signal and Telegram's policies were considered the worst, WhatsApp and Discord are the best, although they all require at least some sort of adjustment to achieve full LGPD compliance.

2.11. Third-Party Sharing of Data and International Data Transfer

Information on third-party data sharing practices and international data transfers, including associated risks, addresses complex aspects of data flows that have significant implications for privacy and legal compliance.[42] Messaging services commonly engage with various third

camera, or photos, so that we can provide you with those features and services. Where we process data based on your consent, you have the right to withdraw your consent at any time."

[41] "Information from optional features. Certain features, like contact syncing, may require that you provide additional information (or grant us access to such information) to make them work. This also includes third-party integrations you choose to enable and the data you authorise those third-party services to share with us. For example, when you link a music streaming account, we may collect information about that account such as the song you are listening to in order to display that information on your profile or as your status (if you have chosen to do so)."

[42] For a detailed analysis of international data transfers regimes see. e.g., Belli, Luca et al, Transferência internacional de dados pessoais na América Latina: rumo à harmonização de normas, 1.ª edição. Rio de Janeiro, RJ: Lumen Juris (2024). https://hdl.handle.net/10438/36141

parties — such as cloud providers, analytics companies, or service subcontractors — and often transfer data across borders. Transparency in disclosing recipients, transfer mechanisms, and applicable safeguards is therefore vital to let users understand the extent and nature of data dissemination beyond the immediate service provider.

Additionally, outlining risks linked to jurisdictional differences in data protection standards, potential governmental access, or security vulnerabilities helps data subjects appreciate the broader context of their data's journey, while complying with security and accountability principles set by article 6 LGPD. In this context, it is interesting to note that all platforms have stated that they share personal data with third parties and transfer data to other countries, and we can observe that, regarding data sharing, there are a wide variety of ways to disclose it.

All analysed platforms declare that personal data sharing is made by suppliers and partners that assist them in their operations (e.g., technological infrastructure, payments or security checks). Apart from Signal, all other platforms share personal data with other companies in their economic group. Discord and Meta declare to share personal data with advertising and marketing companies, without listing or mentioning any example of these companies. Meta's Policy lists more personal data sharing purposes than others, because it consolidates all its products into a single policy. However, it does not relay which categories of personal data are being shared with what types of third parties, nor explicitly mentions any recipients for these purposes. In Meta's Policy, it is possible to find categories of companies that receive personal data, which include everything from game developers to "industry peers, such as other online platforms and technology companies."

	WhatsApp	Facebook Messenger	Instagram	Telegram	Signal	Discord
Mention of third-party sharing	Yes	Yes	Yes	Yes	Yes	Yes
Mention of international transfer	Yes	Yes	Yes	Yes	Yes	Yes

Regarding international data transfers, some of them (Meta and Telegram) name at least 2 countries/regions to which personal data of their users is transferred, although broadly (*"and for other countries"*). None of the platforms has a comprehensive, assertive list of countries, regions or third parties that receive such data transferred internationally. Meta and Discord claim to use standard GDPR clauses, at least for data transfers leaving the European Economic Area. None of them cited the use of binding corporate rules (BCRs) when providing the information that personal data of their users is shared with other companies from the same group. All platforms only state that the transfer may occur to other data centres, allowing the inference that the data are possibly transferred for reasons of redundancy, latency, and availability. Both Signal[43] and Meta[44] mention transfers of personal data to other third parties located in other countries.

2.12. Use of Non-verbal Language

The use of non-verbal language in Policies, through visual elements such as icons, charts, pictograms, videos, and interactive graphics, plays a critical role in enhancing transparency by complementing textual explanations with intuitive clarity. LGPD's article 6 emphasizes the necessity for information to be "clear, precise, and easily accessible," criteria that are challenging to meet through dense legal prose alone. In this context, the use of visual communication resources serves as a cognitive aid that can reduce the complexity of technical and juridical language, making Policies conditions more comprehensible to a diverse audience, including those with limited legal or technical literacy.

Data controllers can facilitate quicker information retrieval, support retention of key concepts, and encourage an active user engagement with data governance through the employment of well-designed non-verbal resources when providing transparency regarding personal data processing practices. This approach not only helps users grasp the ramifications of data processing more effectively but also mitigates information

---

[43] "Other countries where we have or use facilities, service providers, or partners."

[44] "Other countries where our partners, vendors, service providers, and other third parties are located outside of your country of residence."

asymmetry inherent in asymmetrical knowledge relationships between platforms and users, ultimately fostering a more inclusive and user-centred paradigm of data protection communication. Consequently, the strategic use of non-verbal language goes beyond a simple aesthetic consideration, representing a substantive factor in operationalizing meaningful transparency within privacy governance frameworks.

Table 11. Use of non-verbal language

	WhatsApp	Facebook Messenger	Instagram	Telegram	Signal	Discord
Use of non-verbal aids	No	Yes	Yes	No	No	No

Of the platforms studied, Meta is the only one that makes use of visual resources to provide information regarding personal data processing, through the use of images and videos explaining some topics of its Policy. Although WhatsApp is part of Meta's economic group, the same standard has not been adopted, with no multimedia or interactive material available to users.

3. A Limited Compliance

The results presented in the previous sections (summarised in Table 12, below) shed important light on whether and how messaging platforms incorporate transparency, accountability, and explainability into their most immediate privacy communication front, i.e., their Policies. Our effort attempts to go beyond the mere verification of these documents in search of verbatim reproductions of statutory provisions, aiming to discern what is implicit in these Policies, what is omitted and what reinforces transparency, accountability and explainability.

Table 12. Summary of responses over all platforms

Criteria	Whats App	Facebook Messenger	Instagram	Telegram	Signal	Discord
Availability of previous versions of the Policy	Yes	Yes	Yes	Yes	Yes	Yes
Controller's identity	Yes	Yes	Yes	Yes	Yes	Yes
Encarregado's identity	No	No	No	No	No	No
Indicated purpose	Yes*	Yes*	Yes*	Yes	Yes	Yes
Indicated legal basis	No**	No**	No**	No**	No**	No**
Information on ML / AI	No	Yes	Yes	Yes***	N/A§	Yes
Information on ADM	Yes	Yes	Yes	Yes	No	No
Disclosure of types of data processed	Yes	Yes	Yes	Yes	Yes	Yes
Mention of processing of publicly available data	Yes	Yes	Yes	No	No	Yes
Definition of duration of processing	Yes	Yes	Yes	Yes	No	Yes
Explanation of data subjects' rights	Yes	Yes	Yes	Yes	No	Yes
How to withdraw consent	Yes	Yes	Yes	No	No	Yes
Consequences of withdrawal	No	No	No	No	No	Yes
Mention of third-party sharing	Yes	Yes	Yes	Yes	Yes	Yes
Mention of international transfer	Yes	Yes	Yes	Yes	Yes	Yes
Use of non-verbal aids	No	Yes	Yes	No	No	No

* These platforms indicated purpose in a more detailed manner, directly correlating data categories and processing purposes.
** None of the platforms explicitly indicated legal bases contained in articles 7 and 11 of the LGPD; however, in all cases legal bases might be inferred from the wording of the purposes.
*** Mentions purposes of processing personal data that can be interpreted as ML/AI without providing further detail.
§ Does not mention ML/AI at all, thus it is impossible to evaluate its effort to inform on it.

Together, the elements analysed in the previous sections collectively shape the contours of meaningful transparency in the data protection practices of messaging platforms, reflecting legal imperatives and the ethical dimensions of digital privacy. Their analysis reveals the multifaceted nature of transparency, integrating rights, responsibilities, and technological realities into a coherent framework that sustains trust and compliance.

Importantly, the selected group of platforms is particularly relevant for the centrality of the services they provide — messaging services, either as a core activity or as a relevant functionality — for the Brazilian social, political and economic development. In light of their relevance, we believe they deserve a particularly careful scrutiny from data subjects, civil society and regulators. In this respect, Policies play a key function as they can mitigate opacity regarding the way these online services work. However, such documents may also work to further conceal how these services are provided and how personal data are processed for the performance of the main and secondary functions within the platforms.

Worryingly, we observed multiple failures in compliance with the LGPD in different degrees of complexity. As illustrated in the previous sections, there are evident flaws in the implementation of communication efforts, such as not providing accurate information about the consequences of withdrawing consent or about data retention periods, as well as more sophisticated non-conformances, such as insufficiently explaining the use of personal data in the contexts of AI models and related technologies, as well as in automated decision-making.

Some concrete examples demonstrate the challenges of achieving a complete understanding of data processing activities based exclusively on the platform's Policies. Signal's overly short and simplified policy lacks sufficient detail and falls into seemingly contradictory statements, possibly due to insufficient explanations. Although the company has vocally demonstrated in other statements (moxie0, 2017) its concern with designing systems that are conscious of privacy and security, its Policy does not reflect the same level of care. This may stem from a presumption that, due to the nature of their services and how it is structured, their main users are tech-savvy ones, and, thus, completely aware of how a company such as Signal may process their data even if the information is not provided in its Privacy Policy. Other examples

include Discord's mention of automatic data processing for content control without specifying how this occurs and how it relates to transparency and explainability obligations in relation to LGPD's article 20; and the widespread opacity regarding the personal data sharing practices; and Meta's generic transparency about data sharing with third parties and data retention periods.

Notably, for the most part, there is a generalised lack of implementation of simple compliance steps that concretely undermines users' understanding of how their data are processed. These include:

— **Detailed information about data sharing with third parties**. Third parties may include operators or even joint controllers or independent controllers with whom personal data are shared. In some cases, the information provided includes general categories of these third parties, but does not specify purposes, scope, period, contact information, etc. In other cases, the Policies only mention that such sharing may happen. This is crucial for informational self-determination, since being able to have a clear picture of one's personal data's current controllers, as well as the types of processing agents involved in the data flows required to achieve the controllers' processing goals, is a necessary condition for exercising self-determination.
— **Lack of information about international data transfers**. There is a paucity of concrete indications of paucity of concrete indications regarding the destination jurisdictions of personal data flows. This may derive from the lack of specific regulation, since the ANPD only recently drafted its Resolution for International Data Transfers, still not fully applicable. However, the principles and rules of information and transparency regarding international data flows also apply, so the total lack of clarification regarding transfer of personal data represents an insufficient adequacy effort.
— **Purposes, legal bases and rights of the data subject**. Although they constitute an essential part of the LGPD, in some cases, there has not been sufficient transparency regarding the purposes of data processing, legal bases and data subjects' rights. These are fundamental components of a data protection governance framework,

as well as the initial steps in designing a data protection programme, thus, the lack of clarity in Policies regarding these elements, whether due to silence or incompleteness, may indicate potential issues with the platform's governance programmes.

— **Appointment of an *Encarregado* — DPO**. The LGPD and ANPD guidelines are explicit in determining that the identity and contact information of the *Encarregado* must be public. The mere provision of multiple contact channels for a DPO or rights exercises is not a substitute for this obligation.

— **Clarity regarding automated decision-making**. While various forms of uniquely automated decision-making are routinely applied across the studied platforms, and there are somewhat vague mentions of their existence in the Policies reviewed, the efforts of platforms to explain these techniques and how personal data are used in ADM have been insufficient. In some cases, the use of automated decisions was suggested from descriptions of specific aspects of the services, but not explicitly named and explained, nor linked to specific data subject rights and how to exercise them.

The fact that these services operate in Brazil despite presenting both self-evident and more complex forms of non-compliance with the LGPD raises questions about the consequences of these findings and the broader regulatory scenario. Consequences for users of the identified issues amount to concrete challenges to the exercise of their rights as data subjects.

Basic breaches of the duty of information by platforms, such as insufficient information about the *Encarregado*,[45] or about personal data

[45] In fact, the ANPD has recently started inspection processes regarding 20 companies operating in Brazil for lack of clear indication of an *Encarregado*. Among these, X and Telegram can be found. The importance of an *Encarregado* is explained by Fabricio Lopes, General Coordinator of Inspection of the ANPD, as "[t]he absence of an *Encarregado* or an effective communication channel prevents data subjects from exercising their rights and compromises transparency in the processing of personal information. This scenario harms both the data subjects and the performance of the ANPD, which depends on this dialogue to ensure compliance with the LGPD" (ANPD, 2024c)"container-title":"Autoridade Nacional de Proteção de Dados","language":"pt-br","title":"ANPD fiscaliza 20 empresas por falta de

sharing practices and international transfers, impair the data subject's knowledge, and consequently their effective control over their data — thus, impairing informational self-determination. The same can be said about insufficient information on data processing purposes and legal bases, which may also indicate possible gaps in the controller's own adequacy efforts beyond the public-facing privacy policy. Finally, issues involved in clarity of presenting and explaining user rights and particularly complex data protection and technical functions, such as AI/ML and ADM, deepen the informational imbalance between controller and data subject. In all, users enjoy a reduced sphere of rights when information, transparency and explainability are affected.

In terms of the broader regulatory scenario, non-compliance with the LGPD persists in Brazil due to a combination of socio-economic and regulatory factors. Enforcement mechanisms[46] have seen limited use, and limited resources at the disposal of the ANPD, the Brazilian Data Protection Authority, contribute to insufficient oversight. Further attention to the role of the *Encarregado*, international data flows, data security practices, and automated data processing, among others, seem both necessary and urgent.

Importantly, the ANPD has applied only six sanctions through administrative processes, although it has acted more frequently through other non-coercive means (i.e., oversight, monitoring and recommendations, etc.) (ANPD, 2024a). This included a joint action concerning WhatsApp's privacy policy in Brazil, which may have influenced

---

Encarregado e canal de comunicação adequado","URL":"https://www.gov.br/anpd/pt-br/assuntos/noticias/anpd-fiscaliza-20-empresas-por-falta-de-encarregado-e-canal-de-comunicacao","author":[{"literal":"ANPD"}],"accessed":{"date-parts":[["2024",12,16]]},"issued":{"date-parts":[["2024",12,13]]}}],"schema":"https://github.com/citation-style-language/schema/raw/master/csl-citation.json"} .

[46] Under Brazil's LGPD, administrative sanctions imposed by the national authority include a formal warning with a deadline for corrective actions; monetary penalties consisting of a 'simple' fine of up to 2 % of the private-sector entity's gross revenue (capped at R$ 50 million per infraction) and a daily fine that respects the same overall limit; mandatory publicization of the confirmed violation; technical measures such as blocking the offending personal data until it is regularised and outright elimination of those data; and operational restrictions ranging from a partial suspension of the affected database for up to six months (renewable once) to a suspension of the entire data-processing activity for the same period, as well as a possible partial or total prohibition on any related data-processing activities (art. 52, LGPD).

the messaging service's adequacy efforts (although some basic faults remain according to our analysis) (ANPD, 2021a, 2021b). Addressing these issues requires a multifaceted approach, including bolstering the ANPD's capacity, enhancing public awareness campaigns, and incentivizing compliance through a combination of the Authority's sanctioning and educational capacities. In this perspective, we hope that this study provides useful material to support the activity of the ANPD.

4. Conclusion

This article presented some preliminary results and analyses that pave new paths for future research, while also uncovering important findings regarding the efforts (or lack thereof) of messaging platforms to comply with LGPD requirements. The examination of these platforms' Privacy Policies revealed that there are significant gaps in the fundamental elements of transparency regarding data processing. Surprisingly, many of the basic information expected in LGPD compliance were not adequately covered by the platforms' Policies, demonstrating a lack of thoroughness in addressing core legal obligations. This shortfall is compounded by an insufficiency in transparency and explainability in how these platforms describe the widespread use of AI and ML tools to automate data processing.

More broadly, we assert that current approaches to informing users and data subjects tend to fall short — they often stop at merely offering generic statements about data collection, without fully connecting the dots between the technical functioning of the services and the required legal safeguards under the LGPD.

As we have emphasised, the act of adequately informing data subjects is an essential feature of the Brazilian data protection framework. It should be more than the mere presentation of basic — and sometimes incomplete — information; it should rather constitute a meaningful explanation that relates abstract concepts to the concrete instructions allowing the data subjects to understand fully the purposes of data processing and how to enjoy their rights. This means platforms need to provide clear, accessible, and context-specific information that demystifies how personal data are collected, used, stored, and shared, as well as how users can exercise their rights. Without this, data subjects remain in the

dark about the actual impact of the utilized services, reducing the effectiveness of LGPD's protections in practice.

From a broader analytical perspective, one of the fundamental questions raised by this study concerns the very purpose and nature of the Privacy Policies published by these platforms. Clearly, the investigation unveiled contrasting approaches towards these documents, suggesting that the "Policy" may mean very different things to different platforms. On one hand, some platforms opt for excessive simplicity in their policies, which may superficially satisfy legal requirements but effectively diminish the importance of these instruments. This raises a critical inquiry: Does this simplification reflect a deliberate strategy to assign less weight to legal instruments relative to the underlying technical architectures? Such an approach risks side-lining the core principles embedded in the law, undercutting users' granular understanding and effective control over their data.

On the other hand, certain platforms present policies laden with verbose, repetitive, often confusing and convoluted language, but without indicating concrete means to fully enable the exercise of LGPD rights. This complexity might originate either from a genuine intention to be comprehensive or from a strategic design choice aimed at creating opacity through confusion. The consequences of such design are profound: instead of enlightening users, the dense, bureaucratic, or jargon-heavy text serves as a barrier to understanding, effectively undermining the transparency and accountability goals of the LGPD. This duality — between overly simplistic and excessively complex policies — points to an important area of tension in the current compliance landscape.

Moreover, it is important to interrogate the design language and underlying intent behind these policies. How are these documents constructed, both in terms of language and structure? These policies do not exist in a vacuum; rather, they are 'living artifacts' that hold meaning in their ongoing interactions with the various actors involved in the data-processing ecosystem — ranging from regulators and users to developers and corporate stakeholders. A meticulous study of these documents as autonomous objects of analysis can unearth the latent interests that shape them and expose the underlying reasons for the deficiencies observed.

This exploratory study thus contributes to the ongoing discourse by shedding light on the multi-dimensional nature of data protection compliance in Brazil's messaging app sector. Rather than merely cataloguing compliance or non-compliance, it seeks to understand the subtler dynamics behind policy formulation and presentation, highlighting why basic LGPD principles may remain obscured and ineffectively communicated.

Looking forward, this investigation opens several promising avenues for further research. Future studies could delve deeper into the socio-technical interplay between policymaking, technological design, and legal requirements, especially as regards the implications of embedding generative AI to improve the set of functions offered by the apps. Additionally, empirical research involving user perceptions and experiences could reveal how differently the presentation of these documents (and information) actually impacts users' comprehension and trust. Such insights would be invaluable for regulators, developers, and advocacy groups striving to enhance the transparency, accountability, and ultimately, the efficacy of data protection under the LGPD.

In conclusion, ensuring that Privacy Policies (or Notices) are both legally robust and meaningfully communicative is not simply a matter of checking regulatory boxes; it is a vital component of fostering trust and protecting fundamental rights. Messaging platforms play a crucial role in mediating this relationship, and their policies must evolve beyond superficial compliance toward genuine transparency and engagement with data subjects. Only then can the LGPD be fully implemented in practice, establishing a data protection culture that is not only aspirational but also effective.

5. References

ANPD. (2021a, May 7). *Cade, MPF, ANPD e Senacon recomendam que WhatsApp adie entrada em vigor da nova política de privacidade*. Autoridade Nacional de Proteção de Dados. https://www.gov.br/anpd/pt-br/assuntos/noticias/cade-mpf-anpd-e-senacon-recomendam-que-whatsapp-adie-entrada-em-vigor-da-nova-politica-de-privacidade

ANPD. (2021b, May 14). *ANPD divulga orientações aos usuários sobre a nova política de privacidade do Whatsapp*. Autoridade Nacional de

Proteção de Dados. https://www.gov.br/anpd/pt-br/assuntos/noticias/a-nova-politica-de-privacidade-do-whatsapp

ANPD. (2024a). *Balanço: 4 anos*. https://www.gov.br/anpd/pt-br/documentos-e-publicacoes/anpd-balanco-4-anos.pdf

ANPD. (2024b, July 16). *Regulamento sobre a atuação do encarregado pelo tratamento de dados pessoais (Resolução CD/ANPD No 18)*. https://www.in.gov.br/web/dou

ANPD. (2024c, December 13). *ANPD fiscaliza 20 empresas por falta de Encarregado e canal de comunicação adequado*. Autoridade Nacional de Proteção de Dados. https://www.gov.br/anpd/pt-br/assuntos/noticias/anpd-fiscaliza-20-empresas-por-falta-de-encarregado-e-canal-de-comunicacao

Apacible-Bernardo, A., & Fischer, L. (2024, March 19). Identifying global privacy laws, relevant DPAs. *IAPP*. https://iapp.org/news/a/identifying-global-privacy-laws-relevant-dpas

Article 29 Working Party. (2018). *Guidelines on transparency under Regulation 2016/679* (No. WP260 rev.01). European Commission. https://www.edpb.europa.eu/system/files/2023-09/wp260rev01_en.pdf

Bioni, B. R. (2019). *Proteção de dados pessoais: A função e os limites do consentimento* (1a). Forense.

Brasil. (2018, August 14). *Lei Geral de Proteção de Dados (13.709/2018)*. https://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/L13709compilado.htm

CJF. (2022). *Enunciado n. 689*. Consulta de Enunciados - IX Jornada de Direito Civil. https://www.cjf.jus.br/enunciados/enunciado/1828

DataGuidance & B/Luz Adv. (2019). *Comparing privacy laws: GDPR v. LGPD*. https://www.dataguidance.com/sites/default/files/gdpr_lgpd_report.pdf

de Teffé, C. S. (2022). *Dados pessoais sensíveis: Qualificação, tratamento e boas práticas*. Editora Foco. https://books.google.com/books?hl=en&lr=&id=atCLEAAAQBAJ&oi=fnd&pg=PT4&dq=TEFF%C3%89,+Chiara+Spadaccini+de.+Dados+pessoais+sens%C3%ADveis.+S%C3%A3o+Paulo:+Editora+Foco,+2022,+p.45&ots=7uURfpbjHa&sig=p4uQS_3CbqPY85e3yL3awpqRxVQ

Doneda, D. (2019). *Da privacidade à proteção de dados pessoais* (2nd ed.). Thomson Reuters Brasil.

Greenleaf, G. (2023). Global Data Privacy Laws 2023: International Standards Stall, but UK Disrupts. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.4530145

Kemp, S. (2024, February 23). *Digital 2024: Brazil*. DataReportal – Global Digital Insights. https://datareportal.com/reports/digital-2024-brazil

Lopes, A. K., Vargas, A. G., Lopes, F., Maiolino, I., Carvalho, L. B. de, & Moraes, T. (2022). *Guia orientativo para definições dos agentes de tratamento de dados pessoais e do Encarregado*. ANPD.

Miragem, B. (2019). A Lei Geral de Proteção de Dados (lei 13.709/2018) e o direito do consumidor. *Revista Dos Tribunais*, *1009*(DTR\2019\40668). https://brunomiragem.com.br/wp-content/uploads/2020/06/002-LGPD-e--o-direito-do-consumidor.pdf

Moxie0. (2017, September 26). Technology preview: Private contact discovery for Signal. *Signal Messenger*. https://signal.org/blog/private-contact-discovery/

Paiva, F. (2024). *Panorama: Mensageria no Brasil*. MobileTime/Opinion Box. https://static.poder360.com.br/2024/03/Panorama-Mensageria-MAR-24.pdf

Teixeira, J. P. F. (2021). Legítimo Interesse e os Dados Tornados Públicos e de Acesso Público. In R. de Oliveira & M. Cots (Eds.), *O Legítimo Interesse e a LGPDP* (2nd ed.). Thomson Reuters Brasil. https://bdjur.stj.jus.br/jspui/bitstream/2011/151819/legitimo_interesse_lgpdp_oliveira_2.ed.pdf

WhatsApp. (2023, August 2). *Política de Privacidade Suplementar dos Canais do WhatsApp*. WhatsApp.com. https://www.whatsapp.com/legal/channels-privacy-policy?lang=pt_BR

WhatsApp. (2024, October 8). *Por que e como tratamos seus dados*. WhatsApp.com. https://www.whatsapp.com/legal/brazil-privacy-notice/why-and-how-we-process-data

Freedom of Expression and its Limits in Brazil: A Historical Legislative Overview Between 1964 and 2021

FERNANDA CARVALHO DIAS DE OLIVEIRA SILVA*
NICOLE DE BARROS MOREIRA REIS**

Abstract: This article provides a comprehensive historical and legislative analysis of freedom of expression in Brazil, tracing its evolution from 1964 to 2021. By examining the House of Representatives' Diaries and infraconstitutional legislation, the study reveals how Brazil's legal framework has continually balanced the protection of free speech with the imposition of necessary limitations, reflecting the country's shifting political regimes and technological advancements. The research highlights the transition from overt state censorship during the military dictatorship to the democratic guarantees enshrined in the 1988 Constitution of the Republic and explores how contemporary challenges — such as digital disinformation and judicial intervention — have reshaped the boundaries of expression. Recent events, including high-profile disputes involving digital platforms and landmark judicial decisions, illustrate the ongoing tension between platform autonomy and regulatory oversight.

---

* Contact: fe.cdos@gmail.com. Associate in the Disputes and Investigations team at Slaughter and May. Moderator at arbtech.io and member of the Oxford Artificial Intelligence Society. Holds a BCL/MJur degree from the University of Oxford (2025), a Postgraduate Certificate in Data, Artificial Intelligence and High Legal Performance from Pontifícia Universidade Católica do Paraná (2024), and a Bachelor of Laws from the University of São Paulo (2020). Awarded the Law Faculty Prize for the best performance in Law and Computer Science at the University of Oxford (2025). Solicitor of the Senior Courts of England and Wales and member of the Brazilian Bar Association (OAB-SP).

** Partner at Mattos Filho. Advises in the areas of judicial and administrative litigation and arbitration, with a particular focus on technology and public law. She represents local and multinational companies in litigation involving infrastructure, product liability, data protection, internet and entertainment. Holds a Master of Contract Law (LL.M.) from Instituto de Ensino e Pesquisa (Insper-SP), and a Bachelor of Laws from Pontifícia Universidade Católica de São Paulo (PUC-SP).

Notably, the Supreme Federal Court's 2025 ruling on Article 19 of the Civil Framework for the Internet marks a significant shift towards greater platform accountability, requiring proactive content moderation and robust self-regulation. The article's methodology encompasses a qualitative review of approximately 6,000 legislative records and 33 pieces of legislation, focusing on periods of heightened censorship, democratic transition, and the digital era. The findings demonstrate that limits on freedom of expression have long been considered compatible with democratic values in Brazil, with restrictions evolving to address new societal risks. This historical perspective offers essential insights for understanding current debates and informs future policy considerations in Brazil's dynamic digital landscape.

Keywords: Digital platforms, Content moderation, Censorship, Freedom of expression, Legal history

1. Introduction

This study charts the legislative journey of freedom of expression in Brazil from 1964 to 2021, revealing a dynamic evolution shaped by shifting political regimes and rapid technological change. By analysing the House of Representatives' Diaries (Diário do Congresso Nacional) and infraconstitutional legislation, it concludes that Brazilian legal framework and legislative history highlight that certain limits on freedom of expression have long been considered compatible with democratic values. The ongoing tension between safeguarding free speech and imposing necessary limitations has defined Brazil's legal and political landscape — from the overt state censorship of the military regime to the democratic guarantees enshrined in the 1988 Constitution. This historical perspective provides a useful lens for understanding the contemporary challenges facing freedom of expression in Brazil, especially as new digital realities and judicial interventions continue to reshape the boundaries of speech.

This ongoing evolution is vividly illustrated by events and trends in Brazil's digital public sphere over the past few years. On 6 January 2022, the hashtag #TwitterApoiaFakeNews (TwitterSupportsFakeNews) gained momentum on Twitter (now X) in Brazil. Users expressed their dissatisfaction with the prevalence of anti-vaccination rhetoric and

COVID-19 misinformation. In response, the advocacy group Sleeping Giants Brasil posted the following message:

> *@TwitterBrasil must heed the concerns of its users, who refuse to tolerate the rampant spread of disinformation. We need a clear stance and a reporting mechanism for falsehoods, akin to what is already available in the United States. Use #TwitterApoiaFakeNews and help intensify the pressure.*[1]

This call to action referenced a tool designed to report posts containing misinformation about the COVID-19 pandemic. Twitter confirmed to the Brazilian news portal *G1* that such a feature had been tested since August 2021 in jurisdictions such as the United States, South Korea, and Australia.[2]

A few days later, on 13 January 2022, the suspension of the Twitter account of prominent Brazilian businessman Luciano Hang due to breaches of the platform's policies sparked intense debate. Hang claimed that the ban followed his posting of content by Dr José Augusto Nasser, a right-wing physician who commented on child vaccination policies and discussions in the Brazilian Senate concerning the introduction of a health passport.[3] Hang maintained that this suspension was an affront to his constitutionally protected freedom of expression and opinion.

During the tenure of former President Jair Bolsonaro in Brazil, there were significant efforts to limit content moderation by digital platforms, including proposed amendments and vetoes related to legislation such as Bill No 2330/2023, which aimed to curtail the discretion of these platforms in moderating content.[4]

The dispute involving Elon Musk, owner of X, and Justice Alexandre de Moraes of Brazil's Supreme Federal Court (Supremo Tribunal

---

[1] Sleeping Giants Brasil, Tweet (6 January 2022) https://twitter.com/sleepgiantsbra accessed December 9, 2024.

[2] G1, 'Twitter testa ferramenta para denunciar fake news nos EUA, Coreia do Sul e Austrália' (G1, 17 January 2022) https://g1.globo.com/ accessed on December 9, 2024.

[3] Daniel Porto, Gisele Alecrim and Henrique Andrade, 'Twitter suspende conta de Luciano Hang, dono da Havan' TecMundo (13 January 2022) https://www.tecmundo.com.br/ accessed on December 9, 2024.

[4] On attempts to limit the discretion of platforms to remove content, see Projeto de Lei No 2330/2023 (Brazil) https://www.camara.leg.br/ accessed on December 9, 2024.

Federal, STF), illustrates the ongoing tension between platform autonomy and judicial intervention. Between August and September 2024, Justice Moraes imposed a general block on X in Brazil, in addition to millionaire fines against it on the grounds that the platform would have willfully and persistently circumvented the block through an update that used cloud services offered by third parties, such as Cloudflare, Fastly and Edgeuno, allowing some Brazilian users to illegally access X without the need for a VPN.[5]

These are just some of the starring cases of litigation involving content moderation by digital platforms in Brazil. Courts have dealt with claims involving individual users, companies and Big Tech. The São Paulo State Court of Justice, for instance, has presided over hundreds of cases in which users seek the reinstatement of accounts or removed content from platforms such as Facebook, YouTube, Twitter, and Instagram, judging cases based on very different kinds of legal arguments, and offering a wide range of outcomes.[6]

In 2024, the Superior Court of Justice (Superior Tribunal de Justiça, STJ) issued a landmark ruling affirming that platforms are not only permitted but indeed ought to remove content that violates their policies.[7] This decision reinforces the legitimacy of content moderation practices and underscores the evolving legal consensus that online platforms carry responsibilities commensurate with their role as gatekeepers of digital communication.

Alongside these developments, Brazil has just witnessed the review of Article 19 of the Civil Framework for the Internet (CRFI), which establishes a general liability limitation for content providers, who should only be held accountable for damages resulting from third-party content

---

[5] The Guardian, "Brazil's top court lifts Twitter ban but fines company over failure to block accounts" (19 September 2024) https://www.theguardian.com/technology/2024/sep/19/brazil-twitter-ban-fine-musk-alexandre-de-moraes accessed on December 10, 2024.

[6] For representative cases, see Tribunal de Justiça do Estado de São Paulo case law repository, accessible at https://www.tjsp.jus.br/, accessed on December 9, 2024.

[7] Carlos Affonso de Souza, "STJ decide que YouTube pode (e deve) moderar conteúdo: por que isso importa?" (UOL, 24 October 2024) https://www.uol.com.br/tilt/colunas/carlos-affonso-de-souza/2024/10/24/stj-decide-que-youtube-pode-e-deve-moderar-conteudo-por-que-isso-importa.htm accessed on December 9, 2024.

if they refuse to comply with a court order.[8] On June 26, 2025, Brazil's STF partially declared Article 19 of the CRFI unconstitutional, shifting away from a strict "judicial order only" liability rule. Now, platforms can be held liable for unlawful content or inauthentic accounts just by receiving extrajudicial notice — no court order needed. Moreover, if content has already been deemed unlawful, platforms must remove identical reposts even without a new court decision, and there's a presumption of liability for boosted content or content spread by bots unless the platform can prove they acted diligently.

This landmark decision could mean that what used to be the exception for platform liability — acting without a court order — has become the new normal. While the STF avoided calling it "strict liability," the introduction of numerous legal presumptions nudges Brazil towards an almost automatic system of accountability for internet providers. This new precedent, which also extends to email, closed meeting, and private messaging services (for interpersonal communication), now requires platforms to implement robust self-regulation, transparent reporting channels, and to have legal representation in Brazil, urging all online services to urgently adapt their content moderation practices.

Within this multi-sided context, this study aims to provide a comprehensive examination of the legislative debates concerning freedom of expression and its constraints in Brazil, especially as they inform the regulation of online content. It seeks to establish a historical and conceptual foundation for understanding freedom of expression and its shifting contours over time. As Carlos Maximiliano has observed, an appreciation of legal norms is best attained by examining their historical evolution and the changes they have undergone, thus clarifying the current role and meaning of such norms.[9]

This study focuses on the House of Representatives' Diaries and infraconstitutional legislation (excluding state and municipal laws) spanning from 1964 to 1974 (covering the early and most repressive years of the military regime), 1985 to 1991 (the transition from

---

[8] Lei No 12.965, de 23 de Abril de 2014 (Marco Civil da Internet) DOU 24.04.2014, art 19. Judicial review materials accessible at Supremo Tribunal Federal https://www.stf.jus.br/ accessed on December 9, 2024.

[9] Carlos Maximiliano, *Hermenêutica e Aplicação do Direito* (19th edn, Forense 2002) 114.

dictatorship to democracy), and 2015 to 2021 (when the debate on freedom of expression in digital environments and the spread of misinformation intensified). The research also examines all legislation from 1964 to 2021 which contains the keyword 'freedom of expression'.

It is important to note that the scope of this research extends up to 2021, a deliberate choice reflecting the period during which the primary data collection and analysis for this study were conducted. While acknowledging the rapid and significant developments in the legal and technological landscape concerning freedom of expression since then, this historical demarcation allows for a focused examination of specific legislative trends and conceptual shifts within the chosen timeframe.

The study addresses the following questions: *What concepts and interpretations of censorship have emerged in Brazil's legislative debates from 1964 to 2021, and in what ways have lawmakers approached and justified possible limitations on freedom of expression? Furthermore, to what extent have these legislative perspectives and rationales continued or evolved following the adoption of the 1988 Constitution?*

The first part of this paper analyses extracts from the House of Representatives' Diaries, presenting a concise overview of legislative debates on freedom of expression and its limitations. The subsequent section discusses infraconstitutional rules pertaining to freedom of expression, providing critical reflections on their evolution. Finally, the paper turns to the contemporary discourse around content moderation by internet application providers — informing legal and policy debates in the digital age.

2. Methodology

The methodology employed for this study involved the analysis of two primary sources: (i) editions of the House of Representatives' Diaries, the official publication of the House of Representatives, which encompasses various information, including minutes of committee meetings, plenary sessions, and proposals;[10] and (ii) infraconstitutional

[10] Please see: https://www.congressonacional.leg.br/legislacao-e-publicacoes/glossario-legislativo/-/legislativo/termo/diario_da_camara_dos_deputados_dcd. Accessed on February 4th, 2022.

legislation, excluding state and municipal laws, spanning the period from 1964 to 2021.

In reference to (i), a comprehensive examination was conducted on all editions of the House of Representatives' Diaries spanning three distinct timeframes: (1) from 1964 to 1974, encompassing the commencement of the dictatorial regime through to the conclusion of the era characterized by heightened violence and censorship; (2) from 1985 to 1991, marking the onset of the transition from military dictatorship to the democratic regime; and (3) from 2015 to 2021. This analysis specifically targeted editions that made reference to the term "freedom of expression."[11] The selection of date ranges for analysis was determined by considerations of both temporal constraints and technical feasibility, aimed at offering a comprehensive albeit partial historical perspective of legislative discussions regarding the right to freedom of expression and its boundaries in Brazil. The daily editions were scrutinized using the search platform provided by the House of Representatives, with each year constituting a distinct search.[12]

With regard to (ii), a search was carried out for infraconstitutional legislation based on a search for the term "freedom of expression" on the Planalto[13] Legislation Portal[14] and the Federal Official Gazette[15] state and municipal laws were not included.

---

[11] References to the House of Representatives' Diaries ("Diário do Congresso Nacional") throughout the article vary due to the historical change in the names of the Collection of House of Representatives, which between 1946 and 1995 were called the Diary of the National Congress (Section I, between 1953 and 1995), and from 1995 onwards were called the Diary of the House of Representatives. Available at: http://imagem.camara.leg.br/diarios.asp. Accessed on February 5th, 2022.

[12] Available at: https://imagem.camara.leg.br/pesquisa_diario_basica.asp. Accessed on December 9, 2024..

[13] Please note that 'Planalto' stands for the official residence of the Brazilian Presidency.

[14] Presidência da República (Brasil), 'Legislação' http://www4.planalto.gov.br/legislacao/ accessed on December 10, 2024.

[15] Imprensa Nacional (Brasil), 'Página Inicial' https://www.gov.br/imprensanacional/pt-br accessed on December 10, 2024.

1. Results

a. House of Representatives' Diaries

Approximately 6,000 editions of the House of Representatives' Diaries were analysed, from which 477 were selected[16] that contained

[16] They are: 15.2.1965; 7.10.1967; 14.2.1968; 19.6.1968; 26.11.1968; 13.5.1970; 28.5.1971; 10.6.1972; 3.8.1972; 5.2.1974; 7.3.1974; 9.3.1974; 19.4.1974; 7.3.1985; 15.3.1985; 23.3.1985; 13.4.1985; 12.6.1985; 27.8.1985; 1.10.1985; 4.4.1986; 22.5.1986; 13.6.1986; 26.8.1986; 4.12.1986; 19.5.1987; 30.10.1987; 25.3.1988; 16.6.1988; 20.10.1988; 8.11.1988; 18.11.1988; 25.2.1989; 28.2.1989; 1.3.1989; 16.3.1989; 23.3.1989; 1.4.1989; 5.4.1989; 19.4.1989; 21.4.1989; 28.4.1989; 12.5.1989; 30.5.1989; 7.6.1989; 23.6.1989; 29.6.1989; 29.6.1989; 24.8.1989; 14.9.1989; 15.9.1989; 29.9.1989; 20.10.1989; 25.10.1989; 20.3.1990; 28.3.1990; 24.4.1989; 8.6.1990; 19.6.1990; 10.10.1990; 25.1.1991; 9.3.1991; 28.3.1991; 2.4.1991; 3.5.1991; 28.61991; 7.8.1991; 17.8.1991; 26.9.1991; 25.10.1991; 5.12.1991; 11.12,1991; 12.12.1991; 16.12.1991; 17.12.1991; 4.2.2015; 5.2.2015; 10.2.2015; 11.2.2015; 26.2.2015; 27.2.2015; 5.3.2015; 6.3.2015; 11.3.2015; 19.3.2015; 25.3.2015; 27.3.2015; 1.4.2015; 7.4.2015; 8.4.2015; 10.10.2015; 18.4.2015; 23.4.2015; 28.4.2015; 15.5.2015; 10.6.2015; 12.6.2015; 17.6.2015; 2.7.2015; 8.7.2015; 17.7.2015; 22.8.2015; 10.9.2015; 18.9.2015; 10.10.2015; 22.10.2015; 23.10.2015; 25.11.2015; 19.12.2015; 19.02; 2016; 24.2.2016; 25.2.2016; 6.4.2016; 7.4.2016; 27.4.2016; 29.4.2016; 4.5.2016; 20.5.2016; 2.6.2016; 9.6.2016; 30.6.2016; 8.7.2016; 11.8.2016; 16.8; 2016; 17.8.2016; 20.8.2016; 1.9.2016; 17.9.2016; 26.10.2016; 19.1.2017, 3.2.2017, 17.2.2017, 22.2.2017, 10.3.2017, 14.3.2017, 16.3.2017, 22.3.2017, 23.3.2017, 24.3.2017, 31.3.2017, 1.4.2017, 7.4.2017, 13.4.2017, 19.4.2017, 19.5.2017, 25.4.2017, 27.4.2017, 1.6.2017, 2.6.2017, 3.6.2017, 7.6.2017, 10.6.2017, 13.6.2017, 15.6.2017, 27.6.2017, 28.6.2017, 29.6.2017, 5.7.2017, 6.7.2017, 14.7.2017, 15.7.2017, 2.8.2017, 3.8.2017, 22.8.2017, 25.8.2017, 13.9.2017, 14.9.2017, 20.9.2017, 21.9.2017, 22.9.2017, 27.9.2017, 5.10.2017, 7.11.2017, 4.10.2017, 10.10.2017, 18.10.2017, 19.10.2017, 20.10.2017, 24.10.2017, 25.10.2017, 20.11.2017, 23.11.2017, 28.11.2017, 1.12.2017, 8.12.2017, 21.12.2017; 22.12.2017; 6.2.2018; 7.2.2018; 8.2.2018; 20.2.2018; 28.2.2019; 1.3.2018; 7.3.2018; 8.3.2018; 14.3.2018; 21.3.2018; 27.3.2018; 4.4.2018; 5.4.2018; 12.4.2018; 19.4.2018; 3.5.2018; 11.5.2018; 15.5.2018; 17.5.2018; 24.5.2018; 25.5.2018; 29.5.2018; 30.5.2018; 5.6.2018; 20.6.2018; 26.6.2018; 12.7.2018, 2.8.2018, 8.8.2018, 25.8.2018, 12.9.2018, 19.9.2018, 20.10.2018, 24.10.2018, 31.10.2018; 7.11.2018, 9.11.2018, 13.11.2018, 14.11.2018, 21.11.2018, 22.11.2018, 24.11.2018, 28.11.2018, 29.11.2018, 1.12.2018, 5.12.2018, 7.12.2018, 11.12.2018, 12.12.2018, 13.12.2018 e 20.12.2018; 1.2.2019, 7.2.2019, 13.2.2019, 21.2.2019, 22.2.2019, 23.2.2019, 27.2.2019, 28.2.2019, 1.3.2019, 13.3.2019, 15.3.2019, 16.3.2019, 22.3.2019, 23.3.2019, 27.3.2019, 29.3.2019, 30.3.2019, 2.4;2019, 4.4.2019, 9.4.2019, 10.4.2019, 11.4.2019, 12.4.2019, 16.4.2019, 24.4.2019, 25.4.2019, 27.4.2019, 1.5.2019; 8.5.2019, 11.5.2019, 15.5.2019, 28.5.2019,

relevant passages referring to discussions on freedom of expression in different contexts. The qualitative research, the results of which will be presented below, was based on analysing the relevant passages found in the 477 editions selected.

1964–1974: Military dictatorship and censorship

In the early years of the military regime, references to freedom of expression frequently took the form of criticism by opposition Members of Parliament (MPs) against governmental repression and censorship. In a 1965 session, for example, Deputy João Herculino condemned the

---

29.5.2019, 30.5.2019, 26.6.2019, 3.7.2019, 5.7.2019, 11.7.2019, 18.7.2019, 17.8.2019, 23.8.2019, 24.8.2019, 27.8.2019, 29.8.2019, 30.8.2019, 3.9.2019, 6.9.2019, 11.9.2019, 12.9.2019; 13.9.2019, 18.9.2019, 19.9.2019, 20.9.2019, 26.9,2019, 4.10,2019, 5.102019, 8.10.2019, 9.10.2019, 12.10.2019, 16.10.2019, 17.10.2019, 18.10.2019, 22.10.2019, 23.10.2019, 25.10.2019, 30.10.2019, 31.10,2019, 1.11.2019, 2.11.2019, 6.11.2019, 9.11.2019, 12.11.2019, 13.11.2019, 21.11.2019, 29.11.2019, 3.12.2019, 4.12.2019, 5.12.2019, 12.12.2019, 13.12.2019, 18.12.2019, 19.12.2019; 21.12.2019; 4.2.2020, 6.2.2020, 7.2.2020, 12.2.2020, 19.2.2020, 21.2.2020, 3.3.2020, 6.3.2020, 11.3.2020, 13.3.2020, 8.4.2020, 17.3.2020, 24.4.2020, 29.4.2020, 15.5.2020, 27.5.2020, 28.5.2020, 29.5.2020, 3.6.2020, 4.6.2020, 5.6.2020, 6.6.2020, 10.6.2020, 17.6.2020, 18.6.2020, 19.6.2020, 26.6.2020, 1.7.2020, 2.7.2020, 8.7.2020, 11.7.2020, 15.7,2020, 16.7,2020, 21.7,2020, 23.7.2020, 30.7.2020, 5.8.2020; 6.8;2020, 11.8.2020, 12.8.2020, 14.8.2020, 19.8.2020, 28.8.2020, 2.9.2020, 22.9.2020, 25.9.2020, 30.9.2020, 2.10.2020, 7.10.2020, 28.10.2020, 29.10.2020, 30.10.2020, 5.11.2020, 20.11.2020, 24.11.2020, 4.12.2020, 8.12;2020, 10.12.2020, 11.12.2020, 12.12.2020, 17.12.2020, 19.12.2020; 23.12.2020; 1.2.2021, 3.2.2021, 9.10.2021, 10.2.2021, 12.2.2021, 20.2.2021, 23.2.2021, 24.2.2021, 25.2.2021, 26.2.2021, 27.2.2021, 3.3.2021, 4.3.2021, 6.3.2021; 9.3.2021, 10.3.2021, 11.3.2021, 18.3.2021, 17.3.2021, 24.3.2021, 27.3.2021, 31.3.2021, 1.4.2021, 6.4.2021, 13.4.2021, 14.4.2021, 15.4.2021, 17.4.2021, 20.4.2021, 21.4.2021, 24.4.2021, 27.4.2021, 30.4.2021, 5.5.2021, 7.5.2021, 12.5.2021, 13.5.2021, 14.5.2021, 21.5.2021, 22.5.2021, 26.5.2021, 10.6.2021, 22.6.2021, 17.6.2021, 25.6.2021, 1.7.2021, 8.7.2021, 10.7.2021, 13.7.2021; 15.7.2021, 17.7.2021, 5.8.2021, 7.8.2021, 18.8.2021, 20.8.2021, 21.8.2021, 25.8.2021, 26.8.2021, 1.9.2021, 2.9.2021, 4.9.2021, 9.9.2021, 10.9.2021, 11.9.2021, 16.9.2021, 18.9.2021, 1.9.2021, 2.9.2021, 4.9.2021, 9.9.2021, 10.9.2021, 11.9.2021, 16.9.2021, 18.9.2021, 14.10.2021, 15.10.2021, 20.10.2021, 28.10.2021, 29.10.2021, 9.11.2021, 13.11.2021; 17.11.2021, 18.11.2021; 19.11.2021, 25.11.2021, 3.12.2021, 10.12.2021 e 15.12.2021. Available at: https://imagem.camara.leg.br/pesquisa_diario_basica.asp. Accessed on December 9, 2024.

curtailment of freedoms and political persecution of parliamentarians.[17] Similarly, debates from 1967 reveal MPs voicing concerns that, despite claims of democratic ideals, the government's policies severely restricted freedom of speech and assembly.[18]

From 1968 onward, following the enactment of Institutional Act No 5 (AI-5), parliamentary discourse increasingly endorsed state-imposed limitations on freedom of expression to maintain public order.[19] Discussions on Decree-Law 1.077/1970, for instance, addressed the censorship of books and periodicals, highlighting the tension between alleged national security interests and the preservation of fundamental freedoms.[20]

Between 1970 and 1974, discussions about freedom of expression prevailed in the context of press freedom, political freedom and trade union freedom. For example, in the context of parliamentary discussions on Bill 86/1971, which, among other things, dealt with trade union freedom, deputy Peixoto Filho stated that the concept of freedom with responsibility constituted the norm.

Freedom of expression was also mentioned in the context of the discussion of a decree intended to regulate the general administration of the Post Office in 1972. In this context, deputy José Mandelli emphasised the importance of expanding the communications network and guaranteeing freedom of expression, especially political freedom.[21]

In the same year, deputies Dayl de Almeida, from the ARENA party, and Pacheco Chaves, from the MDB, spoke out about the concept of press freedom in Brazil, after parliamentarians raised controversy about the MDB's claims that its members' political freedoms and freedom of expression were being violated by the military regime. On that occasion, deputy Pacheco Chaves reported that newspapers were censored due to the presence of government censors in the newspaper editorial group itself, or due to the group's fear of reprisals from the State for the content published.[22]

---

[17] *Diário do Congresso Nacional* (1965) 279-280.

[18] *Diário do* Congresso *Nacional* (1967) 6404-6405.

[19] Ato Institucional No 5 (Brazil) 13 December 1968.

[20] *Diário do Congresso Nacional* (1970) 12-13.

[21] *Diário do Congresso Nacional* (1972) 705-706.

[22] *Diário do Congresso Nacional* (1972) 2410.

1985–1991: Transition to democracy and rethinking limits

With the gradual end of the military dictatorship and Brazil's return to democracy, legislative debates reflected a renewed emphasis on dismantling censorship and restoring free speech. MPs critiqued 'self-censorship' practised by media outlets under economic and political pressures, noting that while overt censorship had diminished, the fear of state reprisal lingered.[23]

This issue was addressed by PMDB deputy Luiz Henrique in 1985, a period of effective transition from the military dictatorship to the democratic regime, who described the censorship carried out by newspapers at the time as self-censorship due to economic pressure from the government media:

> "Mr Luiz Henrique: (...) I was a Deputy in this House during the AI-5, during the lamentable and obscure time when newspapers were forced to replace articles that were cut off in the newsroom, removed from the machines by agents of the Federal Censorship and replaced with odes to Camões or photographs of flowers (...) *We are living through censorship and self-censorship. Self-censorship stems from the economic pressure of the government media, from the fear of a broadcasting concessionaire of losing the concession or suffering other sanctions. Yes, we are living through terrible times of censorship, of limiting creativity and artistic, cultural and intellectual production in Brazil* [emphasis added] (...)"[24]

The idea of self-censorship was apparently used to designate the option of newspapers to replace stories and information due to the economic pressure exerted by the State, in view of the fear of losing the broadcasting concession contract signed with the Government, or suffering other sanctions of an economic nature — a concept that in no way resembles the moderation of content by application platforms, based on their business model, as will be better explored in the Conclusion.

During these years, the concept of 'freedom with responsibility' also emerged, as lawmakers contemplated how to ensure freedom of expression without reverting to past abuses. Discussions surrounding proposed

---

[23] *Diário do Congresso Nacional* (1985) 2974.

[24] *Diário do Congresso Nacional* (1985) 2974.

legislative measures, including those concerning press laws and public officials' expression on political matters, underscored the belief that unfettered freedom of expression should not come at the expense of democratic stability and social welfare.[25]

Throughout 1985 there were discussions about the importance of freedom of expression, as well as the existence of limits to its exercise — a consequence of a process of political reopening that generated the anxieties and fears of parliamentarians in relation to the regulation of a right that had been suppressed by the State, not only in practice, but also in the normative field, such as Art. 9 of AI-5, which authorised the Presidency of the Republic to impose press censorship if "necessary for the defence of the Revolution", as well as Decree-Law 1077/1970, which enabled "the moral censorship of books and recreational magazines, but not the political censorship of news or information".

In this context, when giving its opinion on Bill 267/83-SF, which aimed to give military personnel in the paid reserve of the Armed Forces the right to express themselves on matters of a political nature, the National Security Commission pointed out that the objective in question was in line with the right to freedom of expression, provided that such freedom did not jeopardise national security.[26]

Moreover, Congresswoman Lucia Viveiros proposed Bill 5.099/1985, which sought to amend the press law in order to impose sanctions on public servants who, even on the basis of true news or real facts, advertised in the unofficial media of municipal, state or federal governments. This bill passed the scrutiny of the Constitution and

---

[25] *For* example, see José Jorge's speech at the Chamber of Deputies session on 1 October 1985: "Mr José Jorge - This six-month period of conciliation and political pacification has evened out the onslaughts of those who want to radicalise and riot, which demonstrates the national unanimity around the ideal of democracy. Brazilian public opinion is aware of the effort that the political class has made to overcome the forms of authoritarianism that still remain in power, as real obstacles to the regime of freedom with responsibility that all Brazilians have been waiting for for so long. Yes, Mr President, ladies and gentlemen, since authoritarianism is the antithesis of political activity, democracy cannot prevail where it is installed. This has been the essence of the National Congress' struggle and is the permanent goal of every representative of the Brazilian people." Diário do Congresso Nacional (1985) 11070.

[26] *Diário do Congresso Nacional* (1986) 4264.

Justice Committees and the Communications and Information Technology Committee.[27]

Still in relation to the delimitation of the right to freedom of expression, the manifesto of the National Campaign for Women's Rights contained demands regarding the new Federal Constitution and which included, in the Education and Culture section, the provision for the duty of the State to ensure "freedom of thought and expression: freedom of production, distribution and dissemination of cultural products by the media, *provided that they do not convey prejudices and discriminatory stereotypes* [emphasis added]."[28]

The House of Representatives' Diaries also reveal that the desire to combat censorship and protect freedom of expression was almost always accompanied by statements opposing the State's desire to protect the information to which Brazilian society had access. As an illustration, the manifesto of the Movement for the End of Censorship, which included representatives from civil society, presented on 19 May 1987, stated that the Movement did not "accept that the state should be able to guard the Brazilian population, deciding on their behalf what they can see, read or listen to" being instead "convinced that only in this way will Brazilians be able to fully exercise their citizenship."[29]

Furthermore, in the years leading up to the promulgation of the Constitution, freedom of expression remained predominantly linked in legislative debates to freedom of the press and political freedoms (such as the Opposition's right to power and the right to vote). For instance, the justification for Bill 257/1987, authored by deputy Fausto Rocha, which established National Freedom of Expression Day, highlighted that freedom of press is part of a broader freedom of expression, since it was "through the media that society learns about what is happening in the country and in the world and, on the basis of the information received, exercises its inalienable right to decide [between freedom of expression and freedom of the press in the legislative sphere]."[30]

---

[27] *Diário do Congresso Nacional* (1986) 5748.

[28] *Diário do Congresso Nacional* (1986) 8166.

[29] *Diário do Congresso Nacional* (1987) 1711.

[30] *Diario do Congresso Nacional* (1987) 3312.

Pronouncements of ARENA's deputies Simão Sessim and Manoel Moreira of the PMDB on the right to freedom of expression on the eve of the publication of the Constitution also draw attention to the impartiality, commitment and independence of the press, as well as access to the truth of facts and events, as fundamental elements for ensuring freedom of expression:

> "Mr Simão Sessim — The exercise of journalistic activity, in order to acquire credibility with public opinion, *requires above all impartiality, independence and commitment only to the reader [emphasis added]. In so doing, the press places itself as the authentic guardian of freedom of expression, and therefore of democracy, and in fulfilling this important role it acquires consistency as an institution of value in the daily lives of the communities to which it belongs*. (...) Paranhos de Siqueira exemplified, in its entirety, the paradigm of the journalist: *seriousness, responsibility and respect in the process of informing*. [emphasis added]."[31]

> "Mr Manoel Moreira — Mr President, ladies and gentlemen, research shows and experience proves that teaching and journalism are concrete instruments for guaranteeing the freedom of a people. The first, if well practised, provides citizens with a path to the inexhaustible sources of knowledge; *the second, if well practised, provides people with direct access to the truth of facts, of events; in short, both represent legitimate, safe channels that lead to democracy* [emphasis added]."[32]

About a month after the publication of the Federal Constitution, which took place on October 5, 1988, in a debate on new perspectives and challenges in the light of the new constitutional order, deputy Antônio de Jesus, from the PMDB, drew attention to "the issue of the abuses that have been committed by some television stations in broadcasting programmes and commercials full of images and concepts that hurt the moral standards and assault the consciences of thousands of Brazilian families," in order to request that this issue be addressed by the

---

[31] *Diário do Congresso Nacional* (1988) 1749.

[32] *Diário do Congresso Nacional* (1988) 1765.

Chamber of Deputies on future occasions.[33] In his speech, the deputy suggested some conceptual guidelines to address the issue of the limits of freedom of expression based on a systematic interpretation of the new Constitution (art. 5, items IV and IX), as shown below:

> "Mr Antônio de Jesus (...) - *A so-called freedom that does not respect the limits imposed by the legitimate right of others — who are also free — will be as pernicious and destructive as any regime of exception* [emphasis added]. Any country that wants to democratise and develop in all aspects — economic, social and cultural — needs to find this balance, translating it into clear rules of coexistence. We have made these general and philosophical considerations in order to address an issue that, *despite having received satisfactory constitutional treatment, has not, in our view, been dealt with in practice in the best interests of society* [emphasis added]. This is the issue of the abuses that have been committed by some television stations in broadcasting programmes and commercials full of images and concepts that hurt moral standards and attack the consciences of thousands of Brazilian families. By enshrining the principles of the free expression of thought and the free manifestation of intellectual, artistic, scientific and communication activity, regardless of censorship, embodied in sections IV and IX of Article 5, the new Constitutional Charter intended to guarantee Brazilians an elementary right recognised throughout the civilised world, putting an end to the obscurantism that had reigned until then. *However, these freedoms need to be considered from the perspective of the social function they aim to fulfil*, i.e. *social development and people's well-being, and not be raised as absolute flags* [emphasis added]. *In the specific case of mass communication, the Constitution itself, in Chapter V of Title VIII, by reaffirming the principle of freedom of expression and expressly prohibiting any and all political, ideological or artistic censorship, has imposed rules to guide the correct use of vehicles that propagate information* [emphasis added], especially radio and television. (...) *It is therefore necessary for the law to urgently create*

[33] *Diário do Congresso Nacional* (1988) 3854.

> *the mechanisms for the defence of the person and the family advocated by the Constitution and for society, through its legitimate channels of expression, to repudiate the perfidy of businessmen who, brutalised by the vision of profit, have lost awareness of their duty as citizens* [emphasis added]. The media provide a public service; the exploitation of television channels and radio stations is a concession from the State. The interests of the community, which wants leisure, information and culture, must therefore prevail in their use, and not be attacked in their moral conscience."[34]

The discussion about controlling the content broadcast by television stations continued at the session on 21 February 1989. When discussing the bill sent by the Executive Branch to regulate art. 220 of the Constitution, PFL deputy Costa Ferreira also pointed out that, in his opinion, the maturing of the precept of freedom of expression would lead to a situation in which the material broadcast by broadcasters would be at the discretion of their social responsibility, under penalty of violating arts. 220, §3, item II and 221, item IV of the Constitution-which deal with guarantees for the individual and the family in the face of radio and television programmes that disrespect ethical and social values. See:

> "Mr Costa Ferreira — (...) In 1989, after a personal appeal from the Minister of Justice to those responsible for the broadcasters, there was moderate broadcasting, which we hope to improve much more, without prejudice to news, leisure and the variety of programming. The broadcasters themselves set limits on what should be shown, without being coerced by the public authorities and without being bound by draconian legal provisions, as was the case in the past. *It is possible that getting used to the precepts of freedom of expression will lead to a situation where no legal prohibition is necessary, leaving it up to the broadcasters' social responsibility to ensure the quality of the material they broadcast* [emphasis added], under penalty of violating the legal precepts mentioned below: 'Art. 220. (...) § 3 II, establishes legal means that guarantee the person and the family the possibility of defending themselves against radio and television programmes

[34] *Diário do Congresso Nacional* (1988) 3854-3855.

> or programming that contradict the provisions of art. 221. Art. 221, IV — to respect the ethical and social values of the individual and the family."[35]

Since then, discussions about the limits of the media's freedom of expression — with television stations as the protagonists — have continued to take place, especially in the face of alleged abuses related to issues such as content that is inappropriate for children and adolescents, content that exploits sex and obscenity, and content that violates 'ethical principles.' In all the legislative discussions examined, the exercise of limiting freedom of expression prevailed (content that offended against personal and family rights, for example) — not seen as absolute — without ever returning to the censorship exercised during the military dictatorship. In this sense, check out the speech by PMDB deputy Neif Jabur on 17 August 1991, in defence of the proposal to amend the Constitution to allow censorship of television stations "when abuses are proven:"

> "Speech by Mr Neif Jabur — Over the next few days, Mr President, I will be seeking to collect the signatures needed to present a proposal for an amendment to the Constitution, *whereby television stations will have their programmes censored when there is evidence of sexual abuse, deviant behaviour and other obvious signs of aggression against ethical principles* [emphasis added]. We are not considering — and it is important to make our objective clear — re-establishing censorship in Brazil, in political and ideological terms. After all, we know very well the history of political censorship practised during dictatorships, its methods, distortions and consequences. (...) Now, when the National Constituent Assembly ensured, in art. 220 of the Magna Carta promulgated in October 1988, that "the manifestation of thought, creation, expression and information, in any form, process or vehicle, will not suffer any restriction," *its purpose was only to prevent the emergence of acts imposing the application of censorship, but it was never to open up the abuses currently committed*. [emphasis added]."[36]

---

[35] *Diário do Congresso Nacional* (1989) 2425.

[36] *Diário do Congresso Nacional* (1991) 14150-14151.

Also in relation to the control of the content broadcast by telecoms stations, it is worth highlighting the discussions, from 1989 onwards, to the effect that freedom of expression should be accompanied by a commitment on the part of journalistic organisations to the right to information and the truth. See the examples below:

> "Mr Antônio Salim Curiati — Article 220 of the Federal Constitution states that the expression of thought, creation, expression and information, in any form, process or vehicle, shall not be subject to any restriction... *The freedom of expression guaranteed by the Constitution must be accompanied by a commitment to the public and to the news broadcast in the media* [emphasis added]."[37]

> "Mr Edivaldo Holanda — (...) The free circulation of information, the right to inform, the modulation and even the sound in the treatment of facts and news are the responsibility of companies in the field.... *Freedom of information is linked to freedom of thought* (...) *Freedom of expression and the right to information must coexist. Freedom of expression is the right of those who use it, but the right to information reaches and encompasses the public to whom it is addressed* [emphasis added]."[38]

2015-2021: Digital era and new debates

Regarding the diaries analysed between 2015 and 2021, discussions about the regulation of television stations prevailed over the first twenty-five years after the promulgation of the Federal Constitution. However, in the more recent period, debates on freedom of expression shifted focus from traditional media to digital platforms. While early discussions still addressed television and radio programming ethics, a growing concern emerged regarding disinformation and online content moderation.

This was first expressed by the discussion of the CRFI, which presented a new scenario in which the concept of freedom in the digital space was linked not to the absence of laws, but to the guarantee and

[37] *Diário do Congresso Nacional* (1989) 1064.

[38] *Diário do Congresso Nacional* (1990) 2144

preservation of freedoms enabled by technology and the development of the Internet.[39]

However, far from ending the concerns regarding the regulation of freedom of expression in cyberspace in Brazil, the CRFI seems to have only introduced new wave concerns regarding digital platforms, as seen in the post-2014 diaries. From March 2015 onwards, there has been widespread discussion about the exercise of the profession of journalist, and the right of non-journalists to have access to the newsrooms of media outlets — a time when it was also possible to identify a growth in the debate about the risks of disinformation in Brazilian society. An illustration of that are statements made by deputies on the proposed amendment to the Constitution 206/2012, which would make it compulsory to have a degree in Social Communication in order to work as a journalist.[40] [41]

In the following years, the prevailing debates dealt with the limits of freedom of expression from the perspective of conflicts between fundamental rights: for example, between freedom of expression and respect for religious freedom or the right to be forgotten. These tensions often arose when expressive acts were perceived to infringe upon deeply held beliefs, personal reputation, or the desire for past information to remain private, particularly in an age where digital footprints are indelible. In addition, there were discussions on the right to demonstrate as part of the right to freedom of expression, acknowledging the importance of public assembly as a form of collective speech. Debates also covered the right of fans of sporting events to freedom of expression, such as displaying banners or chanting slogans,[42] and notably, the right of funk artists to disseminate songs known as "proibidões" — music often characterized by explicit or controversial lyrics that challenged social norms, prompting discussions on the boundaries of artistic freedom and cultural expression.[43]

---

[39] Instituto de Tecnologia & Sociedade do Rio (ITS Rio), *O Marco Civil da Internet: da Construção à Aplicação* (2017) https://itsrio.org/wp-content/uploads/2017/02/marco_civil_construcao_aplicacao.pdf accessed on December 10, 2024.

[40] *Diário do Congresso Nacional* (2015) 187.

[41] *Diário do Congresso Nacional* (2015) 265.

[42] *Diário do Congresso Nacional* (2018) 63

[43] *Diário do Congresso Nacional* (2018) 208.

In relation to internet application providers, discussions about freedom of expression from 2015 onwards, when the CRFI was already in force, began to focus not only on cyber-crimes, such as libel and slander against federal deputies, but also on the clash between the idea of absolute freedom of expression and combating the spread of *fake news*. On May 17th, 2018, for example, deputy Irmão Lazaro, of the PSC, stated that the main challenge for the legislative branch, the press and the electoral justice system was to prevent the spread of *fake news* "without curbing freedom of expression and with the speed needed in an election campaign," given the speed at which fake news circulates on the main social networks.[44]

Regarding latest years (2020-2022), in parallel to the debate about moderating content on social networks and combating *fake news*, there were also discussions about combating various forms of threats to press freedom, such as threats and attacks on journalists on social networks, and even recent attacks on journalists perpetrated by former President of Brazil, Jair Bolsonaro.[45]

b. Legislative landscape

Throughout the research conducted, we found 33 pieces of legislation containing the term "freedom of expression" — including laws, decree-laws and legislative decrees — namely: Decree 10.222/2020; Decree 977/1993; Decree 10.088/2019; Decree 9.637/2018; Decree 9.603/2018; Decree 9.522/2018; Law 13.709/2018; Decree 9.306/2018; Decree 9.202/2017; Decree 8.845/2016; Decree 8.827/2016; Decree 8.799/2016; Law 13.284/2016; Decree 8.707/2016; Decree 8.521/2015; Law 12.965/2014; Law 12.852/2013; Law 12.663/2012; Law 12.485/2011; Decree 7.518/2011; Law 12.343/2010; Decree 6.949/2009; Decree 6.728/2009; Legislative Decree 186/2008; Decree 6.177/2007; Decree 592/1992; Law

---

[44] *Diário do Congresso Nacional* (2018) 208.

[45] In this regard, see Bill 2914/2020, authored by PDT deputy Paulo Ramos, which aimed to increase the penalty for homicide and bodily injury offences for victims who are journalists and broadcasters in the course of their duties. Available at: https://www.camara.leg.br/proposicoesWeb/prop_mostrarintegra?codteor=1897857&filename=PL+2914/2020. Accessed on December 9, 2024.

8.313/1991; Decree 99.710/1990; Decree-law 972/1969; Law 5.250/1967; Decree 55.929/1965; Decree 27.784/1950; and Decree 25.696/1948.

Analysing the content of the legislation found, 5 expressly provided for some kind of limitation or caveat to the exercise of freedom of expression: Decree 55.929/1965, Decree 99.710/1990, Decree 592/1992, Law 12.663/2012 and Law 13.284/2016.

These provisions, despite guaranteeing the right to freedom of expression, authorise the government to curb certain acts which, although being a form of expression, generate risks to the rights of third parties, society or the State.

Decree 55.929/1965, Decree 99.710/1990 and Decree 592/1992 deal with international conventions that guarantee the right to freedom of expression, but allow it to be restricted in given circumstances:

> (i) Decree 55.929/1965, which promulgated the Convention on Territorial Asylum, signed on 28 March 1954, guarantees asylees the right to freedom of expression of thought recognised by the inhabitants of the state which granted the asylum, which cannot be the subject of a complaint by another state, except in the case where the manifestations represent incitement to the use of force or violence against the government of the complaining state;[46]
>
> (ii) Decree 99.710/1990, which promulgated the Convention on the Rights of the Child, guarantees children subject to the jurisdiction of States Parties the right to freedom of expression (art. 13.1), authorising its restriction, subject to legal provision, for the respect of the rights of others and for the protection of

[46] "Art. VII. Freedom of expression of thought, which domestic law recognises for all inhabitants of a state, may not be the subject of a complaint by another state based on concepts publicly expressed against it or its government by asylees or refugees, *except where such concepts constitute systematic propaganda inciting the use of force or violence against the government of the complaining state* [emphasis added]."

national security, public order, or public health and morals (art. 13.2);[47] [48] and

(iii) Decree 592/1992, which enacted the International Covenant on Civil and Political Rights, in its articles 19.1 and 19.2, guarantees everyone the right to freedom of expression, but provides that this right may be subject to restrictions provided by law to ensure respect for the rights of third parties and to protect national security, public order, health or morals (article 19.3).[49]

Laws 12.663/2012[50] and 13.284/2016[51] provide for measures relating to sporting events held in Brazil, namely the 2013 FIFA Confederations Cup, the 2014 FIFA World Cup and World Youth Day (Law

[47] "Article 13 (1) The child shall have the right to freedom of expression. This right shall include freedom to seek, receive and impart information and ideas of all kinds, regardless of frontiers, either orally, in writing or in print, through the media of the arts or by any other means chosen by the child. (2) *The exercise of this right may be subject to certain restrictions, which shall be solely those provided for by law and deemed necessary: a) for respect for the rights or reputations of others, or b) for the protection of national security or public order, or for the protection of public health or morals* [emphasis added]."

[48] As the Federal Supreme Court clarified in 2016, this is an expression of the legislator's concern to harmonise the right to broad freedom of expression and the duty to protect children morally. See: STF, ADI 2.404/DF, rel. Min. Dias Toffoli, j. 31.8.2016.

[49] "Art. 19 (1) No one shall be subjected to any form of harassment for his opinions. (2) Everyone shall have the right to freedom of expression; this right shall include freedom to seek, receive and impart information and ideas of all kinds, regardless of frontiers, either orally, in writing, in print, in the form of art, or through any other media of his choice. *(3) The exercise of the right provided for in paragraph 2 of this Article shall entail special duties and responsibilities. Consequently, it may be subject to certain restrictions, which must, however, be expressly provided for by law and which are necessary in order to: a) ensure respect for the rights and reputation of other persons; b) protect national security, public order, health or morals* [emphasis added]."

[50] "Provides for measures relating to the FIFA Confederations Cup 2013, the FIFA World Cup 2014 and the World Youth Day - 2013, which will be held in Brazil; amends Laws Nos. 6.815, of 19 August 1980, and 10.671, of 15 May 2003; and establishes the granting of bonuses and special monthly allowances to the players of the 1958, 1962 and 1970 World Cup-winning teams."

[51] "Provides for measures relating to the 2016 Olympic and Paralympic Games and related events, which will be held in Brazil; and amends Law No. 12.035, of 1 October 2009, which 'institutes the Olympic Act, within the scope of the federal public administration', and Law No. 12.780, of 9 January 2013, which 'provides for tax measures relating to the holding, in Brazil, of the 2016 Olympic Games and the 2016 Paralympic Games."

12.663/2012), and the 2016 Olympic and Paralympic Games related events. Art. 28 of both laws, which contains similar wording, determines the conditions for access and permanence in the official venues of the events.[52]

Unlike the regulations above, which concern the promotion and regulation of the right to freedom of expression, these laws seem more focused on limiting the freedom of expression of event participants to prevent conduct considered harmful. They prevent them, for example, from carrying symbols of a discriminatory nature (item IV), chanting discriminatory slurs (item V) and inciting or committing acts of violence of any kind (item VIII), among others, under penalty of not being allowed to enter or being immediately removed from the premises (§2). It is only then that the provisions in question reserve for the recipients of the laws "the constitutional right to the free exercise of manifestation and full freedom of expression in defence of the dignity of the human person" (§1).

Finally, the other 28 pieces of legislation analysed concern exclusively the provision or reinforcement of the constitutional guarantee of freedom of expression or enshrine it as a principle of the law in question (as is the case, for example, of Decree 10.088/2019, which deals with the promulgation of International Labour Organisation conventions

---

[52] "Art. 28 — The following are conditions for access and stay of any person in the Official Competition Venues, among others: (...) IV — *not to carry or display posters, flags, symbols or other signs with offensive, racist, xenophobic messages or that encourage other forms of discrimination; V - not to chant discriminatory, racist or xenophobic slogans or songs* [emphasis added]; VI — not to throw objects of any kind inside the sports grounds; VII — not to carry or use fireworks or any other pyrotechnic devices or devices that produce similar effects, including instruments equipped with laser beams or similar, or that can emit them, except for a team authorised by FIFA, or a person or entity appointed by FIFA for artistic purposes; VIII — *not to incite or practise acts of violence, whatever their nature* [emphasis added]; IX — not to invade or incite the invasion, in any way, of the area restricted to competitors, press representatives, authorities or technical teams; and X — not to use flags, including bamboo poles or similar, for purposes other than festive and friendly demonstrations. § Paragraph 1 — *The constitutional right to the free exercise of demonstration and full freedom of expression in defence of the dignity of the human person is reserved. § Paragraph 2 Failure to comply with the conditions set out in this article will result in the person being unable to enter the Official Competition Venue or being immediately removed from the premises*, without prejudice to other administrative, civil or criminal sanctions. [emphasis added]."

and recommendations, and Decree 9.637/2018, which establishes the National Information Security Policy).

3. Conclusion

This study has charted the legislative journey of freedom of expression in Brazil from 1964 to 2021, revealing a dynamic evolution shaped by political shifts and technological advancements. Our analysis of the House of Representatives' Diaries and infraconstitutional legislation has illuminated the historical tensions between safeguarding free speech and imposing necessary limitations, from the overt state censorship of the military regime to the democratic guarantees enshrined in the 1988 Constitution. This historical understanding is not merely academic; it provides an essential lens through which to comprehend the contemporary challenges facing freedom of expression in Brazil.

Indeed, the past few years, particularly beyond our primary research scope of 2021, have underscored the relentless pace of this evolution. The recent landmark judgment by the STF on Article 19 of the CRFI exemplifies this ongoing transformation. By largely moving beyond the strict "judicial order only" prerequisite for platform liability, the STF has ushered in a more nuanced, risk-based approach to content moderation. This pivotal decision, compelling platforms to assume greater proactive responsibility for specific categories of harmful content — from child sexual exploitation to anti-democratic acts — reflects a critical judicial response to the scale and speed of online harms. This, alongside the active debates surrounding new legislative proposals, like Bill No. 2330/2023, concerning online disinformation and digital intermediary accountability, firmly places Brazil at the forefront of global efforts to balance free expression with public safety in the interconnected digital sphere.

It is precisely this dynamic interplay between established constitutional principles and emergent digital realities that defines Brazil's current normative framework for freedom of expression. The analysis of legislative debates in Brazil's House of Representatives from 1964 to 2021 reveals enduring concerns about freedom of expression, its scope, and its limitations. In the formulation of the 1988 Constitution, the legislator was predominantly dealing with a concept of censorship

rooted squarely in the experiences of the military dictatorship: direct, state-imposed prior restraint on media, artistic, and intellectual output. Historically, the term "censorship" was employed to describe such state-driven efforts, including ending concession contracts with broadcasters or influencing editorial boards during the military government era. Its architects meticulously crafted provisions like Articles 5 and 220 to dismantle these mechanisms and prevent their return, focusing on state non-interference and ensuring that the government could not arbitrarily vet, alter, or ban content before its public dissemination. This understanding of censorship was primarily vertical — an abuse of state power over individual expression. Notably, within this historical period, there were no legislative discussions regarding contractual or private constraints on freedom of expression, such as those potentially exercised by digital platforms today.

However, the 21st century has introduced entirely new dimensions to this discussion, fundamentally broadening what we now debate as "censorship." The advent of digital platforms has shifted the paradigm, extending the concern beyond the State's direct hand to encompass the actions of private entities — the internet application providers. When these platforms, through their terms of service, content policies, and algorithmic decisions, remove or restrict user-generated content, this introduces a complex form of "private" or "platform" censorship. This dynamic is further complicated by phenomena like the pervasive spread of disinformation and hate speech. These forms of expression, while seemingly "free," can cause profound societal harm, forcing a critical debate: when does the management of such content by platforms become a necessary protective measure against defined harms, and when does it cross the line into undue expressive curtailment?

Throughout these decades, safeguarding freedom of expression was viewed as integral to preserving other freedoms: press freedom, political liberties, trade union rights, and artistic expression. Freedom of the press received particular attention, with debates continuing after democratization about protecting journalists and ensuring accurate information. Similarly, workers' speech focused on trade union rights, and artistic freedom, especially in television and film, remained a central topic from the military era through the democratic period. Parliamentarians repeatedly emphasized concerns about indecent programming and its impact

on minors. After 1989, legislative measures restricting media freedoms were consistently framed not as censorship, but as necessary modifications of expressive rights to reconcile them with the rights of children, adolescents, and religious communities. These constraints sought to reflect the ethical and social values outlined in the Constitution, such as those contained in Article 221(IV). Importantly, no recorded debates addressed the ability of private entities to impose similar restrictions during these periods.

Moreover, the democratic reopening and the drafting of the new Constitution did not assume absolute freedom of expression. Neither the State nor private actors were explicitly barred from imposing certain limits. Before the Constitution's publication, parliamentarians stressed the importance of ensuring accurate information to enable meaningful political participation, reinforcing a strong link between freedom of expression and access to reliable facts. This connection, also noted by scholars such as Caldas and Fonteles and Conrado, suggests that the historical understanding of freedom of expression did not encompass the dissemination of disinformation. Furthermore, during the Constituent Assembly, civil society groups, including the National Campaign for Women's Rights, advocated for the safeguarding of freedom of expression, contingent upon the condition that it does not perpetuate prejudices or discriminatory stereotypes.

The legislative research has unearthed a noteworthy aspect of the Brazilian legal framework, which allows for the restriction of content moderation by application providers, particularly when such actions are taken to address abuses resulting in violations of the rights of third parties, the welfare of society, or the interests of the Brazilian State. This discovery challenges some of the assumptions put forth during contemporary content moderation debates. Furthermore, Laws 12,663/2012 and 13,284/2016, enacted for sporting events in Brazil, further illustrate these nuances. They authorize immediate removal of individuals from venues for discriminatory language or incitement of violence, measures often echoed in social network terms and conditions.

In sum, the Brazilian legal framework and legislative history highlight that certain limits on freedom of expression have long been considered compatible with democratic values. Such restrictions are not tantamount to the censorship of the military dictatorship era. Rather,

they reflect a constitutional mandate to balance free speech and other rights and interests. The "self-censorship" described by MPs in the post-dictatorship era, driven by economic pressure or fear of reprisal, also resonates with modern discussions about how platforms' policies and economic models can inadvertently shape or restrict online discourse. Accordingly, arguments likening contemporary regulations or platform moderation to past authoritarian censorship misread the Constitution's intentions. They overlook the reality that, historically, freedom of expression has always been understood as subject to reasonable constraints aimed at preserving democratic integrity, social welfare, and the protection of vulnerable groups. This historical overview, therefore, provides a vital framework for understanding this enduring democratic struggle and informs future policy considerations in an ever-evolving digital landscape.

4. References

Caldas, P. F. (1997). *Private life, freedom of the press and moral damage*. São Paulo: Saraiva.

Decree no. 8827/2016 (2016). Provides for the implementation, on national territory, of United Nations Security Council Resolution 2290 (2016) of 31 May 2016, which extends the sanctions regime imposed on South Sudan. Federal Official Gazette. Available at: http://www.planalto.gov.br/ccivil_03/_ato2015-2018/2016/decreto/d8827.htm.

Legislative Decree no. 186/2008 (2008). Approves the text of the Convention on the Rights of Persons with Disabilities and its Optional Protocol, signed in New York on 30 March 2007. Federal Official Gazette. Available at: http://www2.senado.leg.br/bdsf/handle/id/99423.

Decree no. 10.088/2019. (2019). Consolidates normative acts issued by the Federal Executive Branch that provide for the enactment of International Labour Organization (ILO) conventions and recommendations ratified by the Federative Republic of Brazil. Federal Official Gazette. Disponível em: https://www.in.gov.br/en/web/dou/-/decreto-n-10.088-de-5-de-novembro-de-2019-231356812.

Decree no. 10.222/2020. (2020). Approves the National Cyber Security Strategy. Federal Official Gazette. Disponível em: https://www.in.gov.br/en/web/dou/-/decreto-n-10.222-de-5-de-fevereiro-de-2020-241828419.

Decree no. 25.696/1948 (1948). Orders the execution of the acts signed in Montreal, on 09/10/1946, on the occasion of the twenty-ninth session of

the General Conference of the International Labour Organisation. Official Gazette of the Union. Available at: https://legislacao.presidencia.gov.br/atos/?tipo=DEC&numero=25696&ano=1948&ato=1690TV65UMrR1T133.

Decree no. 27.784/1950 (1950). Promulgates the Convention on the Privileges and Immunities of the United Nations, adopted in London on 13 February 1946 at the United Nations General Assembly. Federal Official Gazette. Available at: https://www.planalto.gov.br/ccivil_03/decreto/antigos/d27784.htm.

Decree no. 5.250/1967 (1967). Regulates freedom of expression of thought and information. Federal Official Gazette. Available at: https://www.planalto.gov.br/ccivil_03/leis/l5250.htm.

Decree no. 55.929/1965 (1965). Promulgates the Convention on Territorial Asylum. Federal Official Gazette. Available at: https://www.planalto.gov.br/ccivil_03/decreto/1950-1969/d55929.htm.

Decree no. 592/1992 (1992). International Acts. International Covenant on Civil and Political Rights. Promulgated. Official Federal Gazette. Available at: https://www.planalto.gov.br/ccivil_03/decreto/1990-1994/d0592.htm.

Decree no. 6.177/2007 (2007). Promulgates the Convention on the Protection and Promotion of the Diversity of Cultural Expressions, signed in Paris on 20 October 2005. Federal Official Gazette. Available at: https://legislacao.presidencia.gov.br/atos/?tipo=DEC&numero=6177&ano=2007&ato=7eeATUU9ENRpWT149.

Decree no. 6.728/2009 (2009). Regulates art. 29, items I, II and III, of Law no. 6.015, of 31 December 1973, which deals with public registers, and makes other provisions. Federal Official Gazette. Available at: https://www.planalto.gov.br/ccivil_03/_ato2007-2010/2009/decreto/D6728.htm.

Decree no. 6.949/2009 (2009). Promulgates the International Convention on the Rights of Persons with Disabilities and its Optional Protocol, signed in New York on 30 March 2007. Federal Official Gazette. Available at: https://legislacao.presidencia.gov.br/atos/?tipo=DEC&numero=6949&ano=2009&ato=8dec3Y61UeVpWT233.

Decree no. 7.518/2011 (2011). Provides for the implementation on national territory of Resolution 1975 (2011), adopted by the United Nations Security Council on 30 March 2011, which, among other provisions, calls on the parties involved in the post-election political crisis in Côte d'Ivoire to recognise the election of Mr Alassane Dramane Ouattara, urges Mr Laurent Gbagbo to withdraw from the political process, reiterates its firm condemnation of all violence against the civilian population in the country and establishes a sanctions regime against specified individuals. Federal Official Gazette. Available at: https://legislacao.presidencia.

gov.br/atos/?tipo=DEC&numero=7508&ano=2011&ato=16fgXUE9UM-VpWT875.

Decree no. 8.521/2015 (2015). Provides for the implementation, on national territory, of United Nations Security Council resolution 2161 (2014) of 17 June 2014, which deals with sanctions against individuals, groups, initiatives and entities of Al-Qaeda and associates. Federal Official Gazette. Available at: https://www.planalto.gov.br/ccivil_03/_ato2015-2018/2015/decreto/d8521.htm.

Decree no. 8.707/2016 (2016). Provides for the implementation on national territory of United Nations Security Council Resolution 2206 (2015) of 3 March 2015 establishing a sanctions regime on South Sudan. Federal Official Gazette. Available at: https://legislacao.presidencia.gov.br/atos/?tipo=DEC&numero=8707&ano=2016&ato=137ATVU1EeZpWTbfc.

Decree no. 8.845/2016 (2016). Provides for the implementation on national territory of United Nations Security Council Resolution 2293 (2016) of 23 June 2016 renewing the sanctions regime applicable to the Democratic Republic of Congo. Federal Official Gazette. Available at: https://www.in.gov.br/materia/-/asset_publisher/Kujrw0TZC2Mb/content/id/21910886/UCEQlTzKXPyVi6cWuD3q0ksQ.

Decree no. 8799/2016 (2016). Provides for the implementation, on national territory, of United Nations Security Council Resolution 2253 (2015) of 17 December 2015, which updates and strengthens the sanctions regime imposed by Resolution 1267 (1999) concerning the Islamic State in Iraq and the Levant and Al-Qaeda. Official Federal Gazette. Disponível em: https://www.in.gov.br/materia/-/asset_publisher/Kujrw0TZC2Mb/content/id/21288891/do1-2016-07-07-decreto-n-8-799-de-6-de-julho-de-2016-21288638.

Decree no. 9.202/2017. (2017). Provides for the implementation, on national territory, of United Nations Security Council Resolution 2368 (2017) of 20 July 2017, which updates and strengthens the sanctions regime in force against individuals and entities associated with the Islamic State in Iraq and the Levant and Al-Qaeda. Federal Official Gazette. Available at: https://legislacao.presidencia.gov.br/atos/?tipo=DEC&numero=9202&ano=2017&ato=8a5oXVU1UeZpWT578.

Decree no. 9.306/2018. (2018). Provides for the National Youth System, established by Law no. 12,852 of 5 August 2013. Federal Official Gazette. Available at: pág.1https://legislacao.presidencia.gov.br/atos/?tipo=DEC&numero=9306&ano=2018&ato=cf3cXR65UeZpWT8c0.

Decree no. 9.522/2018. (2018). Promulgates the Marrakech Treaty to facilitate access to published works for persons who are blind, visually impaired or otherwise print-disabled, signed in Marrakech on 27 June 2013.

Federal Official Gazette. Available at: https://legislacao.presidencia.gov.br/atos/?tipo=DEC&numero=9522&ano=2018&ato=a1cUTSU9UeZpWT494.

Decree no. 9.603/2018. (2018). Regulates Law no. 13.431, of 4 April 2017, which establishes the system for guaranteeing the rights of children and adolescents who are victims or witnesses of violence. Federal Official Gazette. Available at: https://legislacao.presidencia.gov.br/atos/?tipo=DEC&numero=9603&ano=2018&ato=5a7gXRE1keZpWTf1d.

Decree no. 9.637/2018. (2018). Establishes the National Information Security Policy, provides for information security governance, and amends Decree no. 2.295, of 4 August 1997, which regulates the provisions of art. 24, caput, item IX, of Law no. 8.666, of 21 June 1993, and provides for the waiver of bidding in cases that may compromise national security. Federal Official Gazette. Disponível em: https://www.in.gov.br/materia/-/asset_publisher/Kujrw0TZC2Mb/content/id/56970098/do1-2018-12-27-decreto-n-9-637-de-26-de-dezembro-de-2018-56969938.

Decree no. 977/1993. (1993). Provides for pre-school assistance for the dependents of civil servants in the direct, autarchic and foundational Federal Public Administration. Federal Official Gazette. Available at: https://www.planalto.gov.br/ccivil_03/decreto/antigos/d0977.htm.

Decree no. 99.710/1990 (1990). Promulgates the Convention on the Rights of the Child. Federal Official Gazette. Available at: https://www.planalto.gov.br/ccivil_03/decreto/1990-1994/d99710.htm.

Decree-Law 972/1969 (1969). Provides for the exercise of the profession of journalist. Federal Official Gazette. Available at: https://www.planalto.gov.br/ccivil_03/decreto-lei/del0972.htm.

Diário do Congresso Nacional (2015). 19 March 2015. Available at: http://imagem.camara.gov.br/Imagem/d/pdf/DCD00201503190004000000.PDF#page.

Diário do Congresso Nacional (2015). 25 March 2015. Available at: http://imagem.camara.gov.br/Imagem/d/pdf/DCD00201503250004400000.PDF#page.

Diário do Congresso Nacional (2018). 17 May 2018. Available at: http://imagem.camara.gov.br/Imagem/d/pdf/DCD00201805170007200000.PDF#page=

Diário do Congresso Nacional (2018). 25 August 2018. Available at: http://imagem.camara.gov.br/Imagem/d/pdf/DCD00201808250013300000.PDF#page=

Diário do Congresso Nacional (1965). 15 February 1965. Accessed at: http://imagem.camara.gov.br/Imagem/d/pdf/DCD15FEV1965.pdf#page=.

Diário do Congresso Nacional (1967). 7 October 1967. Available at: http://imagem.camara.gov.br/Imagem/d/pdf/DCD07OUT1967.pdf#page.

Diário do Congresso Nacional (1970). 13 May 1970. Available at: http://imagem.camara.gov.br/Imagem/d/pdf/DCD13MAI1970.pdf#page=.

Diário do Congresso Nacional (1972). 03 August 1972. Available at: http://imagem.camara.gov.br/Imagem/d/pdf/DCD03AGO1972.pdf#page=.

Diário do Congresso Nacional (1972). 06 May 1972. Available at: http://imagem.camara.gov.br/Imagem/d/pdf/DCD06MAI1972.pdf#page=.

Diário do Congresso Nacional (1985). 13 April 1985. Available at: http://imagem.camara.gov.br/Imagem/d/pdf/DCD13ABR1985.pdf#page.

Diário do Congresso Nacional (1986). 13 June 1986. Available at: ttp://Imagem.Camara.Gov.Br/Imagem/D/Pdf/Dcd13jun1986.Pdf#Page=.

Diário do Congresso Nacional (1986). 22 May 1986. Available at: http://imagem.camara.gov.br/Imagem/d/pdf/DCD22MAI1986.pdf#page=.

Diário do Congresso Nacional (1986). 28 August 1986. Available at: http://imagem.camara.gov.br/Imagem/d/pdf/DCD28AGO1986.pdf#page=.

Diário do Congresso Nacional (1987). 19 May 1987. Available at: http://imagem.camara.gov.br/Imagem/d/pdf/DCD19MAI1987.pdf#page.

Diário do Congresso Nacional (1987). 30 October 1987. Available at: http://imagem.camara.gov.br/Imagem/d/pdf/DCD30OUT1987.pdf#page=.

Diário do Congresso Nacional (1988). 18 November 1988. Available at: http://imagem.camara.gov.br/Imagem/d/pdf/DCD18NOV1988.pdf#page.

Diário do Congresso Nacional (1988). 13 May 1988. Available at: http://imagem.camara.gov.br/Imagem/d/pdf/DCD13MAI1988.pdf#page.

Diário do Congresso Nacional (1989). 12 April 1989. Available at: http://imagem.camara.gov.br/Imagem/d/pdf/DCD21ABR1989.pdf#page=.

Diário do Congresso Nacional (1989). 16 March 1989. Available at: http://imagem.camara.gov.br/Imagem/d/pdf/DCD16MAR1989.pdf#page=.

Diário do Congresso Nacional (1990). 28 March 1990. Available at: http://imagem.camara.gov.br/Imagem/d/pdf/DCD28MAR1990.pdf#page=.

Diário do Congresso Nacional (1991). 17 August 1991. Available at: http://imagem.camara.gov.br/Imagem/d/pdf/DCD17AGO1991.pdf#page=.

Diário do Congresso Nacional (n.d.). Available at: https://imagem.camara.leg.br/pesquisa_diario_basica.asp.

Fonteles, E. M. P. & Conrado, R. M. (2019). Truth as an immanent limit to freedom of expression: fake news and the risk to democracy (pp. 282-407). In: National Association of Prosecutors of the Republic. *Themes of the Public Prosecutor's Office: agreements in the justice system and freedom of expression*. Brasília: ANPR.

G1. (2022). *Understand why users are accusing Twitter of supporting fake news*. Available at: https://g1.globo.com/tecnologia/noticia/2022/01/05/entenda-por-que-usuarios-estao-acusando-o-twitter-de-apoiar-fake-news.ghtml.

Glossary of Legislative Terms (n.d.). Available at: https://www.congressonacional.leg.br/legislacao-e-publicacoes/glossario-legislativo/-/legislativo/termo/diario_da_camara_dos_deputados_dcd.

Instituto de Tecnologia & Sociedade do Rio (ITS Rio), *O Marco Civil da Internet: da Construção à Aplicação* (2017) https://itsrio.org/wp-content/

uploads/2017/02/marco_civil_construcao_aplicacao.pdf accessed on December 10, 2024.

Law no. 12.343/2010 (2010). Establishes the National Culture Plan - PNC, creates the National System of Cultural Information and Indicators – SNIIC and makes other provisions. Federal Official Gazette. Available at: https://legislacao.presidencia.gov.br/atos/?tipo=LEI&numero=12343&ano=2010&ato=2cdUzYq1keVpWTdd1.

Law no. 12.485/2011 (2011). Provides for conditional access audiovisual communication [...]. Federal Official Gazette. Available at: https://legislacao.presidencia.gov.br/atos/?tipo=LEI&numero=12485&ano=2011&ato=-b1aITQU1UMVpWT558.

Law no. 12.663/2012 (2012). Provides for measures relating to the 2013 FIFA Confederations Cup, the 2014 FIFA World Cup and World Youth Day - 2013 [...]. Federal Official Gazette. Available at: https://legislacao.presidencia.gov.br/atos/?tipo=LEI&numero=12663&ano=2012&ato=5a6c3YU1kMVpWT9a2.

Law no. 12.852/2013 (2013). Establishes the Youth Statute and provides for the rights of young people, the principles and guidelines of public youth policies and the National Youth System - SINAJUVE. Federal Official Gazette. Available at: https://legislacao.presidencia.gov.br/atos/?tipo=LEI&numero=12852&ano=2013&ato=560ATWU50MVpWT43d.

Law no. 12.965/2014 (2014). Establishes principles, guarantees, rights and duties for the use of the Internet in Brazil. Federal Official Gazette. Disponível em: https://www.in.gov.br/materia/-/asset_publisher/Kujrw0TZC2Mb/content/id/30054611/do1-2014-04-24-lei-n-12-965-de-23-de-abril-de-2014-30054600.

Law no. 13.284/2016 (2016). Provides for measures relating to the 2016 Olympic and Paralympic Games and related events [...]. Federal Official Gazette. Disponível em: https://www.in.gov.br/web/dou/-/lei-no-13-284-de-10-de-maio-de-2016-21174921.

Law no. 13.709/2018. (2018). General Data Protection Law. Drafted by Law 13,853 of 2019. Federal Official Gazette. Disponível em: https://www.in.gov.br/materia/-/asset_publisher/Kujrw0TZC2Mb/content/id/36849373/do1-2018-08-15-lei-no-13-709-de-14-de-agosto-de-2018-36849337.

Law 8313/1991 (1991). It re-establishes the principles of Law no. 7.505, of 2 July 1986, institutes the National Programme to Support Culture (Pronac) and makes other provisions. Federal Official Gazette. Available at: https://legislacao.presidencia.gov.br/atos/?tipo=LEI&numero=8313&ano=1991&ato=2f4ITSU9UMFpWTdd0.

Maximiliano, C. (2002). *Hermeneutics and the Application of Law*. 19th ed. Rio de Janeiro: Forense, 2002.

Porto, D., Alecrim, G. & Andrade, H. (2022). *Twitter suspends Luciano Hang's account for violating rules* (Luciano Hang press release). CNN Brasil. Available at: https://www.cnnbrasil.com.br/politica/twitter-suspende-conta-de-luciano-hang-por-violacao-de-regras/.

Bill no. 2630/2020. (2020). Establishes the Brazilian Law on Internet Freedom, Responsibility and Transparency. Available at: https://www.camara.leg.br/proposicoesWeb/fichadetramitacao?idProposicao=2256735.

Bill no. 3227/2021. (2021). Amends Law no. 12.965, of 23 April 2014, and Law no. 9.610, of 19 February 1998, to provide for the use of social networks. Available at: https://www.camara.leg.br/proposicoesWeb/prop_mostrarintegra;jsessionid=node0a0hgy949t7fj1w56w5q6f2x7248603664.node0?codteor=2076539&filename=PL+3227/2021. Accessed on 4 March 2022.

Bill no. 5.099/1985 (1985). Amends provisions of the Press Law, for the publication in all media of municipal, state and federal government propaganda. Available at: https://www.camara.leg.br/proposicoesWeb/prop_mostrarintegra;jsessionid=F49EFE36A04DED2AE641F3704BF1F280.proposicoesWebExterno1?codteor=1160253&filename=Dossie+-PL+5099/1985.

Sleeping Giants Brazil (2022). Available at: https://twitter.com/slpng_giants_pt?ref_src=twsrc%5Egoogle%7Ctwcamp%5Eserp%7Ctwgr%5Eauthor.

Twitter. (n.d.) *Rules and policies*. Available at: https://help.twitter.com/pt/rules-and-policies#twitter-rules.

A responsabilidade das plataformas pelo conhecimento (presumido): subsídios para uma compreensão do artigo 16.º, n.º 3, do Regulamento dos Serviços Digitais

INÊS NEVES*

Resumo: Este artigo avança subsídios para uma compreensão e aplicação prática da norma que resulta da conjugação do n.º 3 do artigo 16.º com o artigo 6.º do DSA, debruçando-se, em particular, sobre o alcance da responsabilidade dos prestadores de serviços intermediários (em particular, plataformas em linha), na sequência de uma notificação relativa à presença de conteúdo putativamente ilegal nas suas plataformas ou interfaces. Através de uma abordagem eminentemente jurídica, e próxima, quer dos princípios e do enquadramento de Direito da União Europeia no que se refere à responsabilidade das plataformas, quer da dogmática dos direitos fundamentais, procura-se resposta para um conjunto de questões de estudo. Qual a natureza e o *telos* da presunção de conhecimento do n.º 3 do artigo 16.º do DSA? Que condições deverão estar preenchidas para que se possa presumir o conhecimento da ilegalidade de conteúdos em linha pelo prestador de serviços intermediários, e, em especial, para que se lhe possa ser exigida uma atuação, sob pena de perda da isenção de responsabilidade pelos conteúdos de terceiros? Em particular, que bitola deverá ser adotada na determinação da *suficiência* das informações prestadas pelo destinatário do serviço-notificante? Poder-se-á encontrar, no n.º 3 do artigo 16.º ou no próprio artigo 6.º do DSA, ancoragem para sustentar a necessidade de um regime diferenciado, distinguindo entre conteúdos ilegais evidentes ou manifestos, e conteúdos cuja avaliação de (i)legalidade carece de uma análise aprofundada que o prestador não está em condições de (nem tem

* Doutora em Direito. Professora Auxiliar Convidada da Faculdade de Direito da Universidade do Porto e Investigadora Integrada do CIJ – Centro de Investigação Interdisciplinar em Justiça. Associada do CEJURE – Centro Jurídico do Estado. ORCID: https://orcid.org/0000-0003-0448-2951. Contacto: ineves@direito.up.pt

legitimidade para) fazer (no imediato ou sem mais) e que, portanto, não será de molde a fundar uma responsabilidade pelo conhecimento (presumido)? Que elementos devem ser considerados, pelas entidades competentes, na avaliação da responsabilidade da plataforma? Que medidas de execução deverão ser preferidas em contextos de incerteza (quanto ao efetivo conhecimento pela plataforma)? Finalmente, tendo em conta o princípio da não exclusividade da responsabilidade das plataformas, qual deverá ser o papel de autoridades administrativas e jurisdicionais perante informações ou conteúdos cuja ilegalidade não seja manifesta, gerando dúvidas sobre o grau de diligência esperado da plataforma?

É para estas questões que o texto procura avançar respostas ou princípios de um enquadramento, não só proporcional e razoável à luz dos diferentes interesses presentes, como, e bem assim, compatível com o objetivo de garantir a efetivação dos direitos fundamentais em linha. Em particular, uma vez delimitado o alcance da presunção de conhecimento consagrada no artigo 16.º, n.º 3, do DSA – presunção ilidível e apenas acionada perante notificações suficientemente precisas, fundamentadas, e atinentes a conteúdos manifestamente ilegais –, desenvolve-se a proposta de uma abordagem diferenciada. Nesta perspetiva, considera-se necessário distinguir entre conteúdos manifestamente ilegais e aqueles cuja qualificação dependa de uma apreciação jurídica mais desenvolvida e, em princípio, humana e pública. Em articulação com esta proposta substantiva, sublinha-se a importância de uma maior coordenação entre plataformas e autoridades competentes, assegurando uma intervenção acrescida das autoridades administrativas e jurisdicionais na tarefa de qualificação jurídica de conteúdos.

Palavras-chave: serviços digitais; plataformas em linha; isenções de responsabilidade; mecanismos de notificação e ação; conhecimento; ilegalidade manifesta; direitos fundamentais; *public enforcement*; *private enforcement*.

Abstract: This article provides a theoretical and practical framework for understanding the rule resulting from the combination of Articles 16(3) and 6 of the DSA. In particular, it focuses on the scope of the liability of intermediary service providers (in particular, online platforms) following a notification regarding the presence on their platforms or interfaces of content that may be illegal under the relevant

legislation. By adopting an approach that is both legally sound and aligned with the principles and framework of European Union law regarding the liability of platforms and the fundamental rights doctrine, we aim to provide answers to a series of study questions. What is the nature of the presumption of knowledge set forth in Article 16(3) of the DSA? What conditions must be met for an intermediary service provider to be presumed to have knowledge of the illegality of online content, and in particular for it to be required to take action, under penalty of losing its exemption from liability for third-party content? What criteria should be employed when determining the adequacy of the information provided by the recipient of the service (the notifier)? Could Article 16(3) or Article 6 of the DSA itself serve as a basis for advocating a differentiated regulatory approach, distinguishing between content that is clearly or manifestly illegal and content whose legality may require an assessment that the provider is not in a position to make immediately or without further ado, and which, therefore, must not give rise to liability for presumed knowledge? What factors should the relevant authorities consider when evaluating the platform's knowledge and liability? In contexts of uncertainty as to whether the platform may be deemed to have actual knowledge, which enforcement measures should be preferred? Ultimately, in light of the principle of non-exclusive responsibility of platforms, what should be the role of administrative and judicial authorities when confronted with information or content that is not manifestly illegal, particularly in instances of uncertainty regarding the degree of diligence expected of the platform?

The text aims to provide responses and advance principles for a structured approach that is not only proportionate and reasonable in light of the various interests at stake, but also consistent with the objective of ensuring the realisation of fundamental rights online. In particular, once the scope of the presumption of knowledge enshrined in Article 16(3) of the DSA has been delimited – a rebuttable presumption that is triggered only by notifications that are sufficiently precise, well-founded, and relate to manifestly illegal content – a differentiated approach is proposed. From this perspective, it is considered necessary to distinguish between manifestly illegal content and that whose qualification requires a more developed legal assessment, one that is, in principle, human and public in nature. In conjunction with this substantive

proposal, the paper underscores the importance of greater coordination between platforms and competent authorities, ensuring enhanced involvement of administrative and judicial bodies in the legal qualification of online content.

Keywords: digital services; online platforms; exemptions from liability; notice and action mechanisms; knowledge; manifest illegality; fundamental rights; public enforcement; private enforcement.

1. Introdução

1.1. Aproximação ao quid em estudo

Com o Regulamento (UE) 2022/2065 do Parlamento Europeu e do Conselho de 19 de outubro de 2022 relativo a um mercado único para os serviços digitais ("Regulamento dos Serviços Digitais", "Regulamento"ou "DSA")[1], pretendeu o legislador europeu, através de um conjunto de regras harmonizadas[2], "assegurar um ambiente em linha seguro, previsível e fiável, combatendo a difusão de conteúdos ilegais em linha e os riscos sociais que a difusão de desinformação ou de outros conteúdos pode gerar, e no qual os direitos fundamentais consagrados na Carta sejam eficazmente protegidos e a inovação seja facilitada"[3]. Para esse efeito, e dado que o "comportamento responsável e diligente dos prestadores de serviços intermediários é essencial para um ambiente em

---

[1] União Europeia. *Regulamento (UE) 2022/2065 do Parlamento Europeu e do Conselho de 19 de outubro de 2022*, relativo a um mercado único para os serviços digitais e que altera a Diretiva 2000/31/CE (*Regulamento dos Serviços Digitais*). *JOUE*, L 277, 27 out. 2022, p. 1-102. Disponível em: http://data.europa.eu/eli/reg/2022/2065/oj. Acesso em: 15 out. 2025. Para uma análise artigo a artigo, cf. Wilman, Folkert; Kalėda, Saulius Lukas; Loewenthal, Paul-John. *The EU Digital Services Act*. Oxford: Oxford University Press, 2024. Disponível em: http://dx.doi.org/10.1093/law/9780198892847.001.0001. Acesso em: 15 out. 2025.

[2] Sobre este efeito de harmonização e preempção da legislação nacional, cf. Husovec, Martin. "The DSA as a Cornerstone of the EU Single Market". In: *Principles of the Digital Services Act*. Oxford: Oxford University Press, 2024. Disponível em: http://dx.doi.org/10.1093/law-ocl/9780192882455.003.0017. Acesso em: 15 out. 2025.

[3] Cf. considerando 9, DSA. Sobre este objetivo, e a regulação digital que antecedeu o DSA, cf. Wilman, Folkert; Kalėda, Saulius Lukas; Loewenthal, Paul-John. "Introduction: Origins and Objectives of the DSA". In: *The EU Digital Services Act*. Oxford: Oxford University Press, 2024. Disponível em: http://dx.doi.org/10.1093/law/9780198892847.003.0001. Acesso em: 15 out. 2025.

linha seguro, previsível e fiável e para permitir aos cidadãos da União e a outras pessoas o exercício dos seus direitos fundamentais garantidos pela Carta dos Direitos Fundamentais da União Europeia"[4], o DSA introduz uma mudança de paradigma no que se refere à responsabilidade dos prestadores de serviços intermediários.

Além da previsão de regras e de condições determinativas da isenção de responsabilidade por conteúdos ilegais fornecidos pelos destinatários do serviço (atualizando a Diretiva sobre o comércio eletrónico)[5], o DSA consagra, *ex novo*, "um conjunto claro, eficaz, previsível e equilibrado de obrigações harmonizadas de devida diligência para os prestadores de serviços intermediários"[6], que seguem uma lógica escalonada, adequada ao tipo, à dimensão e à natureza do serviço intermediário em causa, algumas das quais acionadas na sequência de alertas, nomeadamente por destinatários do serviço ou outras entidades, para a presença de atividades ou conteúdos ilegais em linha.

O aditamento desta dimensão de responsabilidade adjetivo-procedimental representa uma evolução demonstrativa da importância e da necessidade de colaboração das plataformas[7] no combate à disseminação de conteúdos ilegais em linha[8], enquanto *perigo* que reclama uma abordagem cooperativa entre atores públicos e privados. Além do mais,

---

[4] Cf. considerando 3, DSA.

[5] Sobre esta "importação", cf. Peguera, Miquel. "The Platform Neutrality Conundrum and the Digital Services Act". *IIC – International Review of Intellectual Property and Competition Law*, 2022. Disponível em: http://dx.doi.org/10.1007/s40319-022-01205-7. Acesso em: 15 out. 2025. *Vd.*, também, sobre as isenções de responsabilidade em geral, Wilman, Folkert; Kalėda, Saulius Lukas; Loewenthal, Paul-John. "Liability of Providers of Intermediary Services". In: *The EU Digital Services Act*. Oxford: Oxford University Press, 2024. Disponível em: http://dx.doi.org/10.1093/law/9780198892847.003.0003. Acesso em: 15 out. 2025.

[6] Cf. considerando 40, DSA. Sobre estas obrigações, cf. Wilman, Folkert; Kalėda, Saulius Lukas; Loewenthal, Paul-John. "Implementation, Cooperation, Penalties and Enforcement". In: *The EU Digital Services Act*. Oxford University Press, 2024. Disponível em: http://dx.doi.org/10.1093/law/9780198892847.003.0005. Acesso em: 15 out. 2025.

[7] Apesar de as plataformas em linha não esgotarem os prestadores de serviços intermediários abrangidos pelo DSA, a sua centralidade justifica a utilização em *quasi*-sinonímia, no título e em texto.

[8] Sobre a necessidade e utilidade social desta dimensão de processo devido, cf. Husovec, Martin. "Content Moderation: Outline". In: *Principles of the Digital Services Act*. Oxford: Oxford University Press, 2024. Disponível em: http://dx.doi.org/10.1093/law-ocl/9780192882455.003.0010. Acesso em: 15 out. 2025.

reforça a qualificação do Regulamento dos Serviços Digitais como "*a legal instrument that affirms the digital rule of law in the European Union. The DSA is a cornerstone in the EU's effort to ascertain the digital rule of law and the digital due process*"[9].

Está-se perante uma mudança necessária, legítima e justificada pelos deveres de proteção que fluem das normas de direitos fundamentais consagrados, quer nos textos constitucionais dos Estados-Membros, quer, e em particular, na Carta dos Direitos Fundamentais da União Europeia ("CDF")[10], que vinculam o legislador europeu a adotar enquadramentos normativos e regulatórios capazes de assegurar a proteção e a efetividade dos direitos fundamentais em linha, se necessário através da responsabilização de operadores privados (e, com ela, a limitação dos respetivos direitos-liberdades).

Não obstante, e como o próprio legislador reconhece no considerando 27 do DSA, "é importante recordar que, apesar do papel geralmente importante desempenhado por tais prestadores, a abordagem do problema dos conteúdos e atividades ilegais em linha não deverá incidir exclusivamente na sua responsabilização e nas suas responsabilidades". Com esta ressalva em mente, considera-se crucial sinalizar os juízos subjetivos e os *cinzentos* ínsitos à moderação de conteúdos em linha e, em particular, à tarefa de determinação da ilegalidade de particulares elementos de informação. Na prática, poderão ser muito diversos os conteúdos com os quais a plataforma se depara (na sequência de uma investigação voluntária, ou de uma notificação por um utilizador). A título de exemplo, não restarão dúvidas da diferença entre conteúdos que envolvem uma ameaça iminente para a vida ou a segurança das pessoas, *vis-à-vis* àqueles que, em contraste, se vejam a meio caminho (num *intermezzo*) entre o exercício de direitos como a liberdade de expressão, a liberdade de criação artística, o humor e a sátira, por um lado, e a lesão

---

[9] Cf. Rocha, Tiago Morais. "Digital Services Act: Towards the Digital Rule of Law". In: Enes, Graça; Neves, Inês; Rocha, Tiago Morais (eds.). *A Digital Europe for Citizens (DigEUCit 2023)*. Cham: Springer, 2026. p. 189-208. Disponível em: https://doi.org/10.1007/978-3-032-02500-5_11. Acesso em: 15 out. 2025, p. 211.

[10] União Europeia. *Carta dos Direitos Fundamentais da União Europeia*. *JOUE*, C 202, 7 jun. 2016, p. 389-405. Disponível em: http://data.europa.eu/eli/treaty/char_2016/oj. Acesso em: 15 out. 2025.

de direitos pessoais, como o direito ao bom nome, à imagem, à reputação e à reserva da intimidade da vida privada e familiar, por outro[11].

Porque assim é, torna-se necessário procurar – para os preceitos do DSA contendentes com a possível responsabilidade das plataformas por conteúdos de terceiros –, uma interpretação razoável e proporcionada.

Trata-se de desiderato particularmente relevante num contexto de crescente *civic vigilantism*, com o incremento das ações populares (propostas por associações de defesa dos consumidores *atípicas*), e que poderá sujeitar as plataformas a *tarefas diabólicas* de ponderação e resolução de conflitos entre direitos fundamentais de diferentes titulares, que aquelas possam ser chamadas a resolver como árbitros, e de cuja resposta resulte a frustração de alguma das expectativas de sinal contrário em presença.

1.2. Delimitação e sequência

Com este objetivo em mente, debruça-se o presente texto sobre um dos preceitos e soluções normativas cuja interpretação e aplicação prática se afiguram justificativas de particulares cautelas e, acima de tudo, carecidas de diferenciações, que não deixarão de impactar na responsabilidade e na responsabilização das plataformas por conteúdos de terceiros. Referimo-nos à norma que resulta da leitura conjugada do n.º 3 do artigo 16.º com o artigo 6.º do DSA, no que se refere às plataformas em linha.

Com efeito, entre as obrigações de devida diligência introduzidas no DSA – especialmente *tradutoras* da responsabilidade (acrescida) dos prestadores de serviços de alojamento virtual, incluindo as plataformas em linha –, está a criação de mecanismos de notificação e ação, que permitam "a qualquer pessoa ou entidade notificá-los da presença, no seu

[11] No sentido da diferença entre *easy* e *hard cases*, e considerando que os *Large Language Models* (LLM) poderão desempenhar um papel relevante na pré-filtragem de conteúdos e na recondução dos casos mais complexos à análise humana, cf. HUANG, Tao. "Content Moderation by LLM: From Accuracy to Legitimacy". *Artificial Intelligence Review*, v. 58, art. 320, 2025. Disponível em: https://doi.org/10.1007/s10462-025-10820-9. Acesso em: 15 out. 2025.

serviço, de elementos específicos de informação que a pessoa ou a entidade considere ser conteúdo ilegal"[12].

Fiel à harmonização pretendida, o legislador europeu prevê um conjunto de requisitos aplicáveis ao *design* dos mecanismos de ação e de notificação, que impactam na liberdade de empresa dos prestadores[13], restringindo a sua margem de conformação no que se refere à arquitetura digital dos seus serviços e plataformas. O objetivo é, afirma-se no considerando 52 do DSA, assegurar "o tratamento atempado, diligente e não-arbitrário das notificações com base em regras uniformes, transparentes e claras e que prevejam garantias sólidas para proteger os direitos e interesses legítimos de todas as partes afetadas, nomeadamente os seus direitos fundamentais garantidos pela Carta, independentemente do Estado-Membro em que estejam estabelecidas ou residam e do domínio do direito em questão".

Além da limitação da liberdade de *design* e gestão da plataforma, também a opção de escolha e a liberdade de ação do prestador de serviços – entre concordar ou não com a avaliação do utilizador, agindo ou não em conformidade – surge *atenuada*. Com efeito, determina o n.º 3 do artigo 16.º do DSA que as notificações do utilizador "dão lugar a um conhecimento efetivo ou a um alerta para efeitos do artigo 6.º relativamente ao elemento específico de informação em causa quando permitem a um prestador diligente de alojamento virtual identificar a ilegalidade da atividade ou das informações em causa sem um exame jurídico pormenorizado". Por outras palavras, o legislador europeu *presume* o conhecimento da atividade ou conteúdo ilegal pela plataforma, sempre que uma notificação contenha *i*) informações suficientes para permitir a um prestador diligente de serviços de alojamento virtual identificar, *ii*) sem um exame jurídico pormenorizado, que *iii*) é evidente que o conteúdo é ilegal.

Ao associar às notificações dos utilizadores, nos termos do n.º 3 do artigo 16.º do DSA, um verdadeiro *trigger effect* de responsabilidade, de sinal contrário à inexistência de obrigações gerais de vigilância ou de apuramento ativo dos factos[14], e, bem assim, contrastante com o prin-

---

[12] Cf. n.º 1 do artigo 16.º, DSA.

[13] Reconhecida no artigo 16.º, CDF.

[14] Cf. artigo 8.º, DSA.

cípio de circunscrição da obrigatoriedade de atuação ao cumprimento de ordens judiciárias ou administrativas[15], o legislador europeu deixa as plataformas perante dois cenários possíveis.

Ou *i*) não se encontram preenchidos os pressupostos do n.º 3 do artigo 16.º do DSA, caso em que não se poderá mobilizar a *presunção de conhecimento*, mantendo-se a regra da isenção de responsabilidade do prestador, nos termos do artigo 6.º do DSA (sem prejuízo da sua eventual responsabilidade pelo incumprimento de algum dos requisitos de devida diligência impostos pelo DSA[16]); ou *ii*) a notificação, pelo seu conteúdo e pelo tipo de ilegalidade (manifesta ou evidente), permite acionar a *presunção de conhecimento efetivo*, forçando o prestador a atuar com diligência no sentido de suprimir ou desativar o acesso aos conteúdos ilegais[17], sob pena de, não o fazendo, poder ser considerado responsável pelas informações ilegais armazenadas na sua plataforma.

É sobre esta responsabilidade pelo conhecimento presumido que o texto se debruça, procurando – através da metodologia típica da investigação científica em Direito – respostas para um conjunto de interrogações com relevo prático, seja para as próprias plataformas, que deverão beneficiar de segurança jurídica no que se refere ao alcance dos seus deveres de *compliance* e de diligência; seja para os utilizadores, que não deverão ter de recear a sua participação em linha, num contexto conflitual acrescido e facilitado pela anonimidade; seja para as próprias autoridades administrativas e judiciárias, chamadas a dirimir litígios que abandonam o paradigma das relações bilaterais, deslocando-se para o domínio do poligonal, com a intervenção de uma pluralidade de partes, entre as quais (diferentes) destinatários e utilizadores dos serviços em linha, e prestadores de serviços intermediários.

Depois de um capítulo introdutório dedicado aos mecanismos de notificação e ação, tal como consagrados no artigo 16.º do DSA, o enfoque do texto recentra-se sobre a norma do n.º 3 do artigo 16.º, procurando articulá-la com o disposto no artigo 6.º do DSA, no qual estão previstas as condições de isenção de responsabilidade aplicáveis aos

---

[15] Cf. artigo 9.º, DSA.

[16] Por exemplo no que se refere ao *design* dos mecanismos de notificação – *inter alia*, a obrigação de facilitar a apresentação de notificações fundamentadas, i.e., com uma explicação das razões – cf. considerando 53, DSA.

[17] Conforme preceitua a alínea *b*) do n.º 1 do artigo 6.º, DSA.

prestadores de serviços de alojamento virtual (incluindo as plataformas em linha).

Laborando sobre a teleologia apontada à norma resultante da conjugação dos dois preceitos, avançam-se, depois, subsídios para uma hermenêutica proporcional e razoável da presunção de conhecimento consagrada no n.º 3 do artigo 16.º do DSA, tendo em conta o teor das notificações e, em particular, o tipo de conteúdos sinalizados. Em particular, reputa-se necessário distinguir entre conteúdos cuja ilegalidade é evidente ou manifesta e cenários (dúbios) a propósito dos quais se considera que a plataforma – independentemente da respetiva capacidade técnica e operacional – nem terá legitimidade ou competência nem oferecerá as garantias de imparcialidade e de conhecimento, necessárias para uma aferição justa, não arbitrária e proporcional da (i)legalidade do conteúdo, e para uma subsequente atuação em conformidade.

Finalmente, procura o texto avançar um conjunto de notas de cautela que, em sede de aplicação e execução do Regulamento, deverão acompanhar e *atuar* a referida hermenêutica, nomeadamente pelas autoridades responsáveis pelo *public enforcement* do DSA (coordenadores dos serviços digitais e Comissão Europeia), e, bem assim, balizar e enquadrar eventuais pretensões indemnizatórias em contextos de *private enforcement*.

Perante a multiplicidade de *zonas cinzentas*, e recordando o princípio de responsabilidade partilhada a que já se fez referência, não sem desconsiderar os limites da delegação privados de tarefas eminentemente estaduais, propõe-se, como solução possível, a intensificação de vias e canais de diálogo entre prestadores de serviços intermediários e autoridades administrativas e judiciárias competentes, por forma a assegurar soluções expeditas, mas, e simultaneamente, conformes e efetivamente protetoras da multiplicidade de direitos fundamentais em presença.

2. Dos mecanismos de notificação e ação no artigo 16.º do DSA: em particular, a presunção de conhecimento no n.º 3

2.1. Uma breve introdução ao artigo 16.º do DSA

Reconhecendo que os "prestadores de serviços de alojamento virtual desempenham um papel especialmente importante na luta contra os conteúdos ilegais em linha, uma vez que armazenam informações fornecidas pelos destinatários do serviço e a pedido destes e, normalmente, dão a outros destinatários acesso às mesmas, por vezes em grande escala"[18], o legislador europeu considera importante que todos eles, "independentemente da sua dimensão, criem mecanismos de notificação e ação facilmente acessíveis e de utilização simples, que facilitem a notificação de elementos específicos de informação que a parte notificante considere constituírem conteúdos ilegais ao prestador de serviços de alojamento virtual em causa ('notificação'), nos termos da qual esse prestador pode decidir se concorda ou não com a avaliação e se pretende suprimir os conteúdos ou bloquear o acesso aos mesmos ('ação')"[19] [20].

É no artigo 16.º do Regulamento dos Serviços Digitais que se encontram previstos os mecanismos de notificação e ação, aí se prevendo um conjunto de regras uniformes, transparentes e claras, explicadas pelo objetivo de assegurar um tratamento atempado, diligente e não-arbitrário das notificações, e que se espera que funcionem, também, como

---

[18] Cf. considerando 50, DSA.

[19] *Ibidem*.

[20] Saliente-se, a este propósito, que a preocupação (de harmonização) do legislador europeu se dirige, sobretudo, aos conteúdos ilegais, pese embora reconheça que os mesmos mecanismos de notificação e ação possam também ser aplicáveis a conteúdos que violem os termos e condições do serviço de alojamento virtual (e que não configurem conteúdos ilegais). Este ponto fulcral de preocupação é sobretudo visível no considerando 50, exigindo que os mecanismos de notificação e ação aplicáveis a conteúdos ilegais devam, "pelo menos, ser tão fáceis de encontrar e utilizar como mecanismos de notificação de conteúdos que violem os termos e condições do serviço de alojamento virtual", regra que aponta para a possibilidade de um tratamento de favor dos primeiros, em relação aos segundos. Sobre a necessidade de regras mais específicas que evitem a atual fragmentação de abordagens e práticas, cf. Drolsbach, Chiara Patricia; Pröllochs, Nicolas. "Content Moderation on Social Media in the EU: Insights from the DSA Transparency Database". In: *Proceedings of the ACM Web Conference 2024 (WWW '24)*. New York: Association for Computing Machinery, 2024. Disponível em: http://dx.doi.org/10.1145/3589335.3651482. Acesso em: 15 out. 2025.

garantias sólidas para a proteção dos direitos e interesses legítimos de todas as partes afetadas[21].

O preceito é aplicável a todos os prestadores de serviços de alojamento virtual (doravante, também "prestadores"), incluindo, como atores protagonistas, as plataformas em linha (de muito grande dimensão), de que aqui se trata em particular.

Como resulta do nele disposto, os prestadores de serviços de alojamento virtual deverão criar e desenhar mecanismos que, obedecendo a um conjunto de requisitos mínimos, permitam a qualquer interessado notificá-los da presença de conteúdos ilegais no respetivo serviço em linha. Entre as condições exigidas no que se refere ao desenho dos referidos mecanismos, está a facilidade do acesso, da utilização e da apresentação de notificações, as quais deverão ser avançadas em termos suficientemente precisos e adequadamente fundamentados, condições que naturalmente impactam na liberdade de *design* e de conformação do prestador, cuja *interface* deverá ser pensada, desenhada e instituída, por forma a permitir que as notificações e os seus autores apresentem esse conteúdo mínimo[22].

Reconhecendo, porém, esta dimensão da liberdade de empresa do prestador de serviços – a liberdade de conformação da sua plataforma –, e a variabilidade de possibilidades que se abrem a este propósito[23], reconhece o legislador europeu, no artigo 44.º, n.º 1, alínea *a*), do DSA, a importância de normas facultativas estabelecidas por organismos de normalização europeus e internacionais pertinentes, relevantes para nortear o exercício dessa liberdade em termos alinhados com as expectativas do legislador, sem, porém, aniquilar o núcleo essencial de uma liberdade de cujo exercício – em *concorrência* com soluções *concorrentes* de outros *players* – poderá até resultar a introdução no mercado de arquiteturas digitais particularmente inovadoras e (mais) alinhadas com os direitos fundamentais a garantir.

---

[21] Cf. considerandos 51 e 52, DSA.

[22] Cf. n.os 1 e 2 do artigo 16.º, DSA.

[23] Essa liberdade decorre, desde logo, da liberdade de empresa, enquanto direito fundamental consagrado no artigo 16.º da Carta dos Direitos Fundamentais da União Europeia.

Além do conjunto de requisitos contendentes com o desenho e a arquitetura dos mecanismos, o artigo 16.º do DSA obriga, ainda, o prestador a agir na sequência das notificações[24].

Com efeito, é possível extrair do artigo 16.º do DSA um conjunto de obrigações suficientemente precisas e determinadas que impõem ao prestador deveres de atuação e obrigações comunicacionais.

Desde logo, tem-se o envio, "sem demora injustificada", de um aviso de receção da notificação (quando a notificação contenha os dados de contacto do notificante)[25]. Segue-se a adoção de uma decisão atempada, diligente, não arbitrária e objetiva[26]. Num terceiro momento, impõe-se a notificação da decisão final ao notificante, "sem demora injustificada", e acompanhada de "informações sobre as possibilidades de reparação relativas a essa decisão"[27] e, bem assim, sobre a eventual utilização de meios automatizados no tratamento das notificações e/ou tomada de decisões[28].

---

[24] Os deveres de ação constantes do artigo 16.º devem ser complementados com outros, também vinculativos para o prestador de serviços visado. Exemplos são os do artigo 17.º do DSA, que obriga os prestadores de serviços abrangidos a apresentar uma exposição de motivos clara e específica a todos os destinatários do serviço afetados por eventuais restrições adotadas, *inter alia*, na sequência de uma notificação apresentada nos termos do artigo 16.º do DSA, incluindo, caso seja estritamente necessário, a identidade do notificador. Cumpre, ainda, realçar o dever de informação às autoridades responsáveis pela aplicação da lei, consagrado no artigo 18.º do DSA, e aplicável aos casos de conhecimento de qualquer informação que levante suspeitas da prática (consumada, iminente ou possível) de um crime grave que envolva uma ameaça à vida ou à segurança de uma pessoa ou pessoas. Além destes, há que não ignorar a existência de regimes especiais, por exemplo, aquele constante do artigo 18.º do *Regulamento (UE) 2024/1083 do Parlamento Europeu e do Conselho, de 11 de abril de 2024*, que cria um regime comum para os serviços de comunicação social no mercado interno e que altera a Diretiva 2010/13/UE (*Regulamento Europeu relativo à Liberdade dos Meios de Comunicação Social*). *JOUE*, L 1083, 17 abr. 2024. Disponível em: http://data.europa.eu/eli/reg/2024/1083/oj. Acesso em: 15 out. 2025.

[25] Cf. n.º 4 do artigo 16.º, DSA.

[26] Cf. n.º 6 do artigo 16.º, DSA. Importa referir que, nos termos do artigo 22.º do DSA, "as notificações apresentadas por sinalizadores de confiança, agindo dentro do seu domínio de competências designado, através dos mecanismos referidos no artigo 16.º, têm prioridade e são tratadas e objeto de uma decisão sem demora indevida".

[27] Cf. n.º 5 do artigo 16.º, DSA.

[28] Cf. n.º 6 do artigo 16.º, DSA.

Pese embora a mobilização de um conceito indeterminado para o enquadramento temporal destes deveres – a "demora (in)justificada" –, a margem de atuação reconhecida ao prestador é mais aparente do que efetiva, desde logo em razão do relevo votado ao *dia em que o destinatário é informado*. Considere-se, a título de exemplo, a obrigação de os fornecedores de plataformas em linha concederem aos destinatários do serviço acesso a um sistema interno eficaz de gestão de reclamações, que, ao abrigo do disposto no artigo 20.º do DSA, deverá permanecer disponível durante um período mínimo de seis meses, contados a partir do momento da informação da decisão final ao notificante, nos termos do n.º 5 do artigo 16.º do DSA[29], assim conformando, ainda que indiretamente, a celeridade esperada do prestador.

2.2. Em particular, a presunção de conhecimento no artigo 16.º, n.º 3, em articulação com o artigo 6.º do DSA

Sem prejuízo da importância dos diferentes números do artigo 16.º, o presente texto justifica uma atenção particular ao disposto no seu n.º 3, e que importa ler em conjugação com o considerando 53 do DSA, do qual resulta que: "[s]empre que uma notificação contenha informações suficientes para permitir a um prestador diligente de serviços de alojamento virtual identificar, sem um exame jurídico pormenorizado, que é evidente que o conteúdo é ilegal, deverá considerar-se que a notificação dá origem ao conhecimento efetivo ou ao conhecimento da ilegalidade".

Em termos próximos, lê-se no n.º 3 do artigo 16.º do DSA: "[c]onsidera-se que as notificações referidas no presente artigo dão lugar a um conhecimento efetivo ou a um alerta para efeitos do artigo 6.º relativamente ao elemento específico de informação em causa quando permitem a um prestador diligente de alojamento virtual identificar a ilegalidade da atividade ou das informações em causa sem um exame jurídico pormenorizado".

O n.º 3 do artigo 16.º do DSA corporiza, pois, uma *presunção*, isto é, uma ilação que o legislador tira de um facto conhecido (a notificação de um conteúdo potencialmente ilegal, com particulares requisitos) para

[29] Cf. n.os 1 e 2 do artigo 16.º, DSA.

firmar um facto desconhecido (o alerta ou o conhecimento efetivo, pelo prestador de serviços intermediários, da presença de conteúdo ilegal).

Na medida em que nada em contrário resulta do DSA, e inclusive pela fragilidade prática da relação entre facto conhecido e desconhecido relativamente a particulares conteúdos [aqueles cuja (i)legalidade possa não ser patente ou manifesta], considera-se estar em causa uma presunção relativa, e, portanto, ilidível mediante prova em contrário, de onde resulta a necessidade de reconhecer à plataforma, uma vez demandada por autoridades administrativas ou judiciais competentes, o direito a refutar o respetivo conhecimento, nomeadamente através de elementos demonstrativos de dúvida razoável em relação à natureza (ilegal) das informações ou conteúdos em questão.

A compreensão do sentido e do alcance do n.º 3 do artigo 16.º não prescinde, porém, de uma análise do artigo 6.º do DSA, preceito para o qual aquele remete.

O artigo 6.º do DSA consagra as condições necessárias para que o prestador de um serviço de alojamento virtual, isto é, e recordando, "um serviço da sociedade da informação que consista na armazenagem de informações prestadas por um destinatário do serviço"[30], possa beneficiar de uma isenção de responsabilidade pelas informações armazenadas a pedido de um destinatário do serviço[31]. Segundo o n.º 1 do artigo 6.º do DSA, o prestador apenas não será responsabilizado pelas informações de terceiros, armazenadas no seu serviço, quando "[n]ão tenha conhecimento efetivo da atividade ou conteúdo ilegal e, no que se refere a uma ação de indemnização por perdas e danos, não tenha conhecimento de factos ou de circunstâncias que evidenciem a ilegalidade da atividade ou do conteúdo" e/ou, "[a] partir do momento em que tenha conhecimento da ilicitude [ou seja alertado para a natureza ilegal dos mesmos][32], atue com diligência no sentido de suprimir ou desativar o acesso aos conteúdos ilegais".

Segundo o considerando 22 do DSA, o "prestador pode tomar conhecimento efetivo dos conteúdos em causa, ou ser alertado para a natureza

---

[30] Cf. n.º 1 do artigo 6.º, DSA.

[31] Sobre estas condições, cf. Husovec, Martin. "Liability Exemptions: Specific Services". In: *Principles of the Digital Services Act*. Oxford: Oxford University Press, 2024. Disponível em: http://dx.doi.org/10.1093/law-ocl/9780192882455.003.0007. Acesso em: 15 out. 2025.

[32] Cf. considerando 22, DSA.

ilegal dos mesmos, através, nomeadamente, de investigações realizadas por iniciativa própria ou de notificações que lhe sejam apresentadas por pessoas ou entidades nos termos do presente regulamento".

Pese embora não o afirme expressamente, crê-se que, além destas duas hipóteses, a saber, o conhecimento adquirido através de (*i*) investigações próprias ou de (*ii*) notificações de destinatários do serviço, nos termos do artigo 16.º do DSA, será ainda de acrescentar um terceiro cenário de potencial conhecimento, resultante de (*iii*) outras notificações (por exemplo, a *citação para* ou a participação em diligências preparatórias de uma ação de responsabilidade civil).

Por seu turno, sendo o artigo 6.º do DSA igualmente claro na diligência imposta ao prestador – a supressão ou a desativação do acesso aos conteúdos ilegais –, não nos parece, também, que as possibilidades expressamente abertas ao prestador (para manter a sua isenção de responsabilidade) se quedem por aí. Pelo contrário, afigura-se-nos que uma atuação do prestador que vá além disso (optando, por exemplo, pela suspensão ou cessação da prestação do serviço, ou pelo encerramento da conta do destinatário do serviço)[33] poderá também não afetar a manutenção da isenção de responsabilidade, contanto que a concreta resposta se afigure proporcional *in casu*[34].

De tudo o exposto, resulta ser o artigo 6.º, n.º 1, do DSA concretização da teleologia subjacente às isenções de responsabilidade dos

---

[33] Cf. artigo 17.º, DSA.

[34] Em apoio desta interpretação, vejam-se os considerandos 54 e 55 do DSA, dos quais resulta que, na sequência da receção de uma notificação que lhe permita concluir que as informações fornecidas por um particular destinatário constituem conteúdo ilegal (ou incompatível com os seus termos e condições), o prestador de serviços de alojamento virtual deverá atuar "com diligência no sentido de suprimir ou desativar o acesso aos conteúdos ilegais", podendo, em particular, "decidir suprimir ou bloquear o acesso às informações fornecidas por um destinatário do serviço, ou restringir de outra forma a sua visibilidade ou monetização", restrição essa que "pode consistir na despromoção nos sistemas de classificação ou de recomendação, bem como na limitação da acessibilidade de um ou mais destinatários do serviço ou no bloqueio do utilizador de uma comunidade em linha sem que o mesmo disso tenha conhecimento ('bloqueamento oculto' — *shadow banning*)" ou, ainda, "através da suspensão ou cessação do pagamento monetário ou da receita associados a essa informação". Questão distinta é a de saber se, com essa sua atuação, não lesa o princípio da proporcionalidade, indo além do estritamente necessário para garantir a remoção de conteúdos ilegais em linha e, bem assim, a preservação da isenção de responsabilidade.

prestadores de serviços intermediários (por conteúdos de terceiros). Como resulta do considerando 18 do DSA, as isenções pressupõem que o prestador presta "os serviços de forma neutra, através de um tratamento meramente técnico e automático das informações prestadas pelo destinatário do serviço", não desempenhando um "papel ativo que lhe permita ter conhecimento ou controlo dessas informações". Isto mesmo explica, aliás, que as isenções se não apliquem "às informações fornecidas não pelo destinatário do serviço, mas pelo próprio prestador do serviço intermediário"[35].

É neste contexto que importa compreender o sentido e o alcance da norma que resulta do n.º 3 do artigo 16.º conjugado com o artigo 6.º do DSA, e, com ela, do afastamento do princípio de isenção de responsabilidade por conteúdos de terceiros, em razão de o prestador – uma vez alertado, por um utilizador, para a presença de conteúdo ilegal em linha, nos termos do n.º 3 do artigo 16.º do DSA – não ter atuado "com diligência no sentido de suprimir ou desativar o acesso aos conteúdos ilegais"[36]. Em particular, considera-se que essa *diligência* deverá ser lida em conjugação com outras exigências consagradas no Regulamento, entre as quais a de que, seja na conceção, seja na aplicação, seja no cumprimento das restrições, os prestadores de serviços intermediários atuem "de forma não arbitrária e não discriminatória", tendo em conta e em respeito pelos "direitos e interesses legítimos dos destinatários do serviço, incluindo os direitos fundamentais consagrados na Carta"[37].

É dessa articulação e indagação aprofundada que se tratará nos próximos pontos.

---

[35] Cf. considerando 18, DSA.

[36] Cf. alínea *b*) do n.º 1 do artigo 6.º, DSA.

[37] Cf. considerando 47, DSA. *Vd.*, também, considerando 22, DSA.

3. Subsídios para uma hermenêutica proporcionada e razoável do artigo 16.º, n.º 3, do DSA: entre a suficiência das notificações e o tipo e/ /ou teor dos conteúdos

3.1. Aproximação ao telos

A diligência esperada do prestador na sequência de uma notificação ou alerta referente(s) à presença de conteúdo ilegal em linha deve ser aferida e balizada pela teleologia do DSA.

Resulta, a este propósito, do seu considerando 51 que, "[t]endo em conta a necessidade de ter devidamente em conta os direitos fundamentais garantidos pela Carta de todas as partes interessadas, qualquer medida tomada por um prestador de serviços de alojamento virtual na sequência da receção de uma notificação deverá ser estritamente direcionada, no sentido de que deverá servir para suprimir ou bloquear o acesso a elementos específicos de informação considerados conteúdos ilegais, sem afetar indevidamente a liberdade de expressão e de informação dos destinatários do serviço"[38]. A necessidade de tomar em conta os direitos e interesses legítimos de todas as partes envolvidas é também afirmada, em particular, no n.º 4 do artigo 14.º do DSA, preceito que impõe uma atuação diligente, objetiva e proporcionada do prestador, na

[38] No que aos fornecedores de plataformas em linha de muito grande dimensão e aos motores de pesquisa em linha de muito grande dimensão se refere, aplicam-se, ainda, deveres adicionais de gestão de riscos sistémicos, que incluem "mobilizar os meios necessários para atenuar diligentemente os riscos sistémicos identificados na avaliação dos riscos, no respeito dos direitos fundamentais", nomeadamente tendo em conta a "necessidade de evitar restrições desnecessárias à utilização do seu serviço, tendo em devida conta os potenciais efeitos negativos nesses direitos fundamentais. Esses fornecedores deverão prestar especial atenção ao impacto na liberdade de expressão." – cf. considerando 86, DSA e artigos 34.º e 35.º, DSA. Segundo o n.º 3 deste último artigo, a Comissão, em cooperação com os coordenadores dos serviços digitais, pode emitir diretrizes, contendo, nomeadamente, boas práticas e recomendações de eventuais medidas, "tendo devidamente em conta as possíveis repercussões das medidas nos direitos fundamentais de todas as partes envolvidas consagrados na Carta". Aos deveres relacionados com os riscos sistémicos somam-se, ainda, obrigações acrescidas em casos de crise, cuja identificação e aplicação deverão também ter em conta as respetivas implicações reais ou potenciais para os direitos e interesses legítimos de todas as partes em causa – cf. artigo 36.º, DSA.

aplicação e execução das restrições impostas à utilização do seu serviço e respeitantes às informações prestadas pelos destinatários do serviço[39].

Ciente da *i*) teia de interesses ínsita aos mecanismos de notificação e ação; da *ii*) complexidade da moderação de conteúdos em linha, e, bem assim, dos *iii*) riscos de *overblocking*, inarredáveis perante a multitude de direitos fundamentais potencialmente afetados por uma medida de supressão ou bloqueio de acesso a elementos específicos de informação[40], teve o legislador europeu o cuidado de salvaguardar a liberdade de ação da plataforma.

Com efeito, esclarece a este propósito o considerando 50 do DSA que, na sequência de uma notificação, o prestador "pode decidir se concorda ou não com a avaliação e se pretende suprimir os conteúdos ou bloquear o acesso aos mesmos («ação»)". Por seu turno, resulta dos considerandos 51 e 52 do DSA que a obrigação de dar seguimento às notificações em tempo útil dependerá do tipo de conteúdos ilegais visados pelas notificações e da urgência da tomada de medidas, mais se prevendo que "as notificações deverão, regra geral, ser dirigidas aos prestadores de serviços de alojamento virtual de que se possa razoavelmente esperar que tenham capacidade técnica e operacional para agir".

Todos estes elementos devem *informar* uma leitura e uma aplicação proporcionais dos artigos 6.º, n.º 1, e 16.º, n.º 3, do DSA que, em nosso ver, não traduzem a pura e absoluta imposição, ao prestador de serviços intermediários, de uma particular atuação na sequência de uma notificação indicativa da presença de conteúdos putativamente ilegais, sob pena de uma automática responsabilização pelos referidos conteúdos de terceiros. Pelo contrário, o que resulta desses preceitos é, "apenas", a exigência de adoção de uma especial *diligência* ou consideração pelo prestador.

Assim, e em particular, posto perante a notificação de conteúdos putativamente ilegais de terceiros na sua plataforma, considera-se ficar o prestador de serviços intermediários obrigado a (*i*) avaliar, confirmando

---

[39] Sobre a necessidade de operacionalização do artigo 14.º do DSA à luz do enquadramento aplicável aos direitos fundamentais, cf. QUINTAIS, João Pedro; APPELMAN, Naomi; Ó FATHAIGH, Ronan. "Using Terms and Conditions to Apply Fundamental Rights to Content Moderation". *German Law Journal*, v. 24, p. 1-20, 2023. Disponível em: http://dx.doi.org/10.1017/glj.2023.53. Acesso em: 15 out. 2025.

[40] Cf. considerando 51, DSA.

ou infirmando, a alegação de ilegalidade ínsita à notificação e, apenas em caso de confirmação da ilegalidade do conteúdo, a (*ii*) adotar medidas de supressão ou desativação de acesso aos referidos conteúdos ilegais, nos termos do artigo 6.º, n.º 1, al. *b*), do DSA. Do exposto resulta, pois, que a atuação exigida em (*ii*) se encontra *dependente* da aferição conduzida em (*i*), podendo a responsabilidade do prestador circunscrever-se então à primeira etapa (nomeadamente quando da sua avaliação não resultem indícios seguros da ilegalidade do conteúdo).

Reitere-se que em todos os momentos e relativamente a todas as potenciais vias de atuação, o prestador de serviços se encontra vinculado ao respeito e garantia dos direitos fundamentais de todas as partes envolvidas e, portanto, não apenas aqueles dos autores da notificação, mas, e bem assim, também aqueles dos titulares da informação.

A sustentar esta nota de enquadramento, e pese embora apenas salientado a propósito das decisões de atuação contra conteúdos ilegais[41] adotadas por autoridades judiciárias ou administrativas nacionais, está o considerando 36, apontando claramente no sentido da necessidade de assegurar "o equilíbrio entre o objetivo que a decisão procura alcançar, em conformidade com a base jurídica que permite a sua emissão, e os direitos e interesses legítimos de todos os terceiros passíveis de ser afetados pela ordem, em particular os seus direitos fundamentais consagrados pela Carta". No mesmo sentido, também o considerando 153 do DSA estabelece que, "[n]o exercício dos poderes estabelecidos no presente regulamento, todas as autoridades públicas envolvidas deverão alcançar, em situações em que se verifique um conflito entre os direitos fundamentais pertinentes, um equilíbrio justo entre os direitos em causa, em conformidade com o princípio da proporcionalidade".

Ainda em apoio do que se avança, cumpre salientar que o DSA delega em privados o exercício de funções tipicamente estaduais – parajudiciais –, que, além de coenvolverem tarefas de qualificação jurídica, implicam, também, exercícios de concordância prática entre direitos fundamentais. Ora, podendo essa delegação ser até perspetivada como

[41] Por "decisões de atuação contra conteúdos ilegais" deverão entender-se as decisões de autoridades judiciárias ou administrativas nacionais, nomeadamente as autoridades responsáveis pela aplicação da lei, que ordenem "aos prestadores de serviços intermediários que adotem medidas contra um ou mais elementos específicos de conteúdo ilegal ou que forneçam determinadas informações específicas." – cf. considerando 31, DSA.

necessária ou *inelutável* perante o ritmo vertiginoso e a ubiquidade das relações em linha, não se pode também ignorar que, ao contrário de uma autoridade pública (administrativa ou jurisdicional), um prestador de serviços privado não terá nem a legitimidade nem a competência nem a sensibilidade (jurídica) para encetar este tipo de funções, não pelo menos em todos os casos. De onde, e na mesma ordem de ideias, não poderá também ser por elas, *sem mais*, responsabilizado.

Finalmente, cumpre ainda alertar para a inexistência de uma definição de *conteúdo ilegal* no Regulamento dos Serviços Digitais, o que, perante conceitos ou alcances de ilicitude divergentes entre os Estados-Membros, milita igualmente, pelo menos em cenários de diferença, no sentido de uma particular cautela na leitura da ação exigida nos termos e para os efeitos do artigo 16.º, n.º 3, do DSA[42].

Em suma, a natureza destas funções, em conjugação com a circunstância de serem exigidas a uma entidade privada, conforta(m) a necessidade de uma leitura parcimoniosa da norma que resulta da articulação do n.º 3 do artigo 16.º com o artigo 6.º do DSA. Afinal, de entre os direitos fundamentais garantidos pela CDF cuja proteção e exercício o DSA pretende também proteger, está também a liberdade de empresa[43].

3.2. Densificação das balizas

O legislador europeu avança, nas normas em análise, pistas que permitem enquadrar a solução normativa que resulta dos artigos 16.º, n.º 3, e 6.º do DSA, em termos concordantes com a teleologia e os limites a que já se fez referência. Essas balizas descobrem-se, quer no conjunto de requisitos impostos às notificações que, nos termos do n.º 3 do artigo 16.º do DSA, fundarão a presunção de conhecimento ou alerta

---

[42] Alertando para essa diferença e para a confusão associada, cf. WANG, Tiana. "Delicate Task: Content Moderation and Intermediary Liability in a Post-DSA World". *Berkeley Technology Law Journal*, v. 39, n. 4, p. 1507-1550, 2024. Disponível em: https://doi.org/10.15779/Z38GH9BB6P. Acesso em: 15 out. 2025, p. 1547. Também no sentido da inexistência de definições claras, em termos geradores de incerteza para as plataformas, cf. ENARSSON, Therese. "Navigating Hate Speech and Content Moderation under the DSA: Insights from ECtHR Case Law". *Information & Communications Technology Law*, v. 33, n. 3, p. 384-401, 2024. Disponível em: https://doi.org/10.1080/13600834.2024.2395579. Acesso em: 15 out. 2025.

[43] Cf. considerando 3, DSA e artigo 16.º, CDF.

ao prestador de serviços, quer na particular adjetivação dos conteúdos passíveis de a fundar.

Vejamos.

Dos considerandos 22 e 51 do DSA resulta um conjunto de requisitos que as notificações de pessoas ou entidades deverão preencher para que se possa fundar a presunção de conhecimento efetivo da presença de conteúdos ilegais em linha. Segundo o primeiro considerando, é necessário "que tais notificações sejam suficientemente precisas e adequadamente fundamentadas para permitir a um operador económico diligente identificar, avaliar e, se for caso disso, adotar medidas, de forma razoável, contra os conteúdos alegadamente ilegais". Já o considerando 51 determina que "uma notificação deverá ser estritamente direcionada, no sentido de que deverá servir para suprimir ou bloquear o acesso a elementos específicos de informação considerados conteúdos ilegais, sem afetar indevidamente a liberdade de expressão e de informação dos destinatários do serviço. Por conseguinte, as notificações deverão, regra geral, ser dirigidas aos prestadores de serviços de alojamento virtual de que se possa razoavelmente esperar que tenham capacidade técnica e operacional para agir contra esses elementos específicos". Por seu turno, reitera o considerando 53 que "[s]empre que uma notificação contenha informações suficientes para permitir a um prestador diligente de serviços de alojamento virtual identificar, sem um exame jurídico pormenorizado, que é evidente que o conteúdo é ilegal, deverá considerar-se que a notificação dá origem ao conhecimento efetivo ou ao conhecimento da ilegalidade".

Do exposto extrai-se que as notificações deverão ser *suficientes*, *precisas*, *adequadamente fundamentadas* e *estritamente direcionadas* (a conteúdos específicos).

Por seu turno, dos mesmos considerandos e, bem assim, do próprio n.º 3 do artigo 16.º do DSA, extrai-se a conclusão de que a presunção de conhecimento apenas se aplicará aos conteúdos cuja ilegalidade surja *evidente* a um prestador diligente de serviços de alojamento virtual, sem a necessidade de um exame jurídico pormenorizado. É dizer, apenas fundarão a referida presunção de conhecimento ou alerta, e a consequente exigência de atuação em conformidade (nos termos do artigo 6.º, n.º 1, do DSA), os conteúdos (*i*) cuja ilegalidade seja *patente* (dispensando,

pois, um exame jurídico pormenorizado), (*ii*) a um *prestador diligente* de serviços de alojamento virtual.

Trata-se de balizas não despiciendas.

Começando pela referida em (*i*), não oferece, de facto, dúvida a proliferação de informações e conteúdos em linha cuja qualificação como conteúdo ilegal se vê dependente de uma análise complexa, morosa, circunstanciada e contextual[44], necessariamente variável, não só de acordo com os particulares contextos social, político, jurídico e geográfico de que é originária, como, e bem assim, com a natureza do respetivo autor, ou, ainda, com os usos ou a finalidade da informação, o respetivo conteúdo e forma, incluindo *nuances* e outros elementos contextuais da linguagem e do discurso[45], a que acresce o alcance e a probabilidade ou iminência de dano.

Desde logo, e com a doutrina, importa recordar que o "*DSA is unclear whether the focus must be on illegal – and/or harmful – content. There is no clear definition of what is harmful and what is illegal. This is problematic and should be resolved. The issue is especially concerning in the EU context, given that some contents or behaviors may be illegal in some MSs, and 'not-illegal-but-harmful' in others (i.e. defamation of religion is a criminal offence in Germany, Italy, Poland, and Spain, but not in Denmark and France)*"[46]. Por seu turno, concordamos também com a doutrina que alerta: "*it is unclear to what extent the procedural innovations of the DSA can provide*

---

[44] A título de exemplo, sobre os comportamentos de autolesão não suicida, veja-se Lookingbill, Valerie; LE, Kimanh. "There's Always a Way to Get Around the Guidelines: Nonsuicidal Self-Injury and Content Moderation on TikTok". *Social Media + Society*, v. 10, n. 2, 2024. Disponível em: https://doi.org/10.1177/20563051241254371. Acesso em: 15 out. 2025, considerando: "*While risks of trauma or re-traumatization must be considered in content moderation policies and practices, it is imperative that social media platforms have more nuanced understandings of content type and intention as the effect social media content has on a user is dependent on the context, goal, and set interactions between a user and a system*", p. 4.

[45] A propósito, cf. ARTICLE 19. *Jillian York: The Global Impact of Content Moderation*. 2020. Disponível em: https://www.article19.org/resources/the-global-impact-of-content-moderation/. Acesso em: 15 out. 2025, e ARTICLE 19. *Content Moderation and Freedom of Expression Handbook*. 2023. Disponível em: https://www.article19.org/wp-content/uploads/2023/08/SM4P-Content-moderation-handbook-9-Aug-final.pdf. Acesso em: 15 out. 2025, p. 41 e ss.

[46] Cf. Turillazzi, Aina et al. "The Digital Services Act: An Analysis of Its Ethical, Legal, and Social Implications". *Law, Innovation and Technology*, v. 15, p. 1–94, 2023. Disponível em: http://dx.doi.org/10.1080/17579961.2023.2184136. Acesso em: 15 out. 2025.

predictability and legal certainty concerning online freedom of expression, in the absence of any major harmonization of the substantive law applicable in this very broad and porous area"[47].

O apontamento de cautela afigura-se particularmente premente a propósito de conteúdos constitutivos de crítica, humor, sátira ou *counter-speech*, cuja qualificação não poderá ser pura e simplesmente delegada, por incompatibilidade axiológico-valorativa, em prestadores de serviços intermediários, desde logo pelos efeitos lesivos da supressão ou remoção em direitos fundamentais ponderosos. A esta luz, considera-se, pois, avisado associar a estes conteúdos, pelo menos *prima facie*, uma *dúvida razoável* que não poderá deixar de impactar na determinação e avaliação do alcance da diligência e da responsabilidade do prestador de serviços.

O critério do *prestador diligente*, referido em (*ii*), suscita mais dúvidas, porém. A primeira respeita à questão de saber se a avaliação exigida por conteúdos não *manifestamente* (i)legais afastará (por não ser esse o comportamento esperado de um prestador diligente) o recurso a meios automatizados.

A este propósito, importa salientar que o DSA não proíbe os prestadores de serviços de alojamento virtual de recorrerem a meios automatizados para efeitos da tomada de decisões[48]. Pelo contrário, aceitando esse como um cenário, não só possível, como prototípico[49], o legislador

---

[47] Cf. Ortolani, Pietro. "The Digital Services Act, Content Moderation and Dispute Resolution". *SSRN Electronic Journal*, 2023. Disponível em: http://dx.doi.org/10.2139/ssrn.4356598. Acesso em: 15 out. 2025, p. 21.

[48] Neste mesmo sentido, pese embora sustentando a necessidade de impor ou, pelo menos, incentivar a intervenção humana logo nesta fase, cf. Quintais, João Pedro et al. "Copyright Content Moderation in the European Union: State of the Art, Ways Forward and Policy Recommendations". *IIC – International Review of Intellectual Property and Competition Law*, v. 55, p. 175-176, 2024. Disponível em: http://dx.doi.org/10.1007/s40319-023-01409-5. Acesso em: 15 out. 2025.

[49] Como referem Franco, Mirko; Gaggi, Ombretta; Palazzi, Claudio E. "Integrating Content Moderation Systems with Large Language Models". *ACM Transactions on the Web*, v. 19, n. 2, art. 18, p. 1-21, 2025. Disponível em: https://doi.org/10.1145/3700789. Acesso em: 15 out. 2025, "*[a]lthough content moderation practices vary across different platforms, the underlying idea is common across all OSNs. In particular, on Facebook, potential violations undergo detection through artificial intelligence (AI) classifiers that scrutinize content during the upload process or are flagged by users who come across violating content. If the AI system identifies a potential violation with high confidence, swift removal may occur without additional checks. On the other hand, if the AI classifier detects potential violations with low confidence or if users report the content, human*

europeu limita-se a impor obrigações de transparência. A esta luz, duas interpretações afiguram-se-nos possíveis.

Por um lado, poderá entender-se que, por *prestador diligente* de serviços de alojamento virtual, o legislador europeu se estará a referir a uma subjetividade (humana ou coletiva-empresarial) que, pese embora sem conhecimentos bastantes para proceder à análise contextual e/ou jurídica necessária e exigida por determinados conteúdos, procede a um juízo de avaliação-qualificação não meramente automatizado. Assim perspetivada a norma, dir-se-á que o prestador diligente será aquele que, perante uma notificação com as referidas características, assegura uma "ação" não assente (em exclusivo) em meios automatizados, de onde resulta não poder um prestador de serviços que haja recorrido a meios automatizados ver a sua responsabilidade excluída, alegando a incapacidade de discernimento ou a ocorrência de erro imputável aos meios automatizados utilizados no processo de catalogação e reconhecimento da ilegalidade dos conteúdos.

Por outro lado, e em contraste, poderá assumir-se, pelo menos no que se refere aos fornecedores de plataformas em linha – obrigados, nos termos do artigo 20.º do DSA, à disponibilização de um sistema interno de gestão de reclamações –, que a diligência esperada do prestador de serviços não exigirá uma avaliação "humana", exclusiva, supervisionadora e/ou adicional àquela porventura resultante do recurso a meios automatizados. Isto, atenta a possibilidade de o notificante, insatisfeito com a referida qualificação (assente em meios automatizados) e consequente inércia do prestador, recorrer ao respetivo sistema interno de gestão de reclamações, a propósito do qual o DSA impõe já que as decisões "sejam tomadas sob supervisão de colaboradores devidamente qualificados, e não exclusivamente com base em meios automatizados"[50]. Note-se que, a adotar-se esta segunda leitura, a reclamação subsequente

---

reviewers scrutinize the content. In such cases, if the content breaches community standards, it is referred to a pool of paid reviewers and removal is contingent on consensos among multiple human moderators. Conversely, potential misinformation violations are routed to third-party certified and independent content moderators unless there is an imminent risk of violence or physical harm. In case of disagreement, users can appeal a moderation decision and the content will be reviewed by human moderators who can either uphold or overturn the initial verdict", p. 6.

[50] Cf. n.º 6 do artigo 20.º, DSA.

do notificante não poderá também deixar de impactar na avaliação e na determinação da diligência esperada do prestador de serviços.

A ausência de resposta concludente no Regulamento e a possibilidade teórica de sustentar as duas leituras são um argumento adicional em favor da interpretação parcimoniosa que aqui se sustenta para a norma do n.º 3 do artigo 16.º do DSA (conjugado com o artigo 6.º). Com efeito, se se sustentar que a presunção de conhecimento e a consequente obrigatoriedade de ação devem ver a sua aplicação circunscrita aos conteúdos cuja ilegalidade seja manifesta ou patente, é dizer, àqueles cuja qualificação (jurídica) se não veja dependente de uma análise contextual e/ou jurídica pormenorizada e, em particular, de um juízo humano qualificado, que um prestador de serviços diligente não está em condições de assegurar[51], a *ação* porventura justificativa do recurso a meios automatizados só se colocará em relação a *i*) conteúdos claramente problemáticos ou ilegais e/ou a *ii*) conteúdos cuja legalidade não ofereça dúvidas, caso em que – havendo ação ou inércia por parte do prestador, respetivamente – não poderá o recurso a meios automatizados ser de molde a *desqualificar* a sua diligência[52].

A distinção propugnada em texto encontra ancoragem na jurisprudência dos tribunais europeus e nacionais. A título de exemplo, considerou o Tribunal de Justiça, no caso *Glawischnig-Piesczek* (C-18/18)[53], que, na sequência da declaração de ilicitude de determinados conteúdos por um tribunal, poderá ser exigida às plataformas a remoção de conteúdos

---

[51] Além da incapacidade (jurídica), importa salientar os constrangimentos associados. Com G'sell, importa reconhecer que "complying with the DSA's procedural requirement will be arduous and expensive for platforms that handle millions of publications daily through automated systems" – cf. G'Sell, Florence. "The Digital Services Act (DSA): A General Assessment". *SSRN Electronic Journal*, 2023. Disponível em: http://dx.doi.org/10.2139/ssrn.4403433. Acesso em: 15 out. 2025.

[52] Sobre esta distinção, cf., *inter alia*, Stockinger, Andrea; Schäfer, Svenja; Lecheler, Sophie. "Navigating the Gray Areas of Content Moderation: Professional Moderators' Perspectives on Uncivil User Comments and the Role of (AI-based) Technological Tools". *New Media & Society*, v. 0, n. 0, 2023. Disponível em: https://doi.org/10.1177/14614448231190901. Acesso em: 15 out. 2025. *Vd.*, também, Tobi, Abraham. "Towards an Epistemic Compass for Online Content Moderation". *Philosophy & Technology*, v. 37, p. 109, 2024. Disponível em: https://doi.org/10.1007/s13347-024-00791-3. Acesso em: 15 out. 2025.

[53] Cf. Tribunal de Justiça da União Europeia. *Acórdão de 3 de outubro de 2019, Glawischnig-Piesczek v. Facebook Ireland Limited* (C-18/18, ECLI:EU:C:2019:821).

idênticos ou substancialmente equivalentes. É o mesmo que afirmar que à categoria de conteúdos manifestamente ilegais deverão reconduzir-se aqueles que tenham já sido, diretamente ou por proximidade, declarados como tal por órgãos jurisdicionais. Além destes, e em linha com o decidido pelo *Landgericht Düsseldorf* no processo *2a O 112/23*[54], relativo à divulgação de anúncios fraudulentos através do *Google Ads*, que imitavam o domínio legítimo de uma empresa notificante, poderá igualmente sustentar-se que o tipo de ilicitude em questão – uso indevido de marca e eventual fraude – deverá ser considerado manifestamente ilegal por uma plataforma que possa facilmente detetar tal prática mediante simples verificação técnica ou mecânica, sem necessidade de um juízo jurídico complexo.

Em contraste, conflitos entre particulares relativos a publicações ou conteúdos criados por utilizadores e partilhados em redes sociais situar-se-ão numa *zona cinzenta*, em que a ponderação entre liberdade de expressão e direitos de personalidade exige uma apreciação humana e pública mais desenvolvida, não exigível, sob pena de desproporcionalidade, a um prestador de serviços intermediários.

Em suma, e de tudo o exposto, resulta possível extrair, do n.º 3 do artigo 16.º, conjugado com o artigo 6.º do DSA, um particular dever para o prestador de serviços de alojamento virtual, de avaliação, qualificação e ponderação da informação putativamente ilegal notificada, nos termos do qual poderá aquele concluir no sentido da presença de *i*) conteúdo ilegal patente ou manifesto, caso em que deverá adotar as diligências exigidas pela alínea *b*) do n.º 1 do artigo 6.º do DSA; *ii*) conteúdo de nenhum modo configurável como ilegal, caso em que se deverá limitar a cumprir os deveres comunicacionais do artigo 16.º do DSA (além de outros aplicáveis); ou de *iii*) conteúdo cuja natureza (i)legal não é certa, cenário em que deverá proceder a uma indagação mais profunda, na sequência da qual poderá optar, ora pela suspensão (temporária) do acesso aos referidos conteúdos, ora pela respetiva manutenção, não devendo – à míngua de outros elementos de enquadramento – ser responsabilizado por qualquer uma das escolhas, contanto que assentes num exercício de ponderação proporcional e razoável, tendo em conta,

---

[54] Cf. Landgericht Düsseldorf. *2a O 112/23, Skinport GmbH c. Google Ireland Ltd.*, 15 jan. 2025.

quer o tipo de conteúdos em presença, quer interações subsequentes, seja com o(s) autor(es) da notificação, seja com os titulares da informação. A classificação assim encetada afigura-se necessária para evitar *falsos positivos* e *falsos negativos*, associados, respetivamente, a interferências indevidas na liberdade de expressão e de criação artística em linha, e, por outro lado, a eventuais lesões ou à perpetuação de interferências em direitos de personalidade[55].

4. Do alcance da responsabilidade da plataforma pelo conhecimento (presumido): subsídios para uma responsabilização proporcionada

Como se avançou, a notificação de elementos específicos de informação considerados "conteúdos ilegais" pela parte notificante exigirá, do prestador de serviços de alojamento virtual, (*i*) uma avaliação e qualificação, mais ou menos complexas, na sequência do que se lhe exigirá, também, (*ii*) a adoção de uma decisão quanto às medidas a adotar. Salientou-se, ainda, que, quer o exercício de qualificação, quer a "ação" assumida pelo prestador, quer, em particular, a escolha das concretas medidas a adotar, devem obediência e surgem limitadas por direitos fundamentais.

Em linha com o exposto, considera-se que também o *enforcement* do DSA pelas autoridades competentes deverá ser balizado pelos direitos fundamentais em presença, onde se incluem, além daqueles dos destinatários do serviço (seja na qualidade de notificantes, seja na qualidade de titulares de informações e/ou conteúdos potencialmente ilegais), também, os direitos do prestador do serviço.

Em particular, perante as dúvidas suscitadas pelas *zonas cinzentas* previamente introduzidas, nomeadamente no que se refere à natureza e à qualificação jurídica de alguns conteúdos como "ilegais", a atuação das autoridades responsáveis pela aplicação e execução do Regulamento dos Serviços Digitais e, em particular, a sua discricionariedade deverão ser

---

[55] Alertando para estes riscos, cf. Wei, Johnny Tian-Zheng; Zufall, Frederike; JIA, Robin. "Operationalizing Content Moderation 'Accuracy' in the Digital Services Act". *AI, Ethics, and Society (AIES) Journal*, v. 7, p. 1527–1538, 16 out. 2024. Disponível em: https://doi.org/10.1609/aies.v7i1.31744. Acesso em: 15 out. 2025, p. 1530ff.

atuadas, em tais cenários, com particular autocontenção e proporcionalidade face ao *status* de dúvida e à dimensão dos poderes de *enforcement* que lhes são reconhecidos.

Isto *de iure constituto*. Já *de iure constituendo*, considera-se haver margem para explorar, promover e prever – com particulares vantagens sobretudo em cenários de *dúvida razoável* – mecanismos e canais de diálogo e de transmissão de informações, entre o prestador de serviços intermediários e as autoridades (administrativas ou jurisdicionais) competentes, estas, em melhor posição para auxiliar o prestador no exercício de qualificação e na determinação das medidas exigíveis para dar cumprimento ao DSA.

Concretizando muito brevemente.

4.1. A proporcionalidade e a razoabilidade da aplicação prática da responsabilidade pelo conhecimento (presumido): entre o public *e o* private enforcement

Como se antecipou, o DSA é um ato normativo que compreende dois pilares principais. Por um lado, um elenco de isenções condicionais de responsabilidade. Por outro lado, um conjunto de obrigações de devida diligência. O não cumprimento das condições previstas nos artigos 4.º a 6.º do DSA afasta o benefício da isenção de responsabilidade como *porto seguro*, permitindo, pois, a responsabilização dos prestadores de serviços intermediários por informações (ilegais) de terceiros, transmitidas, acedidas e/ou armazenadas nos respetivos canais, interfaces ou plataformas. Já a violação dos deveres de devida diligência consagrados no DSA, sujeita-o (prestador) a medidas sancionatórias de execução, entre as quais coimas, adotadas pelas autoridades competentes (*public enforcement*), a que acrescem eventuais pretensões indemnizatórias deduzidas por potenciais lesados (*private enforcement*). Esta responsabilidade (pela violação de deveres próprios) é relativamente independente da (isenção de) responsabilidade relativamente aos conteúdos de terceiros[56].

---

[56] Sobre a articulação entre as duas dimensões, num sistema assente num *mixed enforcement*, cf. Husovec, Martin. "The DSA as a Mixed Enforcement System". In: *Principles of the Digital Services Act*. Oxford: Oxford University Press, 2024. Disponível em: http://dx.doi.org/10.1093/law-ocl/9780192882455.003.0019. Acesso em: 15 out. 2025.

Assim, e pese embora as isenções condicionais de responsabilidade previstas no DSA não possam ser entendidas como "uma base positiva para determinar quando um prestador pode ser responsabilizado"[57], certo é que o prestador de serviços intermediários se vê sujeito a diferentes frentes possíveis de responsabilização, quer em contexto de *public enforcement*, regulado no DSA e nas legislações nacionais de execução, quer em contexto de *private enforcement*, também reconhecido no considerando 121[58] e no artigo 54.º do DSA. Deste último resulta que, "[n]os termos do direito da União e nacional, os destinatários do serviço têm o direito de pedir uma indemnização aos prestadores de serviços intermediários no que diz respeito a quaisquer perdas ou danos sofridos devido a uma violação, por parte desses prestadores, das obrigações que lhes incumbem por força do presente regulamento".

Os dois pilares do DSA suprarreferidos – isenções condicionais de responsabilidade e deveres de devida diligência – não são, porém, apesar da sua autonomia, nem excludentes nem incomunicáveis. Exemplificativa disso mesmo é a norma em estudo, resultante da conjugação do n.º 3 do artigo 16.º com o n.º 1 do artigo 6.º do DSA.

Por isso mesmo, e em linha com as dúvidas de hermenêutica assinaladas *supra*, considera-se que igual parcimónia será também de exigir na respetiva aplicação e execução.

Em particular, deverão as autoridades competentes (coordenadores dos serviços digitais ou Comissão Europeia)[59] abster-se de um

---

[57] Cf. considerando 17, DSA.

[58] De acordo com o considerando 121 do DSA, "[s]em prejuízo das disposições relativas à isenção de responsabilidade previstas no presente regulamento no que respeita às informações transmitidas ou armazenadas a pedido de um destinatário do serviço, os prestadores de serviços intermediários deverão ser responsáveis pelos danos causados aos destinatários do serviço devido a uma violação das obrigações estabelecidas no presente regulamento para esses prestadores. A indemnização desses danos deverá estar em conformidade com as regras e os procedimentos estabelecidos na legislação nacional aplicável e sem prejuízo de outras possibilidades de reparação disponíveis ao abrigo das regras de defesa dos consumidores".

[59] Sobre a articulação entre diferentes entidades responsáveis pelo *enforcement* do DSA, cf. Van Cleynenbreugel, Pieter; Mattioli, Pietro. "Digital Services Coordinators and Other Competent Authorities in the Digital Services Act: Streamlined Enforcement Coordination Lost?". *European Law Blog*, 2023. Disponível em: http://dx.doi.org/10.21428/9885764c.f18f26b8. Acesso em: 15 out. 2025. Também Wilman, Folkert; Kalėda, Saulius Lukas; Loewenthal, Paul-John. "Implementation, Cooperation, Penalties and Enforcement". In:

enforcement punitivo, optando, antes, por uma estratégia dialógica e pedagógica. É dizer, ao invés da imediata imposição de sanções [nos termos dos artigos 51.º, n.º 2, als. *c*) e *d*), 52.º e 74.º do DSA] pela não remoção de conteúdos sinalizados como potencialmente ilegais por destinatários do serviço ou outras entidades, deverão as autoridades competentes, perante a existência de dúvida razoável quanto à natureza e qualificação da informação ou conteúdos, optar por *esclarecer* o fornecedor e concretizar as medidas necessárias para assegurar o cumprimento do Regulamento [nos termos dos artigos 51.º, n.º 2, al. *b*), e 73.º do DSA].

E nem se avance, contra o exposto, a inexistência de base normativa no DSA. Desde logo, está-se perante Regulamento que de nenhum modo impõe a aplicação de sanções pecuniárias em caso de incumprimento das suas disposições. Em contraste, veja-se que, de acordo com os considerandos 115 a 117 do DSA, os poderes de investigação e de execução reconhecidos aos coordenadores dos serviços digitais e, sendo caso disso, a outras autoridades competentes deverão estar enquadrados por condições e limites pormenorizados, e, bem assim, exercidos de forma proporcionada "nomeadamente, à natureza e ao dano global efetivo ou eventual causado pela infração ou suspeita de infração"[60].

Em particular no que à aplicação de sanções se refere (que não se esgotam nas coimas e sanções pecuniárias compulsórias), resulta do Regulamento que os Estados-Membros deverão assegurar a sua efetividade, proporcionalidade e dissuasão, "tendo em conta a natureza, a gravidade, a recorrência e a duração da violação, dado o interesse público visado, o âmbito e o tipo de atividades realizadas, bem como a capacidade económica do infrator"[61].

Idênticas exigências de proporcionalidade se erguem, por seu turno e também, à Comissão Europeia, a quem são reconhecidos poderes de investigação e execução exclusivos em relação aos fornecedores de

---

The EU Digital Services Act. Oxford: Oxford University Press, 2024. Disponível em: http://dx.doi.org/10.1093/law/9780198892847.003.0005. Acesso em: 15 out. 2025. Em particular, sobre as dificuldades de um *enforcement* consistente, cf. NANNINI, Luca et al. "Beyond Phase-in: Assessing Impacts on Disinformation of the EU Digital Services Act". *AI and Ethics*, 2024. Disponível em: http://dx.doi.org/10.1007/s43681-024-00467-w. Acesso em: 15 out. 2025.

[60] *Vd*., também, artigo 51.º, DSA.

[61] *Vd*., também, artigo 52.º, DSA.

plataformas em linha de muito grande dimensão (e aos motores de pesquisa em linha de muito grande dimensão). Com efeito, lê-se no considerando 140 do DSA que os referidos poderes deverão ser exercidos "no pleno respeito do direito fundamental de ser ouvido e de ter acesso ao processo no âmbito de um procedimento de execução, do princípio da proporcionalidade e dos direitos e interesses das partes em causa". Sendo que, no que se refere em particular à aplicação de coimas e de sanções pecuniárias compulsórias, previstas no artigo 74.º do DSA, também o considerando 144 apela a particulares exigências de proporcionalidade.

Em suma, os desafios suscitados por conteúdos de (i)legalidade não manifesta ou patente aconselham e exigem a preferência das autoridades responsáveis por uma abordagem de *enforcement* dialógica, com ancoragem normativa clara.

Isto, no plano do *public enforcement*.

Já no plano do *private enforcement*, a responsabilidade por uma aplicação razoável e proporcional do Regulamento deslocar-se-á, em particular, para os juízes e demais autoridades jurisdicionais (porventura, o Ministério Público), a quem caiba ajuizar do mérito de uma ação de responsabilidade civil movida contra um prestador de serviços, por incumprimento dos seus deveres ao abrigo do DSA, e, em particular, por incumprimento do seu dever de "ação", na sequência da notificação por um destinatário do serviço.

Julga-se, desde logo por imperativos de coerência político-jurídica, ser, também nesta sede, de *levar a sério* o disposto na parte final do n.º 3 do artigo 6.º do DSA, cabendo então aos referidos atores judiciários avaliar se a particular notificação, dirigida ao prestador, era de molde – tendo em conta o seu conteúdo e, em particular, o tipo de informação em presença – a permitir a um "prestador diligente de alojamento virtual identificar a ilegalidade da atividade ou das informações em causa sem um exame jurídico pormenorizado"[62].

Dos atores judiciários exigir-se-á, para o efeito, um conhecimento cabal do enquadramento normativo aplicável, e a exploração das diferentes vias potencialmente abertas aos notificantes[63]. *Inter alia*, será

---

[62] Cf. n.º 3 do artigo 16.º, DSA.

[63] Com efeito, importa recordar que, na sequência da inércia do prestador, aos destinatários da decisão será possível apresentar uma reclamação (nos termos do artigo 20.º do DSA),

relevante aferir se, na sequência de uma notificação e consequente inércia do prestador de serviços, o notificante recorreu ao sistema interno de gestão de reclamações consagrado no artigo 20.º do DSA (quando aplicável e se disponível), ou se, *a contrario*, na sequência de uma decisão restritiva (de supressão, restrição ou bloqueio do acesso a uma informação, conta ou elemento do serviço), o próprio interessado (titular da informação) reagiu através do mesmo sistema. E, em ambos os cenários, em que termos se desenvolveu esta interação e que motivos e /ou preocupações a nortearam.

Adicionalmente, deverá o tribunal apreciar também se, ao invés do mecanismo de notificação oferecido pelo prestador de serviços, o destinatário do serviço optou diretamente pela eventual citação da plataforma para uma ação (coletiva) de responsabilidade civil, elemento que poderá indicar um eventual recurso abusivo à ação. Com efeito, pese embora resulte do considerado 59 do DSA que "[o]s destinatários do serviço deverão poder escolher entre o mecanismo interno de reclamação, a resolução extrajudicial de litígios e a possibilidade de intentar, em qualquer fase, processos judiciais", esta alternatividade está claramente pensada para uma fase posterior a uma notificação ao abrigo do artigo 16.º do DSA, o que se compreende, tendo em conta que a ação de responsabilidade civil terá por objeto aferir do (in)cumprimento de uma qualquer "ação" exigida pelo DSA pela plataforma, o que pressuporá um conhecimento prévio, obtido através de qualquer uma das vias já atrás referidas. Assim, crê-se dever o recurso direto a uma ação de responsabilidade civil, sem recurso prévio ao mecanismo do artigo 16.º do DSA, ser elemento relevante em contexto do *private enforcement*, no qual haverá que apurar se esse recurso direto teve por fundamento a inexistência de canais de notificação e ação (o que representa o incumprimento do DSA pela plataforma) e/ou, pelo contrário, a simples inércia consciente e/ou a eventual má-fé ou abuso do direito à ação por parte do putativo lesado.

---

recorrer a um organismo de resolução extrajudicial de litígios (nos termos do artigo 21.º do DSA), ou intentar, em qualquer fase, um processo judicial, possibilidades de que os notificantes deverão ser informados pelo prestador de serviços. Por seu turno, na sua resposta fundamentada aos notificantes, o prestador de serviços poderá e deverá justificar a sua decisão, com base na ilegalidade não manifesta ou patente dos conteúdos em questão – cf. considerandos 55 e 59, DSA.

Listadas algumas indicações que se julgam relevantes para um *enforcement* efetivo, mas, e bem assim, proporcional do DSA, cumpre agora terminar com um conjunto de recomendações sobre as vantagens do diálogo entre prestadores de serviços intermediários e autoridades (públicas) competentes.

4.2. Repensar o diálogo e a repartição de responsabilidades entre o prestador de serviços intermediários e as autoridades competentes

Segundo o artigo 18.º do DSA, aplicável aos prestadores de serviços de alojamento virtual, incluindo as plataformas em linha, "[s]empre que um prestador de serviços de alojamento virtual tome conhecimento de qualquer informação que levante suspeitas de que ocorreu, está a ocorrer ou é suscetível de ocorrer um crime que envolva uma ameaça à vida ou à segurança de uma ou várias pessoas, o prestador de serviços de alojamento virtual informa imediatamente da sua suspeita as autoridades policiais ou judiciárias do ou dos Estados-Membros em causa e fornece todas as informações pertinentes disponíveis"[64].

Pese embora o reporte de conteúdos às autoridades competentes estar previsto, como dever especial, apenas e tão-só para informações que levantem suspeitas de crimes, julga-se que, imbuídos no mesmo espírito, nada haveria a obstar à aplicação de solução próxima e à introdução de canais de diálogo entre prestadores de serviços e autoridades competentes, a propósito de conteúdos cuja ilegalidade não seja manifesta ou patente e/ou cuja legalidade também se não possa asseverar acima de toda a dúvida razoável, isto é, para os *casos cinzentos*.

Em particular, a prever-se um tal canal, daí resultaria que, entre as possibilidades abertas ao prestador, para – perante a comunicação por um destinatário do serviço relativamente à presença de conteúdos ilegais em linha – preservar a sua isenção de responsabilidade, estaria, também, a via da notificação dos referidos conteúdos e situação fática às autoridades competentes[65], comunicação essa que, em caso de dúvida,

---

[64] *Vd.*, também, considerando 56, DSA.

[65] Neste sentido, Inserra, David. "A Guide to Content Moderation for Policymakers". *Policy Analysis*, n. 974, 2024. Disponível em: https://www.cato.org/policy-analysis/guide-content-moderation-policymakers. Acesso em: 15 out. 2025.

poderia ser considerada "suficiente" para dar cumprimento às condições previstas no artigo 6.º do DSA, e, com elas, para a aplicação da isenção de responsabilidade aí prevista.

A dúvida ínsita à qualificação de um conteúdo em linha como (i)legal e, em particular, o alcance das eventuais ações adotadas na sua sequência (como forma de garantir a preservação da isenção de responsabilidade) – potencialmente lesivas de direitos fundamentais (seja do lesado pela informação, seja do autor da informação) – justificam a referida articulação com as autoridades públicas, nacionais ou europeias competentes[66].

Além de se estar perante elemento relevante e idóneo a atestar a boa-fé do prestador – e demonstrativo de como não age de outro modo, apenas em razão de dúvida razoável quanto à (i)legalidade das informações (fator que deverá ser tido em conta na avaliação da respetiva diligência) –, o diálogo a que se alude revestiria também importância inegável do prisma do interesse público (também ele na base do DSA)[67],

---

[66] Com efeito, apesar de os dados constantes do *DSA Transparency Database* suportarem alguma autocontenção das plataformas no que se refere à adoção de restrições adotadas *motu proprio* (e não na sequência de uma notificação), e fundadas na violação ou incompatibilidade com os termos e condições da própria plataforma – o que mitiga os riscos associados aos falsos positivos (*overblocking*) –, a verdade é que, além de nada permitir assegurar a manutenção dessa tendência, importa também prevenir os efeitos nefastos dos falsos negativos (*underblocking*), isto é, a preservação em linha de informações ou conteúdos danosos, designadamente para os direitos de personalidade. Tudo isto reforça a necessidade de uma maior articulação com as autoridades públicas, em especial nos casos em que a determinação da ilicitude dos conteúdos dependa de uma avaliação jurídica contextual. Sobre o *DSA Transparency Database*, cf. https://transparency.dsa.ec.europa.eu/. Também Kaushal, Rishabh; Van De Kerkhof, Jacob; Goanta, Catalina; Spanakis, Gerasimos; Iamnitchi, Adriana. "Automated Transparency: A Legal and Empirical Analysis of the Digital Services Act Transparency Database". In: *Proceedings of the 2024 ACM Conference on Fairness, Accountability, and Transparency (FAccT '24)*. New York: Association for Computing Machinery, 2024. p. 1121--1132. Disponível em: https://doi.org/10.1145/3630106.3658960. Acesso em: 15 out. 2025. Sobre as insuficiências da base de dados, cf. Trujillo, Amaury; Fagni, Tiziano; Cresci, Stefano. "The DSA Transparency Database: Auditing Self-reported Moderation Actions by Social Media". In: *Proceedings of the 28th ACM Conference on Computer-Supported Cooperative Work and Social Computing (CSCW '25)*. New York: Association for Computing Machinery, 2025. Disponível em: https://doi.org/10.1145/3711085. Acesso em: 15 out. 2025.

[67] Cf. considerando 40, DSA.

ínsito ao combate à presença e à disseminação de conteúdos ilegais em linha (independentemente do tipo de ilegalidade), e justificativo de um envolvimento próximo, seja das autoridades nacionais competentes do Estado-Membro a que o conteúdo ilegal se refere (ou no qual o prestador ou o seu representante legal estejam localizados)[68], seja da Comissão Europeia (no caso dos fornecedores de plataformas em linha de muito grande dimensão)[69].

Está-se, reconhece-se, perante solução válida de *iure constituendo*, porquanto naturalmente carecida (na sua efetivação prática) do desenho e da implementação de canais de diálogo que a permitam, e que, no que se refere à opção entre a notificação a uma das autoridades competentes para a supervisão e execução do Regulamento dos Serviços Digitais e/ /ou a uma outra autoridade administrativa, policial ou judiciária, sempre exigirá o esclarecimento de questões de competência. Com efeito, ainda que em relação às primeiras (sobretudo, os coordenadores dos serviços digitais), a possibilidade aqui sustentada possa até não suscitar problemas[70], importa não negligenciar a eventual necessidade e premência da intervenção de outras entidades e a possibilidade de a elas recorrer, nos termos de um enquadramento próprio[71].

Pese embora carecida de concretização ulterior, a solução aqui avançada – diálogo preventivo – antecipa os problemas reconhecidos, *inter alia*, no considerando 39 do DSA, ao prever que "o presente regulamento não deverá impedir as autoridades judiciárias ou administrativas nacionais competentes de emitir, com base no direito da União ou nacional aplicável, uma decisão de reposição de conteúdos sempre que esses conteúdos estiverem em conformidade com os termos e condições

---

[68] Cf. artigo 49.º e ss., DSA.

[69] Cf. artigo 64.º e ss., DSA.

[70] Pese embora focado nos direitos dos destinatários do serviço, o DSA consagra um direito à apresentação de reclamações no artigo 53.º do DSA, cujo canal de efetivação permitirá, de igual modo, a um prestador de serviços apresentar uma preocupação fundamentada.

[71] Desde logo, há propostas no sentido da centralização da apreciação de casos complexos na Comissão Europeia ou no Comité Europeu dos Serviços Digitais, por forma a evitar uma aplicação divergente do DSA, motivada por diferentes tradições nacionais de tutela do discurso e de proteção da liberdade de expressão. Neste sentido, cf. Van De Kerkhof, Jacob. "The DSA's Tower of Babel: On Digital Services Coordinators and Freedom of Expression". *European Journal of Risk Regulation*, p. 1-26, 2025. Disponível em: https://doi.org/10.1017/err.2025.10034. Acesso em: 15 out. 2025.

do prestador do serviço intermediário, mas tenham sido erradamente considerados ilegais por esse prestador e tenham sido suprimidos".

Pergunta-se: ao invés da manutenção de um quadro de incerteza e da promoção de um crescente risco de responsabilização, propício à "remoção por defeito", não exigirão os direitos fundamentais, pelo contrário, a segurança jurídica e a centralização da resolução dos "casos difíceis" nas autoridades judiciárias ou administrativas nacionais competentes, como dita o artigo 9.º do DSA?

Por outras palavras, ao invés da adoção de decisões de reposição (que pressupõem e convivem com a remoção errónea de conteúdos), considera-se necessário *reabilitar* a responsabilidade das autoridades competentes pela adoção das competentes decisões de atuação-remoção, em termos que não só não afetam a isenção de responsabilidade dos prestadores de serviços como contribuem para um ambiente em linha seguro e respeitador dos direitos fundamentais, evitando a censura das liberdades e garantindo, em simultâneo, a proteção de todos os direitos fundamentais.

Em suma, sendo a garantia e a efetivação dos direitos fundamentais uma responsabilidade de todos, não poderá a sua atuação e proteção em linha deixar de assentar numa efetiva partilha de responsabilidades[72], e não numa pura e simples delegação de tarefas.

5. Conclusões

O quadro de isenções de responsabilidade (pelos conteúdos de terceiros) e o conjunto de obrigações de devida diligência consagradas no Regulamento dos Serviços Digitais representam as suas duas faces de Jano. São questões inconfundíveis, porém. Com efeito, uma coisa é aferir da (não) responsabilidade do prestador pelos conteúdos ilegais de terceiros. E outra, diferente, é cuidar da sua responsabilidade pelo (in)cumprimento das obrigações de devida diligência previstas no Regulamento. Porque assim é, a regra é a de que o incumprimento das obrigações de

[72] E não apenas dos riscos. Sobre a necessária partilha de riscos entre diferentes *stakeholders*, a propósito do sistema dos artigos 34.º e 35.º, DSA, cf. HUSOVEC, Martin. "General Risk Management". In: *Principles of the Digital Services Act*. Oxford: Oxford University Press, 2024. Disponível em: http://dx.doi.org/10.1093/law-ocl/9780192882455.003.0015. Acesso em: 15 out. 2025.

devida diligência é fonte de uma responsabilidade específica – pela qual podem, aliás, ser aplicadas sanções e avançadas pretensões indemnizatórias –, inconfundível com a responsabilidade por conteúdos e informações ilegais em linha de terceiros.

A diferença não elimina, porém, a existência de pontos de contacto.

Este artigo debruçou-se sobre um em particular – aquele que resulta da leitura conjugada do n.º 3 do artigo 16.º com o n.º 1 do artigo 6.º do DSA, preceitos estes que, estabelecendo uma presunção de conhecimento da presença de conteúdos ilegais em linha, por um prestador de serviços disso alertado por um destinatário do serviço, exigem desse mesmo prestador uma atuação diligente no sentido da supressão ou desativação do acesso a esses conteúdos "ilegais". Procurou-se, para esta *norma*, uma proposta de interpretação proporcional e razoável, conjugada com a necessidade de uma sua aplicação e execução de "geometria variável".

Retomando as questões de estudo avançadas, considera-se que a presunção de conhecimento, consagrada no n.º 3 do artigo 16.º do DSA, configura uma presunção relativa, podendo o prestador, ainda quando indevidamente acionado, avançar elementos demonstrativos de dúvida razoável quanto à qualificação da informação como ilegal e, portanto, capazes de afastar a sua responsabilidade pela "ação" (ou omissão).

Por seu turno, e em linha com uma hermenêutica proporcional e razoável da norma, sustenta-se a circunscrição do potencial aplicativo da presunção de conhecimento a *i*) notificações suficientes, precisas, adequadamente fundamentadas e estritamente direcionadas (a conteúdos específicos) e *ii*) respeitantes a conteúdos cuja ilegalidade seja manifesta ou patente a um prestador diligente. Em particular, considera-se que somente em tais casos será de exigir a atuação do prestador de serviços, sob a ameaça de, em caso de inércia, não poder beneficiar da isenção de responsabilidade do artigo 6.º do DSA.

A esta luz, e além da importância de uma análise atenta da notificação avançada pelo destinatário do serviço, considera-se que o n.º 3 do artigo 16.º e o próprio artigo 6.º do DSA suportam a existência de um regime diferenciado, distinguindo entre *i*) conteúdos ilegais evidentes ou manifestos, *ii*) conteúdos cuja legalidade não oferece dúvida razoável e *iii*) conteúdos cuja qualificação como (i)legais se afigura carecida de uma avaliação que o prestador poderá não estar em condições de fazer (no imediato ou sem mais).

Além da proposta hermenêutica sustentada, tratou-se, também, do respetivo impacto na aplicação e execução do Regulamento dos Serviços Digitais, quer em contexto de *public enforcement*, quer em contexto de *private enforcement*, considerando relevante a exigência de uma análise contextual séria, por parte de todos os atores públicos envolvidos – desde coordenadores dos serviços digitais e Comissão Europeia, às autoridades judiciárias e administrativas nacionais competentes.

O texto conclui, propondo a necessidade – sobretudo em *cenários cinzentos*, isto é, de (i)legalidade não manifesta ou evidente – de um maior envolvimento, articulação e diálogo entre prestadores de serviços e autoridades administrativas e judiciárias competentes, às quais se considera dever caber, em última linha, a responsabilidade por decisões de qualificação (jurídica) e pela adjudicação de direitos fundamentais, funções que não devem caber, nem podem ser exigidas, sem mais ou em termos desproporcionais, a um prestador de serviços intermediários.

Segundo parte da doutrina, o DSA é exemplificativo de um ato normativo que procura alcançar compromissos extremamente difíceis, complexos e, em alguns casos, verdadeiramente dilemático-diabólicos, seja entre o combate à disseminação de conteúdos ilegais e à desinformação em linha, por um lado, e a liberdade de expressão e de informação, por outro; seja entre a proteção dos utilizadores e da sociedade em geral, por um lado, e os ónus incidentes sobre os prestadores de serviços, com efeitos conexos na concorrência e na inovação, por outro[73].

Foi precisamente com a pretensão de explorar um domínio onde as referidas tensões se fazem sentir de forma particular e qualificada que se procuraram aqui respostas capazes de nortear um enquadramento, não só proporcional e razoável perante os diferentes interesses em presença, como, e bem assim, compatível com o objetivo de garantir a efetivação dos direitos fundamentais em linha. A final, considera-se necessário *devolver* ao Estado uma função que não poderá ser pura e simplesmente entregue a privados – a proteção dos direitos fundamentais –, que

---

[73] Sobre a questão, cf. WILMAN, Folkert. "Het voorstel voor de Digital Services Act". *Nederlands Tijdschrift voor Europees Recht*, v. 27, n. 1-2, p. 27, 2021. Disponível em: http://dx.doi.org/10.5553/nter/138241202021027102002. Acesso em: 15 out. 2025.

importa garantir não (apenas) por via de um sistema automatizado[74], mas através da sensibilidade humana e da intervenção pública.

Referências

ARTICLE 19. *Content Moderation and Freedom of Expression Handbook*. 2023. Disponível em: https://www.article19.org/wp-content/uploads/2023/08/SM4P-Content-moderation-handbook-9-Aug-final.pdf. Acesso em: 15 out. 2025.

ARTICLE 19. *Jillian York: The Global Impact of Content Moderation*. 2020. Disponível em: https://www.article19.org/resources/the-global-impact-of-content-moderation/. Acesso em: 15 out. 2025.

Bleakley, Paul; Martellozzo, Elena; DeMarco, Jeffrey. "Moderating Online Child Sexual Abuse Material (CSAM): Does Self-Regulation Work, or Is Greater State Regulation Needed?". *European Journal of Criminology*, v. 21, n. 2, 2023. Disponível em: https://doi.org/10.1177/14773708231181361. Acesso em: 15 out. 2025.

Drolsbach, Chiara Patricia; Pröllochs, Nicolas. "Content Moderation on Social Media in the EU: Insights from the DSA Transparency Database". In: *Proceedings of the ACM Web Conference 2024 (WWW '24)*. New York: Association for Computing Machinery, 2024. Disponível em: http://dx.doi.org/10.1145/3589335.3651482. Acesso em: 15 out. 2025.

---

[74] Traçando uma boa evolução, cf. Bleakley, Paul; Martellozzo, Elena; DeMarco, Jeffrey. "Moderating Online Child Sexual Abuse Material (CSAM): Does Self-Regulation Work, or Is Greater State Regulation Needed?". *European Journal of Criminology*, v. 21, n. 2, 2023. Disponível em: https://doi.org/10.1177/14773708231181361. Acesso em: 15 out. 2025, onde se assinala que, "*[w]hen the Internet was still in its infancy, self-regulation by online service providers like social media platforms was seen as a practical solution to the challenges posed by the new digital landscape. Without clarity on basic questions like how to manage competing, overlapping, or otherwise unclear jurisdiction – as well as an overarching philosophical preference for an uber-liberal 'cyberanarchy' – permitting cyberspace to operate with minimal interference was not just an appealing option but, in many ways, the only viable option available at the time. However, as the role of the Internet has exponentially expanded to the point that it has become such an intrinsic element of society, this self-regulation has had a substantial impact on public safety. Where before the state and its law enforcement apparatus was responsible for performing this role, the balance has tipped in the Internet age to the point that social media platforms like Facebook, Twitter and TikTok now serve as the first line of defence against harmful content like CSAM, with CMs serving as first responders for the digital age (Roberts, 2014; Bellanova and De Goede, 2021)*", pp. 244-245.

ENARSSON, Therese. "Navigating Hate Speech and Content Moderation under the DSA: Insights from ECtHR Case Law". *Information & Communications Technology Law*, v. 33, n. 3, p. 384-401, 2024. Disponível em: https://doi.org/10.1080/13600834.2024.2395579. Acesso em: 15 out. 2025.

FRANCO, Mirko; GAGGI, Ombretta; PALAZZI, Claudio E. "Integrating Content Moderation Systems with Large Language Models". *ACM Transactions on the Web*, v. 19, n. 2, art. 18, p. 1-21, 2025. Disponível em: https://doi.org/10.1145/3700789. Acesso em: 15 out. 2025.

G'SELL, Florence. "The Digital Services Act (DSA): A General Assessment". *SSRN Electronic Journal*, 2023. Disponível em: http://dx.doi.org/10.2139/ssrn.4403433. Acesso em: 15 out. 2025.

HUANG, Tao. "Content Moderation by LLM: From Accuracy to Legitimacy". *Artificial Intelligence Review*, v. 58, art. 320, 2025. Disponível em: https://doi.org/10.1007/s10462-025-10820-9. Acesso em: 15 out. 2025.

HUSOVEC, Martin. "Liability Exemptions: Specific Services". In: *Principles of the Digital Services Act*. Oxford: Oxford University Press, 2024. Disponível em: http://dx.doi.org/10.1093/law-ocl/9780192882455.003.0007. Acesso em: 15 out. 2025.

HUSOVEC, Martin. "Content Moderation: Outline". In: *Principles of the Digital Services Act*. Oxford: Oxford University Press, 2024. Disponível em: http://dx.doi.org/10.1093/law-ocl/9780192882455.003.0010. Acesso em: 15 out. 2025.

HUSOVEC, Martin. "General Risk Management". In: *Principles of the Digital Services Act*. Oxford: Oxford University Press, 2024. Disponível em: http://dx.doi.org/10.1093/law-ocl/9780192882455.003.0015. Acesso em: 15 out. 2025.

HUSOVEC, Martin. "The DSA as a Cornerstone of the EU Single Market". In: *Principles of the Digital Services Act*. Oxford: Oxford University Press, 2024. Disponível em: http://dx.doi.org/10.1093/law-ocl/9780192882455.003.0017. Acesso em: 15 out. 2025.

HUSOVEC, Martin. "The DSA as a Mixed Enforcement System". In: *Principles of the Digital Services Act*. Oxford: Oxford University Press, 2024. Disponível em: http://dx.doi.org/10.1093/law-ocl/9780192882455.003.0019. Acesso em: 15 out. 2025.

INSERRA, David. "A Guide to Content Moderation for Policymakers". *Policy Analysis*, n. 974, 2024. Disponível em: https://www.cato.org/policy-analysis/guide-content-moderation-policymakers. Acesso em: 15 out. 2025.

KAUSHAL, Rishabh; VAN DE KERKHOF, Jacob; GOANTA, Catalina; SPANAKIS, Gerasimos; IAMNITCHI, Adriana. "Automated Transparency: A Legal and Empirical Analysis of the Digital Services Act Transparency Database". In: *Proceedings of the 2024 ACM Conference on Fairness, Accountability, and Transparency*

(FAccT '24). New York: Association for Computing Machinery, 2024. p. 1121-1132. Disponível em: https://doi.org/10.1145/3630106.3658960. Acesso em: 15 out. 2025.

Landgericht Düsseldorf. *2a O 112/23, Skinport GmbH c. Google Ireland Ltd.*, 15 jan. 2025.

Lookingbill, Valerie; LE, Kimanh. "There's Always a Way to Get Around the Guidelines: Nonsuicidal Self-Injury and Content Moderation on TikTok". *Social Media + Society*, v. 10, n. 2, 2024. Disponível em: https://doi.org/10.1177/20563051241254371. Acesso em: 15 out. 2025.

Nannini, Luca et al. "Beyond Phase-in: Assessing Impacts on Disinformation of the EU Digital Services Act". *AI and Ethics*, 2024. Disponível em: http://dx.doi.org/10.1007/s43681-024-00467-w. Acesso em: 15 out. 2025.

Ortolani, Pietro. "The Digital Services Act, Content Moderation and Dispute Resolution". *SSRN Electronic Journal*, 2023. Disponível em: http://dx.doi.org/10.2139/ssrn.4356598. Acesso em: 15 out. 2025.

Peguera, Miquel. "The Platform Neutrality Conundrum and the Digital Services Act". *IIC – International Review of Intellectual Property and Competition Law*, 2022. Disponível em: http://dx.doi.org/10.1007/s40319-022-01205-7. Acesso em: 15 out. 2025.

Quintais, João Pedro et al. "Copyright Content Moderation in the European Union: State of the Art, Ways Forward and Policy Recommendations". *IIC – International Review of Intellectual Property and Competition Law*, v. 55, p. 175-176, 2024. Disponível em: http://dx.doi.org/10.1007/s40319-023-01409-5. Acesso em: 15 out. 2025.

Quintais, João Pedro; Appelman, Naomi; Ó Fathaigh, Ronan. "Using Terms and Conditions to Apply Fundamental Rights to Content Moderation". *German Law Journal*, v. 24, p. 1-20, 2023. Disponível em: http://dx.doi.org/10.1017/glj.2023.53. Acesso em: 15 out. 2025.

Rocha, Tiago Morais. "Digital Services Act: Towards the Digital Rule of Law". In: Enes, Graça; Neves, Inês; Rocha, Tiago Morais (eds.). *A Digital Europe for Citizens (DigEUCit 2023)*. Cham: Springer, 2026. p. 189-208. Disponível em: https://doi.org/10.1007/978-3-032-02500-5_11. Acesso em: 15 out. 2025.

Stockinger, Andrea; Schäfer, Svenja; Lecheler, Sophie. "Navigating the Gray Areas of Content Moderation: Professional Moderators' Perspectives on Uncivil User Comments and the Role of (AI-based) Technological Tools". *New Media & Society*, v. 0, n. 0, 2023. Disponível em: https://doi.org/10.1177/14614448231190901. Acesso em: 15 out. 2025.

Tobi, Abraham. "Towards an Epistemic Compass for Online Content Moderation". *Philosophy & Technology*, v. 37, p. 109, 2024. Disponível em: https://doi.org/10.1007/s13347-024-00791-3. Acesso em: 15 out. 2025.

TRIBUNAL DE JUSTIÇA DA UNIÃO EUROPEIA. *Acórdão de 3 de outubro de 2019, Glawischnig-Piesczek v. Facebook Ireland Limited* (C-18/18, ECLI:EU:C:2019:821).

TRUJILLO, Amaury; FAGNI, Tiziano; CRESCI, Stefano. "The DSA Transparency Database: Auditing Self-reported Moderation Actions by Social Media". In: *Proceedings of the 28th ACM Conference on Computer-Supported Cooperative Work and Social Computing (CSCW '25)*. New York: Association for Computing Machinery, 2025. Disponível em: https://doi.org/10.1145/3711085. Acesso em: 15 out. 2025.

TURILLAZZI, Aina et al. "The Digital Services Act: An Analysis of Its Ethical, Legal, and Social Implications". *Law, Innovation and Technology*, v. 15, p. 1-94, 2023. Disponível em: http://dx.doi.org/10.1080/17579961.2023.2184136. Acesso em: 15 out. 2025.

UNIÃO EUROPEIA. *Regulamento (UE) 2024/1083 do Parlamento Europeu e do Conselho, de 11 de abril de 2024 (Regulamento Europeu relativo à Liberdade dos Meios de Comunicação Social)*. *JOUE*, L 1083, 17 abr. 2024. Disponível em: http://data.europa.eu/eli/reg/2024/1083/oj. Acesso em: 15 out. 2025.

UNIÃO EUROPEIA. *Regulamento (UE) 2022/2065 do Parlamento Europeu e do Conselho, de 19 de outubro de 2022 (Digital Services Act – DSA)*. *JOUE*, L 277, 27 out. 2022, p. 1-102. Disponível em: http://data.europa.eu/eli/reg/2022/2065/oj. Acesso em: 15 out. 2025.

UNIÃO EUROPEIA. *Carta dos Direitos Fundamentais da União Europeia*. *JOUE*, C 202, 7 jun. 2016, p. 389-405. Disponível em: http://data.europa.eu/eli/treaty/char_2016/oj. Acesso em: 15 out. 2025.

VAN CLEYNENBREUGEL, Pieter; MATTIOLI, Pietro. "Digital Services Coordinators and Other Competent Authorities in the Digital Services Act: Streamlined Enforcement Coordination Lost?". *European Law Blog*, 2023. Disponível em: http://dx.doi.org/10.21428/9885764c.f18f26b8. Acesso em: 15 out. 2025.

VAN DE KERKHOF, Jacob. "The DSA's Tower of Babel: On Digital Services Coordinators and Freedom of Expression". *European Journal of Risk Regulation*, p. 1-26, 2025. Disponível em: https://doi.org/10.1017/err.2025.10034. Acesso em: 15 out. 2025.

WANG, Tiana. "Delicate Task: Content Moderation and Intermediary Liability in a Post-DSA World". *Berkeley Technology Law Journal*, v. 39, n. 4, p. 1507--1550, 2024. Disponível em: https://doi.org/10.15779/ Z38GH9BB6P. Acesso em: 15 out. 2025.

WEI, Johnny Tian-Zheng; ZUFALL, Frederike; JIA, Robin. "Operationalizing Content Moderation 'Accuracy' in the Digital Services Act". *AI, Ethics, and Society (AIES) Journal*, v. 7, p. 1527-1538, 16 out. 2024. Disponível em: https://doi.org/10.1609/aies.v7i1.31744. Acesso em: 15 out. 2025.

WILMAN, Folkert. "Het voorstel voor de Digital Services Act". *Nederlands Tijdschrift voor Europees Recht*, v. 27, n. 1-2, p. 27, 2021. Disponível em: http://dx.doi.org/10.5553/nter/138241202021027102002. Acesso em: 15 out. 2025.

WILMAN, Folkert; KALĖDA, Saulius Lukas; LOEWENTHAL, Paul-John. *The EU Digital Services Act*. Oxford: Oxford University Press, 2024. Disponível em: http://dx.doi.org/10.1093/law/9780198892847.001.0001. Acesso em: 15 out. 2025.

WILMAN, Folkert; KALĖDA, Saulius Lukas; LOEWENTHAL, Paul-John. "Implementation, Cooperation, Penalties and Enforcement". In: *The EU Digital Services Act*. Oxford: Oxford University Press, 2024. Disponível em: http://dx.doi.org/10.1093/law/9780198892847.003.0005. Acesso em: 15 out. 2025.

WILMAN, Folkert; KALĖDA, Saulius Lukas; LOEWENTHAL, Paul-John. "Introduction: Origins and Objectives of the DSA". In: *The EU Digital Services Act*. Oxford: Oxford University Press, 2024. Disponível em: http://dx.doi.org/10.1093/law/9780198892847.003.0001. Acesso em: 15 out. 2025.

WILMAN, Folkert; Kalėda, Saulius Lukas; Loewenthal, Paul-John. "Liability of Providers of Intermediary Services". In: *The EU Digital Services Act*. Oxford: Oxford University Press, 2024. Disponível em: http://dx.doi.org/10.1093/law/9780198892847.003.0003. Acesso em: 15 out. 2025.

Estados privados e soberania digital: a regulação de plataformas globais no Brasil e na América Latina

RODRIGO ARDISSOM DE SOUZA

Introdução[1, 2]

As empresas transnacionais de tecnologia, tais como Google, Meta e Amazon, não apenas acumulam um expressivo poder econômico, mas também assumem o papel de atores autônomos no ordenamento jurídico global. Ao regular fluxos de dados, controlar infraestruturas críticas e estabelecer normas que regem o ambiente digital, tais corporações exercem funções tradicionalmente associadas à soberania estatal. Esse fenômeno as distancia da estrutura que se consolidou no século XX, diminuindo o poder dos Estados-nação e consolidando "empresas paralelas", na medida em que suas práticas desafiam a efetividade dos marcos regulatórios nacionais e internacionais, bem como as dinâmicas clássicas da geopolítica. O presente texto aborda a crescente complexidade do ecossistema digital, caracterizado pela centralidade das grandes plataformas tecnológicas que ultrapassam os limites tradicionais do Estado, inaugurando novos desafios para a governança global no século XXI.

Essa dinâmica reforça a necessidade de políticas públicas consistentes que promovam a inclusão digital e fortaleçam infraestruturas locais. Neste contexto, a soberania digital emerge como uma questão

[1] Este trabalho foi realizado no âmbito do meu estágio de pós-doutorado no Programa Universitário de Estudos sobre Democracia, Justiça e Sociedade (PUEDJS) da Universidade Nacional Autônoma do México (UNAM). Agradeço ao Programa de Bolsas de Pós-Doutorado da UNAM, ao meu orientador John Mill Ackerman Rose e à Coordenação de Humanidades da UNAM pelo apoio concedido a esta pesquisa.

[2] Este texto contou com o apoio de ferramentas de revisão baseadas em modelos de linguagem de larga escala (LLMs), incluindo ChatGPT, Perplexity, Gemini e DeepSeek, utilizadas exclusivamente para revisão redacional e verificação preliminar de informações, sob supervisão e responsabilidade intelectual integral do autor.

estratégica, que transcende o controle técnico-jurídico de infraestruturas tecnológicas. Quando concebida não apenas como problema de governança, mas como um projeto político e ético, a soberania digital permite o desenvolvimento de políticas públicas voltadas à redução da dependência tecnológica externa, ao fortalecimento da infraestrutura nacional de dados e à promoção de um ecossistema digital mais inclusivo. Esta perspectiva demanda novos arranjos regulatórios e institucionais que assegurem, de maneira simultânea, a proteção dos direitos fundamentais, a promoção da inovação local e a autonomia tecnológica dos Estados, atribuindo à soberania digital um papel central no desenvolvimento nacional e na reconfiguração da ordem global.

A infraestrutura digital nacional revela vulnerabilidades críticas que se manifestam de forma particularmente aguda no tratamento de dados sensíveis por instituições públicas e acadêmicas, especialmente diante da ausência de políticas consistentes para a gestão federada de dados em conformidade com a Lei Geral de Proteção de Dados. Essa fragilidade está diretamente conectada à concentração de poder tecnológico nas mãos de poucas corporações globais, que controlam não apenas o processamento de dados, mas também os padrões técnicos e as capacidades operacionais necessárias para geri-los de forma autônoma. A falta de políticas alinhadas à LGPD para a gestão federada de dados ameaça tanto a privacidade quanto a segurança institucional, agravando dependências externas.

A concentração de capacidades técnicas em mãos de poucas empresas privadas acentua essa assimetria, especialmente no que diz respeito às tecnologias fundacionais da inteligência artificial, que reforçam lógicas de opacidade, deslocalização decisória e captura normativa. Ao mesmo tempo, estudos recentes indicam que o país sofre com a descontinuidade de políticas públicas voltadas à retenção e formação de quadros técnicos de excelência, evidenciada pela redução de 16,2% nas bolsas de pós-graduação da CAPES entre 2019 e 2022, com a taxa de cobertura caindo de 42% em 2014 para 32% em 2021 – ponto mais baixo –, recuperando parcialmente para 37% em 2023 (CAPES, 2020; CAPES, 2024), e pela evasão de pesquisadores qualificados para o exterior, o que agrava a dependência externa e compromete a autonomia tecnológica de longo prazo (MCTI, 2023; CNPq, 2024; SBPC, 2024). A articulação entre formação interdisciplinar em áreas críticas – como inteligência

artificial, robótica e computação de alto desempenho – e investimentos públicos estruturais é condição necessária, ainda que não suficiente, para reverter esse quadro.

Os fluxos globais de dados, cujo volume já havia superado o comércio internacional de bens físicos e continuou a crescer exponencialmente, configuram hoje um campo de disputa estratégica (McKinsey Global Institute, 2016, pp. i-iii; OECD 2022, p. 12). O McKinsey Global Institute já destacava que esses fluxos internacionais cresceram mais de 45 vezes entre 2005 e 2020, inaugurando uma nova era de globalização digital orientada por dados. Essa tendência consolidou-se e se intensificou nos últimos anos, com estimativas recentes indicando que os dados transfronteiriços não apenas continuam a impulsionar cadeias globais de valor, inovação e produção, mas já contribuem de maneira significativa para o PIB mundial (OECD, 2023, p. 15; LSEG, 2024). Tais fluxos constituem tanto o substrato técnico da economia digital quanto o motor de uma nova governança transnacional, marcada pela centralidade de atores privados e pela erosão de mecanismos clássicos de responsabilização. O poder acumulado pelas *big techs* não se limita ao controle mercantil: avança também sobre as esferas normativas e simbólicas, influenciando práticas democráticas e decisões de âmbito coletivo.

Este artigo se inscreve em uma investigação jurídica e político-conceitual que examina as transformações contemporâneas da soberania estatal frente ao avanço de regimes normativos privados operados por grandes plataformas tecnológicas. O enfoque metodológico adotado é de caráter crítico e analítico, baseado na revisão de literatura especializada, documentos normativos, resoluções judiciais paradigmáticas e materiais públicos relevantes vinculados à governança digital e às formas emergentes de autorregulação corporativa. Adotamos um desenho qualitativo com três eixos: (*i*) exame de decisões e votos nos Recursos Extraordinários 1.037.396 e 1.057.258 do Supremo Tribunal Federal (STF) e seus efeitos sobre o art. 19 do Marco Civil da Internet (MCI); (*ii*) análise de conflitos entre o Estado brasileiro e plataformas digitais (bloqueios e medidas coercitivas) como casos críticos; e (*iii*) avaliação dos limites da autorregulação corporativa, especialmente do Oversight Board da Meta, à luz de literatura especializada (Douek, 2022, pp. 526-527; Gorwa & Garton Ash, 2020, pp. 286-287; Al Sur, 2022).

A partir de uma perspectiva comparada e situada no contexto latino-americano, o estudo analisa o deslocamento de funções soberanas para atores privados mediante o exame de casos emblemáticos. O recorte temporal concentra-se no período de 2019 a 2025, coincidindo com a intensificação de conflitos regulatórios e o amadurecimento de debates sobre soberania digital na região. Mais do que reconstruir cronologias, o objetivo é mapear as lógicas estruturantes que possibilitam a emergência de "Estados privados" no espaço digital e discutir os limites jurídicos da autorregulação tecnológica ante a erosão dos marcos públicos de controle. Não abordamos neste trabalho regulação setorial de publicidade, mensuração econométrica de impactos ou análise antitruste detalhada, salvo quando essas dimensões interferem diretamente na arquitetura de moderação de conteúdo e na circulação informacional.

Nesse cenário, disputas como a travada entre o Supremo Tribunal Federal (STF) e a plataforma X (antigo Twitter) evidenciam os limites do aparato estatal diante da resistência organizada das corporações globais. O descumprimento reiterado de decisões judiciais por parte da empresa, seguido do bloqueio temporário da plataforma no Brasil em agosto-setembro de 2024, reatualizou o debate sobre os limites constitucionais da liberdade de expressão e o papel do Estado na mediação de discursos digitais em situações de ameaça à ordem democrática. Mais do que um caso isolado, o episódio insere-se numa constelação de conflitos em que plataformas globais operam à revelia das jurisdições nacionais, tensionando a legitimidade da autoridade pública e deslocando o eixo da normatividade.

O julgamento da constitucionalidade do art. 19 do Marco Civil da Internet representa outro ponto de inflexão. Em 26 de junho de 2025, o Supremo Tribunal Federal, por maioria de oito votos a três, fixou tese de repercussão geral que amplia os deveres de cuidado das plataformas e admite responsabilização a partir de notificação extrajudicial em casos de ilícitos gravíssimos (STF, 2025a, Teses de Repercussão Geral RE 1037396 e RE 1057258). O acórdão foi publicado no Diário da Justiça Eletrônico em 5 de novembro de 2025, com 1.323 páginas, 132 dias após a decisão – prazo superior à média de 56,7 dias para casos de repercussão geral, refletindo a complexidade da matéria. A tese fixada tem efeito vinculante e aplicação prospectiva. Assim, descrevemos a tese majoritária, distinguindo-a dos votos individuais e assinalando pontos

pendentes de consolidação. A eventual consolidação desta interpretação pode abrir espaço para um novo paradigma de regulação, mais proativo, mas também mais suscetível a riscos de moderação excessiva. O desafio não está apenas em revisar um regime normativo; está em reconfigurar o lugar do Estado em um ecossistema normativo híbrido, onde as fronteiras entre regulação pública e governança privada se mostram cada vez mais porosas.

Por fim, a consolidação de instâncias normativas autônomas, como o Oversight Board da Meta, ilustra a emergência de formas sofisticadas de autorregulação corporativa. O desenho institucional do Oversight Board limita-o a revisões *ex post* de casos individuais, circunscritas ao ecossistema Meta, sem poderes de auditoria algorítmica nem efeitos vinculantes *erga omnes* (Douek, 2022, p. 535). Tais arranjos não substituem a regulação pública, tampouco garantem *accountability* efetiva. Ao contrário, reforçam a fragmentação da normatividade e esvaziam a esfera pública de mecanismos democráticos de controle. Neste contexto, torna-se urgente pensar a soberania digital não como mera extensão da lógica territorial, mas como fundamento ético-jurídico para a reconstrução de instrumentos coletivos de regulação e justiça no século xxi. Diante desse cenário, o artigo investiga em profundidade como a atuação paraestatal de plataformas digitais desafia princípios constitucionais e examina criticamente respostas regulatórias, argumentando em favor de uma reapropriação estatal da soberania no meio digital mediante critérios verificáveis: dever de cuidado estruturado, transparência auditável, *enforceability* local e separação entre moderação de conteúdo e arquitetura.

O artigo está estruturado em quatro seções principais. A primeira analisa as tensões entre o Estado moderno e a construção de sistemas jurídicos privados, estabelecendo o conceito de "Estados privados". A segunda examina a jurisprudência do STF sobre o art. 19 do MCI. A terceira investiga conflitos jurisdicionais entre Estado e plataformas. A quarta avalia o Oversight Board como paradigma dos limites da autorregulação. A conclusão sintetiza argumentos e propõe critérios para a reapropriação estatal da soberania digital.

Tensões do Estado moderno e a construção privada de sistemas jurídicos

As plataformas digitais, em sua evolução, passaram a exercer um poder regulatório que desafia frontalmente o papel tradicional do Estado moderno. De fato, elas operam como sistemas jurídicos privados, com competência para estabelecer normas, julgar comportamentos e impor sanções em seus ecossistemas. Essa dinâmica ataca frontalmente a estrutura política estatal (Schmitt, 2005 [1922], p. 9), indo além da mera erosão da soberania clássica para uma reconfiguração estrutural do espaço jurídico e político global.

Nesse novo cenário, decisões críticas relacionadas a direitos fundamentais – como a liberdade de expressão, o acesso à informação e a proteção de dados pessoais – estão cada vez mais concentradas em instâncias privadas, alheias aos mecanismos institucionais democráticos e moldadas por lógicas corporativas. A transparência algorítmica, enquanto desdobramento dos direitos constitucionais à informação e à proteção de dados (conforme previsto no art. 5.º, incisos xiv e lxxii, e no art. 220 da Constituição Federal, bem como nos arts. 6.º, 9.º e 20 da Lei Geral de Proteção de Dados), emerge como princípio essencial para assegurar a fiscalização e o controle público sobre decisões automatizadas que afetam direitos individuais e coletivos. Tal deslocamento normativo consolida um poder transnacional que tensiona diretamente os princípios constitucionais que historicamente fundamentaram os Estados-nação (Lessig, 2006, pp. 5-6).

Entendemos por Estados privados as estruturas corporativas transnacionais que, ao controlar infraestruturas digitais críticas e regrar comportamentos *online*, passam a desempenhar papéis típicos do Estado, sem investidura pública ou *accountability* democrática. Este conceito alinha-se às formulações de Cohen (2012, p. 9), que analisa a emergência de "constitucionalismos digitais" configurados por atores privados, e Celeste (2022, p. 77), que examina o *digital constitutionalism* como governança híbrida. Também dialoga com Srnicek (2017, p. 45) sobre capitalismo de plataformas e DeNardis (2014, p. 15) sobre privatização da governança. A diferença central reside na operacionalização mediante critérios verificáveis: dever de cuidado estruturado,

transparência auditável, *enforceability* local e separação entre moderação de conteúdo e arquitetura.

No contexto brasileiro, a resistência sistemática de corporações como Google, Telegram e Meta em cumprir determinações judiciais revela que a questão é política e jurídica, não técnica. Casos concretos que serão analisados na Seção III incluem: (*i*) bloqueios do WhatsApp em 2015 e 2016 por recusa em fornecer dados; (*ii*) resistência do Telegram durante eleições de 2022; e (*iii*) descumprimento da plataforma X em 2023-2024, culminando em bloqueio nacional. O conflito é normativo: disputa por legitimidade entre autoridades públicas e corporações transnacionais. O descumprimento reiterado evidencia tanto a limitação das ferramentas coercitivas estatais quanto uma assimetria estrutural: o Estado atua sob regime público com controles democráticos, enquanto as plataformas impõem regras privadas baseadas em interesses econômicos, alheias a consequências sociais e políticas.

A noção de neutralidade das plataformas dissolve-se diante de análise rigorosa. Elas são agentes normativos ativos que exercem controle substancial sobre a esfera pública digital, definindo quais discursos são amplificados ou silenciados via algoritmos opacos (Gillespie, 2018, pp. 1-23). A curadoria algorítmica é governança privada que molda debate público e afeta instituições democráticas.

Na América Latina, caracterizada por arcabouços regulatórios fragmentados, essas plataformas operam com ampla margem para impor regimes normativos próprios (Da Silva & Núñez Reyes, 2021, pp. 45-50). A ausência de coordenação regional favorece a "soberania paralela", onde multinacionais decidem sobre direitos fundamentais sem controle estatal efetivo. Casos como a reconfiguração do Twitter por Elon Musk ou alterações nos algoritmos do Facebook (Hofheinz, 2023, pp. 45-50) evidenciam que decisões com impacto sistêmico podem ser tomadas por indivíduos ou conselhos privados, sem vínculo constitucional ou compromisso democrático (Srnicek, 2017).

O que está em curso é a transformação estrutural nas formas de mediação social. Plataformas não apenas intermedeiam: definem parâmetros normativos e regimes de visibilidade. Ao estabelecerem regras que se sobrepõem, ignoram ou contradizem legislações nacionais, constroem um regime jurídico extraterritorial que orienta comportamentos segundo lógicas de mercado (Celeste, 2022). A digitalização do espaço

público sob lógicas privadas compromete a legitimidade democrática e dificulta a responsabilização. A coexistência de normatividades estatais e privadas fragmenta a legitimidade democrática. Enquanto os Estados buscam exercer a autoridade soberana, as *big techs* operam com normatividades próprias, desvinculadas de território e resistentes à jurisdição estatal. Esse hibridismo, alimentado por lacunas, *forum shopping* e estruturas transnacionais, limita a eficácia dos instrumentos clássicos do direito internacional. Plataformas emergem como instâncias regulatórias paralelas sem ancoragem territorial ou *accountability*.

O Marco Civil da Internet (Lei 12.965/2014) revela potencial de marcos que buscam integrar corregulação, transparência e responsabilidade compartilhada (Brasil, 2014). Contudo, desafios atuais superam capacidades do arcabouço vigente, exigindo dimensões políticas e técnicas capazes de conter poder concentrado. A tese do excepcionalismo cibernético continua alimentando o debate internacional (Goldsmith & Wu, 2006). A tensão não é apenas conflito de interesses: evidencia urgência de reconfigurar o direito internacional para a proteção efetiva de direitos em ambiente informacional transformado.

A *internet* não se desenvolveu em vácuo institucional. Sua expansão foi viabilizada por incentivos econômicos e estruturas regulatórias públicas, especialmente investimentos estatais em infraestrutura e pesquisa (Mazzucato, 2011, pp. 28-35). Longe de espaço autônomo, o ambiente digital resulta de políticas públicas deliberadas. A ideia de "espaço sem fronteiras" é narrativa funcional que omite o papel decisivo do Estado na arquitetura informacional. O crescimento exponencial, porém, permitiu o surgimento de atores privados com capacidade de controlar infraestruturas críticas e regular interações em escala planetária.

Conceitos clássicos como soberania, jurisdição e territorialidade perdem força analítica diante de redes transnacionais. Interconexão e ubiquidade ultrapassam a capacidade dos Estados de exercerem controle efetivo, pois o desenvolvimento técnico supera a velocidade de adaptação legislativa. A territorialidade jurídica dissolve-se onde indivíduos e corporações elidem responsabilidades, tornando a construção da responsabilidade jurídica instável. Nesse contexto, o espaço digital apresenta-se como domínio autônomo onde limitações normativas tradicionais tornam-se inoperantes. Esse redesenho das formas de mediação social exige respostas institucionais articuladas. É nesse contexto que se

insere o debate sobre a constitucionalidade do art. 19 do Marco Civil da Internet, analisado a seguir.

A tensão tecnopolítica e a constitucionalidade do art. 19 do Marco Civil da Internet

O julgamento da constitucionalidade do art. 19 do Marco Civil da Internet pelo Supremo Tribunal Federal, concluído em 26 de junho de 2025, resultou em decisão de alto impacto que redefine bases da responsabilização das plataformas digitais no Brasil (STF, 2025a). Este processo transcende a esfera técnico-jurídica, envolvendo questões estruturais relativas à soberania digital e ao poder econômico-normativo das plataformas. Originalmente concebido para equilibrar liberdade de expressão e controle judicial, o art. 19 foi intensamente questionado pela proliferação industrial de desinformação sistêmica, discursos de ódio e crimes cibernéticos.

Esse clamor por regulação mais firme atravessa o espectro político-ideológico nacional, indicando que o modelo anterior, ancorado em ordens judiciais *ex post facto*, não satisfazia a percepção social de justiça e proteção. Paralelamente, a tentativa de avançar com o Projeto de Lei (PL) n.º 2.630/2020 (PL das Fake News) esbarrou em pressões políticas, *lobbies* corporativos e polarizações legislativas, resultando no arquivamento em abril de 2024 e aprofundando a paralisia do Poder Legislativo. Essa inação forçou o Judiciário a assumir protagonismo central na tarefa de definir limites para plataformas, atribuição que analistas consideram atípica e de alto risco institucional.

As *big techs* viam no art. 19 anteparo fundamental contra censura prévia, argumentando que o endurecimento regulatório poderia induzi-las à remoção preventiva arbitrária, empobrecendo o debate público. O STF estava diante de um dilema constitucional: como incorporar anseios por responsabilização mais efetiva, prevenir abusos sistêmicos e garantir que a arquitetura jurídica acompanhasse a evolução tecnológica, sem solapar a liberdade de expressão. Em decisão histórica, o STF, por maioria de oito votos a três, declarou o art. 19 parcialmente inconstitucional, estabelecendo nova tese de repercussão geral para a responsabilização de provedores por conteúdo de terceiros (STF, 2025a, Teses RG RE 1037396 e RE 1057258). O acórdão foi publicado no Diário

da Justiça Eletrônico em 5 de novembro de 2025. Descrevemos a tese majoritária, distinguindo-a dos votos individuais e assinalando pontos pendentes de consolidação.

A Corte firmou entendimento de que, para conteúdos ilícitos gravíssimos e manifestos – como incitação a atos antidemocráticos, terrorismo, induzimento ao suicídio, discriminação por raça/religião/identidade de gênero, crimes contra mulher e pornografia infantil –, as plataformas podem ser responsabilizadas civilmente pela não remoção a partir de notificação extrajudicial, desde que comprovada falha sistêmica no dever de cuidado (STF, 2025b; Migalhas, 2025). Para crimes contra a honra e ilícitos de menor gravidade, a exigência de ordem judicial permanece, embora as plataformas sejam incentivadas a remover proativamente quando a ilicitude for clara.

Os votos dos Ministros refletiram complexidade interpretativa significativa. O Ministro Dias Toffoli (relator) argumentou pela inconstitucionalidade parcial, defendendo suficiência da notificação extrajudicial para ilícitos flagrantes. Reconheceu que a *internet* não é "território sem lei" e que a judicialização prévia entrava a celeridade, estendendo a possibilidade a contas falsas, impulsionamentos pagos e conteúdos eleitorais (Toffoli, Voto RE 1037396, 2025; Migalhas, 2025). O Ministro Luiz Fux propôs a responsabilização objetiva e direta sem ordem judicial como regra para ilícitos penais manifestos, enfatizando dever de vigilância e moderação proativa, inclusive para conteúdo impulsionado ou monetizado (Fux, Voto RE 1057258, 2025; Poder360, 2025). O Ministro Luís Roberto Barroso defendeu um regime diferenciado conforme o tipo de ilícito, atribuindo "dever de cuidado" para riscos sistêmicos e a exigindo relatórios de transparência, alinhando-se ao Digital Services Act europeu.

Em contraponto, André Mendonça, Edson Fachin e Kassio Nunes Marques defenderam constitucionalidade plena, argumentando que a exigência de ordem judicial é fundamental para preservar liberdade de expressão e evitar "censura privada" sem controle estatal, alertando que a redefinição caberia ao Legislativo (Mendonça, Voto RE 1037396, 2025; Fachin e Nunes Marques, Votos RE 1057258, 2025; Poder360, 2025; O Globo, 2025). Flávio Dino, Cristiano Zanin e Gilmar Mendes endossaram parcial inconstitucionalidade, reforçando a necessidade de regime diferenciado para *marketplaces*, anúncios pagos e conteúdo manipulado por algoritmos, reconhecendo poder de amplificação

(Poder360, 2025). Alexandre de Moraes propôs equiparar redes sociais a meios de comunicação de massa, exigindo representação nacional, transparência algorítmica e responsabilização solidária por conteúdos ilícitos (Poder360, 2025). Cármen Lúcia considerou que as plataformas não são neutras e devem responder por descumprimento judicial, combinando com dever de cuidado proativo (Poder360, 2025).

A tese majoritária é consenso complexo que busca equilibrar ampliação da responsabilização sem inviabilizar liberdade de expressão. Exige que as plataformas implementem autorregulação obrigatória com sistemas de notificação eficazes, garantia de devido processo para usuários, relatórios de transparência detalhados e canais acessíveis (STF, 2025a). Incluiu obrigatoriedade de sede e representante legal no Brasil e apelo explícito ao Congresso para criar um marco legal permanente atualizado (O Globo, 2025; Migalhas, 2025).

Este veredito não encerra debate, mas o modula para novo patamar, deslocando ênfase para deveres proativos e necessidade de critérios claros para remoção. A redação original do art. 19, ao condicionar a responsabilização à prévia notificação e descumprimento judicial, mostrava-se insuficiente para a disseminação massiva de desinformação, discursos de ódio e práticas danosas (Koops, 2014, p. 250). A interpretação crítica, parcialmente endossada pela tese do STF, apontava "imunidade excessiva" que dificultava a proteção de valores constitucionais essenciais, criando vácuo de *accountability* em ambiente de alto risco (Koops & Galič, 2021, p. 485). O veredito foi recebido por alguns setores como correção histórica necessária. Para juristas como Francisco Britto Cruz, Laura Schertel (2025) e Ronaldo Lemos (2025), a tese de convergência, embora solução judicial provisória, é "ampla e detalhada o suficiente para impedir excessos e coibir que provedores removam conteúdos a bel-prazer". Essa perspectiva contraria alarmes de que a decisão sufocaria a liberdade de expressão de pequenos *blogs* ou Wikipédia. Defensores argumentam que três tipos de provedores continuam sob o art. 19 original: e-mail, videoconferências fechadas e mensageria instantânea (Britto Cruz & Schertel, 2025). Essa modulação reflete uma solução ponderada, ajustando o dispositivo criado para "internet que não existe mais" e promovendo democracia e equidade econômica.

O julgamento transcende a esfera técnico-jurídica, envolvendo questões estruturais de soberania digital e poder econômico-normativo

das plataformas no Brasil (DeNardis, 2014). Em contexto de assimetrias globais no controle de infraestrutura e dados, a decisão do STF abre caminho para redefinir o papel do Estado na governança digital. Ao estabelecer parâmetros que potencialmente priorizam direitos fundamentais ante domínio corporativo, sinaliza uma mudança de paradigma. A controvérsia não se limita ao plano doméstico: insere-se ativamente no debate internacional sobre responsabilidade de intermediários, atuando como marco global na reconfiguração do capitalismo de plataformas.

Embora o Brasil seja reconhecido por legislação pioneira, o país continua enfrentando o desafio de adaptar-se às dinâmicas contemporâneas do ecossistema digital, notadamente escalabilidade, automação e virulência da propagação de ilícitos. A experiência demonstrava que a lentidão de processos judiciais e a exigência de decisão individualizada não ofereciam respostas à velocidade de conteúdos danosos – desinformação, discurso de ódio, incitamento à violência – que se propagam viralmente, ameaçando democracia e segurança. Episódios como ataques de 8 de janeiro de 2023 evidenciaram falha estrutural no modelo (CPMI dos Atos de 8 de Janeiro, 2023).

A discussão sobre a reforma do art. 21 aprofunda esse processo ao avançar sobre a responsabilização civil em contextos sensíveis como divulgação não autorizada de imagens íntimas. Embora represente uma resposta necessária a violações graves, a forma como alterações têm sido propostas – muitas vezes por medidas provisórias ou projetos de tramitação acelerada – revela déficit deliberativo preocupante. A ausência de avaliações de impacto normativo, a escuta multissetorial qualificada e o diálogo com padrões internacionais de direitos humanos compromete a qualidade e legitimidade do processo legislativo.

A interpretação conforme a Constituição do art. 19 – impondo dever de cuidado preventivo e exigindo medidas proativas diante de conteúdos que violem gravemente direitos fundamentais – rompe com o modelo originalmente estabelecido pelo legislador em 2014, centrado na neutralidade, na não responsabilização prévia e na judicialização das ordens de retirada. Essa mudança não resulta de um processo legislativo democrático tecnicamente fundamentado, mas da construção jurisprudencial reativa diante de lacunas estruturais, amplificadas por crises sucessivas no ecossistema informacional.

É preciso reconhecer que o modelo original revelava limitações profundas diante de escala, velocidade e complexidade das interações digitais. Dados empíricos do InternetLab e da Coalizão Direitos na Rede demonstram que a judicialização como única via gerava assimetrias severas de acesso à justiça, ineficiência sistêmica e incapacidade de resposta célere a conteúdos evidentemente ilícitos que permaneciam disponíveis por longos períodos. Segundo levantamentos dessas organizações, usuários em situação de vulnerabilidade enfrentavam barreiras quase intransponíveis para obter decisões judiciais, enquanto conteúdos difamatórios, invasões de privacidade e violência digital prosseguiam sem resposta adequada. A regra da ordem judicial prévia, embora pautada por proteção à liberdade de expressão, criava blindagem desproporcional às plataformas, gerando ambiente permissivo à circulação de danos – inclusive violência política, discurso de ódio e exposição não consentida de imagens íntimas.

Contudo, a guinada hermenêutica promovida pelo STF não se limita a corrigir disfunções operacionais. Promove mutação na arquitetura do Marco Civil sem o devido respaldo legislativo, o que enseja preocupações sob perspectiva da legalidade estrita e reserva de lei. A exigência de "dever de cuidado" genérico, sem delimitação legal de escopo, critérios ou consequências jurídicas, transfere ao Judiciário – e às próprias plataformas – a tarefa de definir padrões de diligência e limites da moderação legítima. Tal transferência gera tensão latente entre necessidade de proteção de direitos e risco de privatização da censura, como alertado por Derechos Digitales (2024) e Nomura, Costa Filho e Ramos, 2025, que apontam opacidade de sistemas automatizados, assimetria de poder informacional e ausência de mecanismos efetivos de contestação e transparência.

Além disso, o julgamento insere-se em contexto institucional marcado por intensificação da atuação do STF em matérias de governança digital, evidenciado pelo Inquérito 4781 ("das Fake News"), instaurado em 2019 e ainda em trâmite. A manutenção de inquérito sob controle direto do Supremo, com base em fundamentos excepcionais como Lei de Segurança Nacional (revogada em 2021), tem gerado controvérsias sobre separação de poderes, devido processo legal e uso expansivo da jurisdição constitucional para tutela do espaço público digital. Embora legitimado por setores que veem na Corte bastião de defesa democrática, o protagonismo também levanta questões sobre *accountability*,

excesso de jurisdição e sobreposição de funções típicas dos Executivo, Legislativo e Judiciário.

A reforma hermenêutica do art. 19 é indissociável do redesenho do papel do Estado no ambiente digital. Trata-se de disputa normativa o profunda que não se reduz à adequação de dispositivos legais, mas tensiona próprio equilíbrio entre poderes, a definição de competências regulatórias e os modos de produção da normatividade em tempos de transnacionalização da infraestrutura comunicacional. As plataformas digitais operam hoje como centros normativos de facto, estabelecendo critérios de visibilidade, exclusão e relevância à margem dos controles institucionais clássicos. Ao redefinir o art. 19 por via judicial, o Brasil posiciona-se como laboratório de experimentação normativa – ainda que sem solidez institucional e legal que tal protagonismo exige.

A discussão sobre a reforma do art. 21 aprofunda esse processo ao avançar sobre a responsabilização civil em contextos sensíveis como a divulgação não autorizada de imagens íntimas. Embora represente a resposta necessária a violações graves, a forma como as alterações têm sido propostas – muitas vezes por medidas provisórias ou projetos de tramitação acelerada – revela déficit deliberativo preocupante. A ausência de avaliações de impacto normativo, a escuta multissetorial qualificada eo diálogo com padrões internacionais de direitos humanos comprometem a qualidade e legitimidade do processo legislativo.

Apesar da narrativa de avanço normativo, a ausência de critérios objetivos, auditáveis e respaldados em legislação primária pode comprometer o próprio objetivo de proteção de direitos, ao fomentar soluções casuísticas, judicializações massivas e decisões contraditórias. O risco de conformar sistema baseado em expectativas vagas de diligência ou modelos de *overcompliance* corporativo é real – podendo resultar em silenciamentos seletivos, discricionariedade algorítmica e erosão da previsibilidade jurídica. Diante disso, qualquer tentativa de revisar o Marco Civil – seja judicial ou legislativa – deve partir da premissa de que a regulação do espaço digital não pode ser reduzida nem à deferência tecnocrática às plataformas nem à lógica punitivista de responsabilização ilimitada. A construção de um novo marco regulatório requer um pacto normativo orientado por princípios constitucionais, estruturado em bases públicas e permeado por mecanismos de participação, controle e revisão. É nesse sentido que o momento atual, marcado por indefinição

quanto aos contornos da decisão do STF e ausência de regulamentação infralegal, deve ser compreendido não como ponto de chegada, mas como processo em disputa, cuja direção dependerá da capacidade institucional de produzir normas legitimadas, equilibradas e coerentes com fundamentos do Estado Democrático de Direito. Esse redesenho jurisprudencial, contudo, só produz efeitos duradouros quando alcança a arquitetura técnica e as cadeias contratuais que estruturam a circulação informacional. É precisamente nesse ponto onde jurisdição encontra infraestrutura que emergem conflitos entre Estado brasileiro e plataformas globais, examinados na seção seguinte.

3. Soberania em xeque no espaço digital: confrontos jurisdicionais e geopolítica da governança

A concepção clássica de soberania estatal, fundamentada na autoridade suprema sobre território e população, encontra-se em reconfiguração diante das dinâmicas intrínsecas ao espaço digital. Lógicas coloniais e práticas extrativistas que há séculos desafiam essa autoridade persistem, transmutadas na exploração massiva de dados e hegemonia das infraestruturas tecnológicas controladas por conglomerados transnacionais. O colonialismo digital reproduz estruturas de exploração e dependência, apropriando-se de vastos volumes de dados e concentrando poder econômico-informacional em poucas plataformas gigantescas. Longe de intermediárias neutras, essas entidades operam como atores hegemônicos capazes de moldar mercados, comportamentos sociais e estruturas normativas, desafiando diretamente a soberania jurisdicional dos Estados, especialmente os periféricos, e acentuando desigualdades estruturais no sistema internacional.

A materialização dessas tensões revela-se em confrontos recentes que evidenciam as dificuldades dos Estados nacionais em fazer valer normas no ambiente digital globalizado. Os bloqueios temporários do WhatsApp no Brasil (2015 e 2016), decorrentes do descumprimento de ordens judiciais em investigações criminais e da recusa da empresa em fornecer dados criptografados, foram marcos iniciais dessa nova etapa de confrontação. Em dezembro de 2015, juiz da 1.ª Vara Criminal de São Bernardo do Campo determinou suspensão do aplicativo por 48 horas após o Facebook (controladora do WhatsApp) não atender a

ordem judicial de fornecer dados em investigação criminal. O bloqueio afetou cerca de 100 milhões de usuários, gerando impacto econômico e social significativo. O episódio repetiu-se em maio de 2016, quando juiz de Sergipe ordenou o bloqueio por 72 horas novamente por descumprimento de ordem judicial em investigação de tráfico de drogas. Embora bloqueios tenham gerado discussões legítimas sobre o impacto desproporcional sobre milhões de usuários, deixaram claro que a não intervenção estatal plena é inviável. Esses precedentes sinalizaram a disposição do Judiciário brasileiro em assegurar cumprimento de normas domésticas mesmo diante de resistência de plataformas globais com políticas internas próprias.

A tensão entre proteção de direitos individuais e salvaguarda de direitos coletivos tornou-se ainda mais aguda diante do avanço de plataformas como Telegram, cuja lógica transnacional frequentemente desafia legislações domésticas. Durante o período eleitoral de 2022, a resistência do Telegram em remover canais que disseminavam desinformação massiva, discursos de ódio e conteúdos claramente ilegais levou o Tribunal Superior Eleitoral (TSE) a aplicar multas milionárias e ameaçar bloqueio da plataforma. O Telegram mantinha-se resistente a cooperar com as autoridades brasileiras, alegando que sua arquitetura descentralizada e criptografia dificultavam a moderação centralizada. O ministro Alexandre de Moraes, então presidente do TSE, determinou a multa diária de R$ 100 mil por descumprimento de ordens e chegou a ameaçar a suspensão do aplicativo. Apenas sob a iminência de bloqueio, o Telegram indicou representante legal no Brasil e começou a atender parcialmente requisições judiciais. Tal desafio ilustra a capacidade dos capitais e arquiteturas tecnológicas de impor normas de conduta próprias, em detrimento das autoridades estatais.

Nesse contexto de redefinição da autoridade estatal sobre fluxos digitais, o Inquérito das Fake News (INQ 4.781), instaurado em 2019 sob relatoria de Alexandre de Moraes, emergiu como o epicentro de disputas centrais para a afirmação da soberania digital no Brasil. A investigação abarcou desde a difusão de desinformação sistêmica até ataques coordenados contra instituições democráticas e a organização de atos antidemocráticos, evidenciando o uso político das plataformas. Casos como de Allan dos Santos e Oswaldo Eustáquio tornaram-se exemplares, ao disseminarem teorias conspiratórias e explorarem falhas na

moderação de conteúdo em redes como o X. Após as eleições de 2022, a desinformação escalou em sofisticação, incorporando estratégias como o uso de VPNs e aplicativos descentralizados, levando à ampliação do escopo da investigação em 2024.

A crise intensificou-se com a resistência reiterada da plataforma X e de seu proprietário, Elon Musk, ao cumprimento de ordens judiciais – incluindo recusa em remover conteúdos ilegais, ausência de representação legal no Brasil e ataques públicos ao STF. Musk utilizou a própria plataforma para criticar decisões judiciais brasileiras, acusando Moraes de "ditador" e alegando censura política. Em resposta, o Supremo aplicou multas diárias milionárias, determinou o bloqueio de ativos financeiros e, de maneira inédita, ordenou a suspensão do funcionamento da plataforma em agosto de 2024, estendendo sanções à Starlink com base na existência de "grupo econômico de fato" (Advocacia-Geral da União [AGU], 2024, p. 23). A decisão de bloquear a Starlink, embora polêmica, baseou-se no entendimento de que as empresas controladas pelo mesmo grupo econômico devem responder solidariamente por descumprimentos. A medida afetou usuários em áreas remotas da Amazônia e outras regiões dependentes de *internet* via satélite, gerando debates sobre proporcionalidade.

Apesar das controvérsias quanto à legalidade e proporcionalidade – especialmente considerando impactos sobre usuários em áreas dependentes da Starlink –, a decisão reafirmou com veemência a autoridade do Estado na regulação do ambiente digital. A Agência Nacional de Telecomunicações (Anatel) teve um papel central, coordenando ações com provedores de *internet* para bloquear o acesso a serviços como Cloudflare, utilizados para burlar sanções. Esse movimento, de alta complexidade técnica e jurídica, revelou a urgência de reforçar mecanismos de coordenação entre agências reguladoras e Poder Judiciário frente a plataformas que operam à revelia das determinações estatais. O bloqueio foi implementado mediante notificação a mais de 20 mil provedores de *internet*, que tiveram um prazo de 24 horas para cumprir a determinação sob pena de uma multa diária de R$ 50 mil.

O embate entre Elon Musk e o STF ultrapassou a esfera nacional, catalisando o debate global sobre os limites da liberdade de expressão *online* e a obrigação das plataformas de moderar conteúdos. Enquanto Musk defendia uma concepção irrestrita de liberdade, alinhada a

paradigmas norte-americanos e indiferente às especificidades constitucionais brasileiras, o STF reafirmou que esse direito fundamental não é absoluto: deve ser compatibilizado com a proteção da ordem democrática, os direitos de terceiros e a integridade institucional do Estado. O Marco Civil da Internet – especialmente à luz da reinterpretação do art. 19 na tese de a repercussão geral fixada pelo STF – foi fundamental na afirmação de modelo que busca o equilíbrio entre liberdade e responsabilização.

Ao enfrentar plataformas que desobedecem a decisões judiciais, o Estado brasileiro sinaliza a necessidade de preservar a estabilidade democrática frente a riscos de desinformação e manipulação algorítmica em larga escala. Nesse sentido, as declarações públicas do presidente Lula, inclusive em fóruns como o G20, reafirmaram a soberania jurídica nacional e o dever das empresas estrangeiras de submeterem-se ao ordenamento brasileiro. Sua posição representa um movimento estratégico que articula política doméstica e projeção internacional, visando uma governança digital global mais equitativa, fundada nos princípios do Estado de Direito, nos direitos humanos e na inovação responsável. Em discurso na Assembleia Geral da ONU em setembro de 2024, Lula defendeu que as "empresas de tecnologia não podem estar acima das leis nacionais" e propôs a criação de mecanismos multilaterais de governança das plataformas.

A constatação de que instrumentos tradicionais de regulação e *enforcement* são insuficientes diante das complexidades e da velocidade do ambiente digital impõe a necessidade de respostas institucionais articuladas. A construção de marcos regulatórios eficazes e justos, capazes de reconhecer a natureza transnacional das plataformas sem abrir mão da soberania jurídica local, exige articulação entre governos, sociedade civil, academia e setor privado. Nesse cenário, a cooperação internacional adquire centralidade para harmonizar princípios, construir parâmetros globais e assegurar que a tecnologia sirva ao bem público, à coesão social e à consolidação das democracias.

A soberania digital, compreendida não como fechamento, mas como capacidade de exercer controle efetivo sobre fluxos e infraestruturas digitais, deve ser pensada como componente essencial do território político e jurídico dos Estados. Essa soberania, mais do que conceito reativo, constitui alicerce de novo pacto regulatório para o século XXI,

no qual a atuação estatal se afirme frente à lógica desregulada dos mercados informacionais globais. A experiência brasileira, embora controversa e ainda em construção, oferece elementos empíricos para repensar a relação entre Estados nacionais e poder corporativo transnacional no espaço digital, apontando tanto para potencialidades quanto para limites e riscos de abordagens judicializadas de regulação.

Diante da ineficácia parcial do *enforcement* estatal e das assimetrias estruturais reveladas nesses conflitos, torna-se necessário avaliar criticamente as alternativas propostas pelas próprias plataformas. A seção seguinte examina o Oversight Board da Meta como caso paradigmático dos limites estruturais da autorregulação corporativa.

Os limites estruturais da autorregulação corporativa: os limites estruturais do Oversight Board

A progressiva delegação de funções regulatórias do Estado a entes privados no espaço digital tem sido legitimada por um discurso recorrente de eficiência, inovação e flexibilidade. Nesse cenário, as plataformas propuseram arranjos de autorregulação que prometem transparência, imparcialidade e respeito aos direitos fundamentais. Caso paradigmático é o Oversight Board da Meta, concebido como instância autônoma para revisar decisões sobre moderação de conteúdo. Sua criação, em 2019, representou uma tentativa institucional de responder a pressões globais por *accountability* e críticas acumuladas sobre violações de direitos, práticas opacas e interferências em processos democráticos.

Tal como Douek (2024) reconhece, o Conselho conquistou significativa "legitimidade sociológica" na mídia e academia, tornando-se rapidamente referência no debate sobre moderação de conteúdos *online*. Contudo, embora publicamente apresentado como independente, o Conselho opera estruturalmente vinculado aos marcos normativos internos da própria empresa, restringindo-se à avaliação *ex post* de casos individuais, sem ingerência direta sobre decisões estratégicas, *design* algorítmico, coleta de dados ou priorização de fluxos informacionais. Douek (2024) reforça esse ponto ao demonstrar que, apesar de avanços procedimentais pontuais, o Conselho evita persistentemente enfrentar questões mais difíceis e estruturais para as quais teria sido criado, priorizando "legitimidade de fachada" sobre solução efetiva de dilemas. Essa

limitação funcional, que Douek chama de "evasão estratégica", impede que o Oversight Board trate dos fatores estruturais que transformam as plataformas em vetores centrais de desinformação, violência simbólica e captura comunicacional das esferas públicas.

Mais do que falha conceitual ou operacional, o modelo de autorregulação corporativa revela utilidade estratégica para deslocar a autoridade normativa. Douek chama a atenção para o caráter "performativo" dessa governança privada: ao simular mecanismos de controle e transparência, tais estruturas neutralizam preventivamente demandas por regulação estatal efetiva, atuando como "biombo" para o poder corporativo. O Oversight Board encarna esse deslocamento – tribunal simbólico que atua em nome de comunidade global abstrata, mas permanece subordinado às racionalidades corporativas da *big tech* estadunidense. Gillespie (2018) e Klonick (2018) já alertavam que a instituição seria menos órgão normativo legítimo do que sofisticada ferramenta de gestão reputacional.

Douek atualiza a crítica mostrando como a legitimidade sociológica adquirida pelo Conselho obscurece o fato de que as decisões não têm força vinculante fora da Meta, tampouco dialogam eficazmente com sistemas jurídicos nacionais, especialmente no Sul Global, onde prevalecem profundas desigualdades informacionais e técnicas associadas à colonialidade tecnopolítica (Kwet, 2019, pp. 1-24). Essa dinâmica corrobora a emergência de "Estados Privados" que, ao estabelecerem as próprias leis e tribunais internos, erodem a soberania jurisdicional e normativa dos Estados-nação, impondo uma "soberania paralela" sobre o espaço digital.

A ausência de poder de auditoria sobre algoritmos, a inexistência de obrigações legais vinculantes de transparência e a incapacidade de provocar mudanças estruturais nos mecanismos das plataformas revelam o limite político da autorregulação corporativa. Como apontado por Srnicek (2017, p. 45), o modelo econômico dessas plataformas é intrinsecamente incompatível com a lógica de justiça informacional, já que se sustenta no engajamento massivo, priorizando conteúdos polêmicos, emocionais e polarizadores. Nesse sentido, Douek (2024) complementa a perspectiva ao argumentar que o Conselho adota métricas formalistas e simplistas para medir a eficácia, desconsiderando o impacto substantivo e material sobre direitos dos usuários.

O discurso da "governança *multistakeholder*" frequentemente invocado para legitimar o Oversight Board torna-se véu que encobre uma profunda assimetria estrutural entre usuários, Estados e empresas, travestindo de pluralismo uma governança essencialmente privada, desigual e autorreferente (Musiani et al., 2016, p. 15). Como demonstrado por Douek (2024), a autorregulação não opera como modelo legítimo de freios e contrapesos, mas como dispositivo estratégico de preempção regulatória. As plataformas resistem ativamente a qualquer forma de escrutínio público, negando acesso a dados sensíveis, utilizando cláusulas contratuais opacas e rebaixando padrões éticos em contextos periféricos.

Gorwa e Garton Ash já indicavam que a governança da liberdade de expressão global não pode depender exclusivamente de corporações com *accountability* voluntária e lógicas mercadológicas. Nesse ponto, as limitações denunciadas por Douek (2024) convergem diretamente com críticas feitas neste texto: a tentativa de construir um tribunal privado como o Oversight Board apenas posterga a necessidade urgente de mecanismos democráticos efetivos que articulem normatividade estatal, participação social e *accountability* jurídica genuína. A inércia regulatória, seja por deferência a modelos privados ou por captura de agenda, solidifica a "soberania paralela" das *big techs*, desafiando a própria fundação do direito público no ambiente digital.

Nesse contexto, recente jurisprudência do Supremo Tribunal Federal brasileiro – especialmente ao reinterpretar o art. 19 do Marco Civil da Internet – surge como contramovimento institucional significativo e estudo de caso emblemático diante da insuficiência estrutural da autorregulação (STF, 2025a, 2025b). Ao reconhecer dever de cuidado proativo das plataformas, independentemente de ordem judicial específica, o STF rompe com a lógica restritiva e formalista incorporada pelo Oversight Board. Enquanto este último enfatiza avaliações *ex post* limitadas e superficiais, a nova abordagem brasileira enfatiza a responsabilização das plataformas por falhas sistêmicas na arquitetura de circulação informacional.

Esse deslocamento do foco da moderação de conteúdo para estruturas invisíveis de produção e circulação de informação constitui ruptura decisiva com a lógica que rege modelos como o Oversight Board, revalorizando o papel do Estado como instância legítima de produção

normativa. Trata-se da reconfiguração da soberania no espaço digital que recoloca o interesse público no centro da governança da informação, alinhando-se inclusive com o robusto Digital Services Act europeu (European Union, 2022), que institui obrigações de transparência, auditoria independente e gestão de riscos, desafiando frontalmente a lógica voluntarista da autorregulação corporativa.

Por fim, é crucial sublinhar que modelos corporativos de autorregulação, como o Oversight Board da Meta, não apenas negligenciam especificidades jurídicas e culturais locais, mas reforçam a lógica de dominação epistêmica. A universalização dos marcos normativos e valores do Norte Global como padrão de conduta digital constitui uma forma concreta de colonialismo digital. Ao reafirmar a soberania estatal sobre o espaço digital, propomos um ato de autodeterminação normativa e defesa dos direitos fundamentais, indispensável à construção de arquitetura informacional democrática, inclusiva e genuinamente plural.

Como adverte Douek, o perigo da abordagem atual não reside na ilegitimidade evidente dessas iniciativas, mas justamente na capacidade de se fazerem passar por modelos legítimos, influenciando futuras instituições de governança digital em direção a soluções formais e performativas, esvaziadas de transformações substantivas e estruturais. Assim, a falácia do Oversight Board não é apenas um problema da Meta, mas um alerta global sobre os limites do controle privado e a urgência da ação soberana. A experiência demonstra que delegar governança digital a tribunais corporativos equivale a consolidar assimetrias de poder, perpetuar opacidade e renunciar à possibilidade de construir um espaço público digital efetivamente democrático.

Conclusão

A presente investigação desvelou que a ascensão de Estados privados no espaço digital não apenas redefine fronteiras da soberania, mas configura um novo paradigma jurídico-político, onde a autoridade normativa é exercida por agentes não estatais com efetiva capacidade de modulação social. Trata-se de mutação estrutural: deslocamento da função regulatória clássica dos Estados para entes corporativos globais, que operam mediante controle de infraestruturas críticas, cadeias logísticas, fluxos informacionais e sistemas algorítmicos que organizam, mediam e

condicionam a experiência cotidiana dos cidadãos (Yeung, 2018, pp. 505–-523). Longe de testemunharmos simples erosão do Estado, observa-se a emergência de topologia normativa assimétrica, na qual corporações transnacionais passam a integrar, disputar e, em certos casos, suplantar funções constitutivas da ordem pública.

Essa reconfiguração manifesta-se em dispositivos jurídicos formais e informais, como termos de uso, *standards* técnicos e mecanismos de moderação de conteúdo, os quais, embora apresentem-se como neutros ou técnicos, impõem regimes normativos próprios, muitas vezes alheios ao controle democrático e à transparência exigível em regimes jurídicos republicanos (Cohen, 2012, pp. 1–23). O resultado é uma cidadania degradada, onde direitos fundamentais são subsumidos à lógica da governança por plataformas e à extração econômica de dados – modelo cuja racionalidade repousa na privatização do espaço público digital e na transformação da vida em recurso (Zuboff, 2019, p. 8; Srnicek, 2017). A cumplicidade institucional que sustentou essa consolidação – marcada por descompasso regulatório, interesses econômicos globalizados e captura de processos decisórios – foi fundamental para a emergência de soberanias paralelas.

É nesse contexto que a análise crítica do Oversight Board da Meta, retomada na introdução e aprofundada ao longo do texto, adquire centralidade como síntese empírica das limitações da autorregulação corporativa. Embora publicamente apresentado como mecanismo de supervisão independente, o Board opera sob marcos normativos delimitados pela própria empresa, sem autoridade real sobre algoritmos, fluxos de monetização ou estratégias de ranqueamento. Sua atuação ilustra uma lógica de governança performativa, em que a encenação da responsabilidade substitui seu exercício efetivo. Como apontado por Al Sur (2022; Celeste, 2022, p. 77), trata-se de "teatro regulatório" no qual a supervisão é instrumentalizada para legitimar regimes opacos de poder informacional. Sua ineficácia revela que a autorregulação, longe de ser solução adequada, tende a reforçar assimetrias e consolidar ausência de *accountability* pública no ambiente digital.

A análise demonstra que a insuficiência regulatória da autorregulação privada clama por ação estatal assertiva – reapropriação soberana da autoridade normativa, especialmente no equilíbrio entre liberdade de expressão, direitos fundamentais e integridade democrática. No Brasil,

essa reapropriação ganha forma na resposta jurisprudencial dada pelo STF ao reinterpretar o art. 19 do Marco Civil, configurando verdadeiro "dever de cuidado" (*duty of care*) das plataformas digitais. Ao deslocar o ônus da tutela de direitos da vítima para a própria arquitetura de funcionamento das plataformas, a Corte reconhece o caráter sistêmico dos danos informacionais e reafirma o papel do Estado como guardião da ordem constitucional no espaço digital.

Essa guinada jurisprudencial posiciona o Brasil na vanguarda global das respostas jurídicas à concentração de poder informacional. Ao exigir a implementação de medidas proativas – como canais acessíveis de denúncia, mecanismos transparentes de moderação e políticas eficazes de prevenção –, o STF alinha-se às tendências internacionais mais avançadas, como o Digital Services Act da União Europeia [Regulation (EU) 2022/2065], sem renunciar à autonomia interpretativa das normas constitucionais nacionais. Trata-se de gesto jurídico de afirmação soberana, que responde a conjuntura marcada por desinformação, violência política e tentativas de subversão democrática com doutrina sistematicamente construída e consequente.

Do ponto de vista geopolítico, delineia-se novo eixo de confrontos normativos: Estados nacionais tentam rearticular sistemas jurídicos em ambiente digital originalmente desenhado para escapar à jurisdição pública. A confrontação entre Estado brasileiro e plataformas como Telegram e X ilustra com precisão a recusa em aceitar um modelo de dominação informacional baseado no fórum privado. O esforço por instituir o direito público digital é uma forma de resistência à colonialidade algorítmica e à lógica extraterritorial que vem o regendo espaço digital (Kwet, 2019, pp. 13). A formação dos Estados privados digitais é sintoma de mutação mais ampla: descentralização e fragmentação da autoridade jurídica global. A governança digital deixou de ser campo técnico ou setorial para tornar-se arena principal das disputas por hegemonia, soberania e redistribuição de poder. Diante disso, reformas incrementais não bastam. O desafio exige refundação normativa que inclua marcos regulatórios efetivos, instâncias multilaterais de controle e incorporação ativa da sociedade civil e saberes locais na definição de direitos e garantias no digital. A luta pela soberania digital não é apenas jurídica: é epistêmica, cultural e civilizatória.

Referências

Advocacia-Geral da União. (2024). Manifestação na ADPF 1.188 (Petição 12.404): reconhecimento de "grupo econômico de fato" e bloqueio de bens [Secretaria-Geral de Contencioso]. https://www.gov.br/agu/pt-br/comunicacao/noticias/ADPF1188.pdf

Agência O Globo. (2025, 11 de junho). Maioria do STF vota para responsabilizar redes por conteúdo ilegal. Exame. https://exame.com/brasil/maioria-do-stf-vota-para-responsabilizar-redes-por-conteudo-ilegal/

Al Sur. (2022). *La moderación de contenidos desde una perspectiva interame ricana*. R3D/Al Sur. https://www.alsur.lat/en/report/content-moderation-interamerican-perspective

Alphabet Inc. (2024). Form 10-K for the fiscal year ended December 31, 2023. U.S. Securities and Exchange Commission. https://abc.xyz/investor/static/pdf/2023_alphabet_10K.pdf

Alphabet Inc. (2025, 4 de fevereiro). Alphabet announces fourth quarter and fiscal year 2024 results [Comunicado de imprensa]. https://abc.xyz/investor/news/news-details/2025/Alphabet-Announces-Fourth-Quarter-2024-and--Fiscal-Year-Results-02-04-2025/default.aspx

Barroso, L. R. (2025). Roberto Barroso — RE 1037396 (Tema 987) [Tese proposta]. Poder360. https://static.poder360.com.br/2025/06/tese-RobertoBarroso-MarcoCivildaInternet-STF.pdf

Brasil. (2014, 23 de abril). Lei n.º 12.965, de 23 de abril de 2014: Marco Civil da Internet. *Diário Oficial da União*.

Britto Cruz, F., & Schertel, L. (2025, 25 de junho). *STF e a responsabilidade das plataformas: Um novo equilíbrio*. JOTA. https://www.jota.info/opiniao-e-analise/artigos/stf-e-a-responsabilidade-das-plataformas-um--novo-equilibrio

Canalys. (2025, 12 de junho). Global cloud infrastructure spending rose 21% in Q1 2025. https://canalys.com/newsroom/global-cloud-q1-2025

CAPES. (2020). *Relatório de Gestão 2020. Coordenação de Aperfeiçoamento de Pessoal de Nível Superior*. https://www.gov.br/capes/pt-br/centrais-de-conteudo/05072021_RelatoriodeGestao2020.pdf

CAPES. (2024). GeoCAPES: Sistema de Informações Georreferenciadas da CAPES. Coordenação de Aperfeiçoamento de Pessoal de Nível Superior. https://geocapes.capes.gov.br

Caplan, R. (2023). Networked platform governance: The construction of the democratic platform. *International Journal of Communication*, 17, 3451--3472. https://ijoc.org/index.php/ijoc/article/view/20035

Castells, M. (2009). *Communication power*. Oxford University Press.

Celeste, E. (2022). *Digital constitutionalism: The role of Internet bills of rights*. *Routledg*e. https://doi.org/10.4324/9781003256908

Comitê Gestor da Internet no Brasil. (2023). *Sistematização das contribuições à consulta sobre regulação de plataformas digitais*. Núcleo de Informação e Coordenação do Ponto BR. https://cgi.br/publicacao/sistematizacao-das-contribuicoes-a-consulta-sobre-regulacao-de-plataformas-digitais/

Cohen, J. E. (2012). *Configuring the networked self: Law, code, and the play of everyday practice*. Yale University Press.

Couldry, N., & Mejías, U. A. (2019). *The costs of connection: How data is colonizing human life and appropriating it for capitalism*. Stanford University Press.

CPMI dos Atos de 8 de Janeiro. (2023). *Relatório final*. Congresso Nacional do Brasil.

Da Silva, F., & Núñez Reyes, G. (2021). *La era de las plataformas digitales y el desarrollo de los mercados de datos en un contexto de libre competencia* (LC/TS.2021/173). Comisión Económica para América Latina y el Caribe (CEPAL). https://repositorio.cepal.org/handle/11362/47540

De Gregorio, G. (2020). Digital constitutionalism: The role of the internet in shaping constitutional rights. *International Journal of Constitutional Law*, 18(1), 130–159. https://doi.org/10.1093/icon/moaa001

DeNardis, L. (2014). *The global war for Internet governance*. Yale University Press.

Derechos Digitales. (2024). *Reporte anual Derechos Digitales 2023: Gobernanza digital y justicia*. https://www.derechosdigitales.org/wp-content/uploads/MemoriaDD-2023.pdf

Dias Toffoli, J. A. (2025). Dias Toffoli — RE 1037396 (Tema 987): "Decálogo contra a violência digital e a desinformação" [Proposta de tese/voto]. Poder360. https://static.poder360.com.br/2025/0"/tese-DiasToffoli-Marco-CivildaInternet-STF.pdf

Douek, E. (2022). Content moderation as systems thinking. *Harvard Law Review*, 136(2), 526–607. https://harvardlawreview.org/print/vol-136/content-moderation-as-systems-thinking/

Douek, E. (2024). The Meta Oversight Board and the empty promise of legitimacy. *Harvard Journal of Law & Technology*, 37(2), 373–445.

Durand, C. (2020). *Technoféodalisme: Critique de l'économie numérique*. La Découverte.

eMarketer. (2024, 4 de abril). *US Amazon ecommerce forecast 2024*. https://www.emarketer.com/content/us-amazon-ecommerce-forecast-2024

eMarketer. (2024, 17 de abril). *Amazon will surpass 40% of US ecommerce sales this year, despite competition in grocery, home improvement*. https://www.emarketer.com/content/amazon-will-surpass-40-of-us-ecommerce-sales-this-year

European Union. (2022). Regulation (EU) 2022/2065 of the European Parliament and of the Council of 19 October 2022 on a Single Market for Digital Services and amending Directive 2000/31/EC (Digital Services Act).

Official Journal of the European Union, L 277, 1-102. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32022R2065

Fachin, E. (2025, 25 de junho). Proposta de tese – Edson Fachin – Marco Civil da Internet [Voto]. Poder360. https://static.poder360.com.br/2025/06/proposta-tese-EdsonFachin-MarcoCivilInternet-25-jun-2025-1.pdf

Fux, L. (2025). Luiz Fux — RE 1057258 (Tema 533) [Proposta de tese/voto]. Poder360. https://static.poder360.com.br/2025/06/tese-LuizFux-MarcoCivildaInternet-STF.pdf

Gillespie, T. (2010). The politics of "platforms". *New Media & Society*, 12(3), 347-364. https://doi.org/10.1177/1461444809342738

Gillespie, T. (2018). *Custodians of the internet: Platforms, content moderation, and the hidden decisions that shape social media*. Yale University Press.

Goldsmith, J. L., & Wu, T. (2006). *Who controls the internet?: Illusions of a borderless world*. Oxford University Press.

Gorwa, R. (2019). The platform governance triangle: Conceptualising the informal regulation of online content. *Internet Policy Review*, 8(2), 1-22. https://doi.org/10.14763/2019.2.1407

Gorwa, R., & Garton Ash, T. (2020). Democratic transparency in the platform society. In N. Persily & J. A. Tucker (Eds.), Social media and democracy: The state of the field (pp. 286-312). Cambridge University Press.

Kingsbury, B., Donaldson, M., & Vallejo, R. (2016). Global administrative law and deliberative democracy. In A. Orford & F. Hoffmann (Eds.), *The Oxford handbook of the theory of international law* (pp. 526–542). Oxford University Press. https://doi.org/10.1093/law/9780198701958.003.0027

Klonick, K. (2018). The new governors: The people, rules, and processes governing online speech. *Harvard Law Review*, 131(6), 1598-1670. https://harvardlawreview.org/print/vol-131/the-new-governors

Koops, B.-J. (2014). The trouble with European data protection law. *International Data Privacy Law*, 4(4), 250–261. https://doi.org/10.1093/idpl/ipu023

Koops, B.-J., & Galič, M. (2021). Unity in privacy diversity: A kaleidoscopic view of privacy definitions. *South Carolina Law Review*, 73(2), 465–500. https://scholarcommons.sc.edu/cgi/viewcontent.cgi?article=4408&context=sclr

Kwet, M. (2019). Digital colonialism: US empire and the new scramble for Africa's data. *Race & Class*, 60(4), 82-96. https://doi.org/10.1177/03063968188 23172

Lemos, R. (2025, junho). No Marco Civil, STF mirou nas big techs e acertou na internet inteira. *Folha de S. Paulo*. Republicado pelo ITS Rio em 1.º de julho de 2025. https://itsrio.org/pt/artigos/no-marco-civil-stf-mirou-nas-big-techs-e-acertou-na-internet-inteira/

Lessig, L. (2006). *Code: Version 2.0*. Basic Books. https://codev2.cc/

Mazzucato, M. (2011). *The entrepreneurial state: Debunking public vs. private sector myths*. Anthem Press.

McKinsey Global Institute. (2016). *Digital globalization: The new era of global flows*. McKinsey & Company.

McKinsey Global Institute. (2022). Global flows: The ties that bind in an interconnected world. McKinsey & Company. https://www.mckinsey.com/capabilities/strategy-and-corporate-finance/our-insights/global-flows-the-ties-that-bind-in-an-interconnected-world

MCTI. (2023). *Indicadores nacionais de ciência, tecnologia e inovação 2023*. *Ministério da Ciência*, Tecnologia e Inovação. https://www.gov.br/mcti/pt-br/acompanhe-o-mcti/indicadores/paginas/publicacoes/arquivos/indicadores_cti_2022.pdf

Mendonça, A. (2025). André Mendonça — RE 1037396 (Tema 987) [Proposta de tese/voto]. *Poder360*. https://static.poder360.com.br/2025/06/tese-AndreMendonca-MarcoCivildaInternet-STF.pdf

Meta Platforms, Inc. (2020, 19 de novembro). *Community Standards Enforcement Report: Third quarter 2020*. https://about.fb.com/news/2020/11/community-standards-enforcement-report-nov-2020/

Migalhas. (2025, 6 de novembro). *STF publica acórdão que derruba dispositivo do Marco Civil da Internet*. https://www.migalhas.com.br/quentes/443948/stf-publica-acordao-que-derruba-dispositivo-do-marco-civil-da-internet

Musiani, F., Cogburn, D. L., DeNardis, L., & Levinson, N. S. (Eds.). (2016). *The turn to infrastructure in Internet governance*. Palgrave Macmillan. https://doi.org/10.1057/9781137483591

Nomura, D. N. S., Costa Filho, J. R., & Ramos, P. H. (2025). *O Preço da Moderação: Impactos no Judiciário e o debate sobre a revisão do Marco Civil da Internet pelo STF* (Policy Briefs Reglab, n. 2). Reglab. https://reglab.com.br/wp-content/uploads/2025/06/rl-preco-da-moderacao.pdf

Nunes Marques, K. (2025, 26 de junho). Voto do ministro Nunes Marques [RE 1057258/MG]. *Poder360*. https://static.poder360.com.br/2025/06/voto-NunesMarques-responsabilizacaodasredes-STF-26-jun-2025.pdf

OECD. (2015). *Data-driven innovation: Big data for growth and well-being*. OECD Publishing. https://doi.org/10.1787/9789264229358-en

OECD. (2022). Going digital to advance data governance for growth and well-being. *OECD Publishing*. https://doi.org/10.1787/e3d783b0-en

Poder360. (2025, 27 de junho). *Leia as íntegras dos votos dos 11 ministros na responsabilização das redes*. https://www.poder360.com.br/poder-justica/leia-as-integras-dos-votos-dos-11-ministros-na-responsabilizacao-das-redes/

Pollicino, O. (2020). Constitutional adjudication and digital transformation: Towards a new constitutional paradigm? *German Law Journal*, 21(4), 564-582. https://doi.org/10.1017/glj.2020.32

RIBEIRO, D. B., & OLIVEIRA, E. F. DOS A. (2024). A distribuição de bolsas da CAPES em tempos de cortes orçamentários. *Temporalis*, 24(47), 35-50. https://doi.org/10.22422/temporalis.2024v24n47p35-50

SBPC. (2024, 22 de abril). Crise na pós-graduação: Evasão de pesquisadores prejudica ciência nacional. Sociedade Brasileira para o Progresso da Ciência. https://portal.sbpcnet.org.br/noticias/crise-na-pos-graduacao-evasao-de-pesquisadores-prejudica-ciencia-nacional/

SCHMITT, C. (2005). *Teologia política: Quatro capítulos sobre a doutrina da soberania* (1922). Del Rey.

SRNICEK, N. (2017). *Capitalismo de plataformas*. Boitempo.

STF. (2025a). Teses de repercussão geral nos RE 1037396 e RE 1057258. Supremo Tribunal Federal.

STF. (2025b). Acórdão nos RE 1037396 e RE 1057258, publicado no Diário da Justiça Eletrônico em 5 de novembro de 2025. Supremo Tribunal Federal.

SYNERGY RESEARCH GROUP. (2025, 1 de maio). *AI helps cloud market growth rate jump back to almost 25% in Q1*. https://www.srgresearch.com/articles/ai-helps-cloud-market-growth-rate-jump-back-to-almost-25-in-q1

TROTTIER, D. (2016). *Social media as surveillance: Rethinking visibility in a converging world*. Routledge.

VAN DIJCK, J., NIEBORG, D., & POELL, T. (2019). Reframing platform power. *Internet Policy Review*, 8(2). https://doi.org/10.14763/2019.2.1414

YEUNG, K. (2018). Algorithmic regulation: A critical interrogation. *Regulation & Governance*, 12(4), 505-523. https://doi.org/10.1111/rego.12158

ZUBOFF, S. (2019). *The age of surveillance capitalism: The fight for a human future at the new frontier of power*. PublicAffairs.

Digital Identity Management: Emerging Technologies and the Future of Security and Privacy

SÍLVIA DE CARVALHO HOMEM[1]

Abstract: In contemporary society, digital identity has become an indispensable component in the interactions between individuals, institutions and digital systems. The advent of technologies such as blockchain, biometric authentication, artificial intelligence, wearable technology and, more recently, the Internet of Bodies (IoB) has had a profound effect on the evolution of personal identification models. The present article puts forward a critical analysis of digital identity management, exploring the legal, ethical and political implications of the expansion of biometrics, the use of genetic data and the growing integration of connected body devices. It has been posited that these transformations are instrumental in the establishment of a novel paradigm of algorithmic surveillance, exerting a direct influence on the realms of informational self-determination and human dignity. The research adopts a legal-normative and multidisciplinary approach, proposing robust safeguards compatible with fundamental rights and the principles of the rule of law.

Keywords: Biometrics, Blockchain, Digital identity, Facial recognition, Fundamental rights, Genetic data, Informational self-determination, Internet of bodies, Wearables

Contents: I. Preliminary Considerations; II. Evolution of Digital Identity Systems III. Blockchain and Self-Sovereign Identity IV. Biometric Authentication: Potential and Risks V. Genetic Data and the Expansion of Biometrics VI. Wearable Technology VII. Surveillance Architectures and the Algorithmic Construction of Digital Identity VIII. Internet of Bodies (IoB): Body, Data and Control in the Age of Cybersurveillance

---

[1] Investigadora do Jusgov (Universidade do Minho).

IX. Emotion Recognition and the Fragility of Automated Interpretation
X. Final Considerations. Bibliographical References.

I. Preliminary Considerations

The accelerated digital transformation has profoundly impacted the methods by which individuals are identified, recognised and authenticated across various domains of social life. Digital identity has become a pivotal component in interactions with the State, commercial entities, and technological platforms, progressively superseding conventional identification models reliant on physical documentation. This transition is driven by emerging technologies such as blockchain, biometrics, artificial intelligence, technological and smart clothing, which promise greater security, convenience and personalisation.

A new phase of this evolution has recently come to the fore: the Internet of Bodies (IoB). This concept refers to the use of devices that are interconnected to the human body, such as clothing, ingestibles, implantables or embedded devices. The purpose of these devices is to collect and transmit biometric, physiological and behavioural data.[2] The Internet of Bodies signifies not merely a technological continuity, but rather an ontological shift[3] in the conceptualisation of the body as an identity interface and a continuous data source. In this new scenario, the body becomes not only an object of identification but also a subject of constant surveillance, creating a field of tension between innovation, freedom and control.

The present article puts forward a critical analysis of digital identity management, with a particular focus on the legal and normative implications arising from the integration of body-embedded technologies such as the Internet of Bodies. The argument advanced is that there is a need for a more articulated legal framework capable of addressing

---

[2] A Celik, KN Salama and AM Eltawil, 'The Internet of Bodies: A Systematic Survey on Propagation Characterisation and Channel Modeling' (2021) IEEE Internet of Things Journal https://www.researchgate.net/pu blication/353270324_The_Internet_of_Bodies_A_Systematic_Survey_on_Propagation_Characterization_and_Channel_Modeling.

[3] Mireille Hildebrandt, *Smart Technologies and the End(s) of Law: Novel Entanglements of Law and Technology* (Edward Elgar 2015); Shoshana Zuboff, *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power* (PublicAffairs 2019).

the specific risks associated with bodily-integrated data systems. Such a framework should complement and strengthen existing instruments (namely the GDPR, the Medical Device Regulation (MDR) and the forthcoming eIDAS 2.0 Regulation) by incorporating enhanced safeguards related to proportionality, purpose limitation, transparency, data minimisation and algorithmic accountability.[4] These principles are essential to ensure that digital identity systems remain consistent with fundamental rights and the individual's informational autonomy in an increasingly datafied society.

Methodologically, the study adopts a legal-normative approach grounded in a systematic bibliographical review and a critical analysis of relevant normative, technical and doctrinal materials at the intersection of law, technology and data protection.

II. Evolution of Digital Identity Systems

Digital identity management has become an increasingly significant field of study in the context of the intensification of interactions mediated by digital technologies. Digital identity systems were initially established on the basis of rudimentary authentication mechanisms, focusing on the integration of identifiers (usernames) and credentials (passwords). The advent of technological evolution resulted in the implementation of models supported by centralised databases, which were typically under the jurisdiction of institutional, public or private entities. Nevertheless, this centralised architecture has demonstrated structural vulnerabilities to cyber-attacks, resulting in unauthorised access and the compromise of sensitive data.[5]

In order to mitigate these weaknesses, the Public Key Infrastructure (PKI) was conceived, based on cryptographic mechanisms for identity authentication and the establishment of secure communication

---

[4] Regulation (EU) 2016/679 (General Data Protection Regulation) [2016] OJ L119/1; Regulation (EU) 2017/745 (Medical Device Regulation); European Commission, *Proposal for a Regulation amending Regulation (EU) No 910/2014 as regards establishing a framework for a European Digital Identity* COM(2021) 281 final.

[5] Paul Ohm, 'Broken Promises of Privacy: Responding to the Surprising Failure of Anonymisation' (2010) 57 UCLA *Law Review*, P. 1701-2010.

channels.[6] Notwithstanding the enhancements implemented by this model, the emergence of social networks and the widespread adoption of federated authentication systems, such as Single Sign-On (SSO), have given rise to novel attack vectors and have substantially augmented the exposure of personal data, thereby escalating the risks associated with digital security and privacy.

III. *Blockchain* and Self-Sovereign Identity

The advent of blockchain technology has emerged as a robust solution for the management of decentralised identities, providing a more secure and tamper-resistant system designed to mitigate the vulnerabilities associated with centralised architectures.[7] In this context, users' personal data are stored on mobile devices (such as smartcards or smartphones with secure hardware) according to the principles of self-sovereign identity. The authenticity of data access is guaranteed by cryptographic mechanisms, ensuring the integrity, privacy and confidentiality of the information. Cryptographic proof of identity is stored in a decentralised manner, based on blockchain technology, thus eliminating the need for third parties to verify identity and reducing the risks associated with data sharing. The self-sovereign identity model is predicated on the notion that it should provide users with exclusive control over their personal data, thereby engendering autonomy with regard to their management and disclosure.

Self-sovereign identity is defined as an identity management system that functions independently of centralised entities. It is characterised by a decentralised architecture that prioritises security, privacy and user self-determination. In this model, the concept of the right to selective disclosure empowers individuals to exercise control over the disclosure of their personal information, in accordance with the principle of informational self-determination, as codified in the General Data Protection Regulation (GDPR).

---

[6] Andrea Renda, *The Legal Framework for eID and Trust Services in the EU* (CEPS Research Report 2018/01, 2018).

[7] Renda, *Legal Framework for eID*.

The primary benefits of blockchain technology can be summarised as follows: firstly, it guarantees the integrity of database records and, secondly, it enables the detection of any improper or unauthorised alteration of data.[8] In practical terms, this means that if a credential, transaction or attribute stored on the blockchain is modified (even slightly) the system automatically flags the discrepancy, as the altered record no longer matches the cryptographic hash validated by the network. For example, an attempt to change the attributes of a digital identity credential or to falsify a timestamped record would be immediately identifiable and rejected by the system. This provides evidence of interference with activity records. Nevertheless, challenges pertaining to interoperability between disparate platforms persist, as do issues relating to the scalability (the capacity to expand or adapt to changes in scale, scaling up or down, without compromising performance or efficiency) of this technology.

The European Union's proposed eIDAS 2.0 Regulation constitutes a practical example of the implementation of these principles, as it provides for the creation of a European digital identity wallet enabling secure access to public and private services. The system is designed to minimise the disclosure of personal information and to enhance user control over digital credentials. In practical terms, blockchain-based trust infrastructures do not store personal data themselves; rather, they record the validity, issuance or revocation status of credentials. This allows users to authenticate directly with a service provider through cryptographic proofs stored locally on their device, without relying on external identity providers. In this decentralised model, the blockchain functions as a distributed trust anchor, enabling the verification of credentials while circumventing traditional authentication intermediaries.[9] The judicious management of information disclosure, in conjunction with the tenet of informational self-determination, has been demonstrated to curtail the risks of data exposure and augment confidence in the identity verification process.

---

[8] *Ibid*.

[9] European Commission, *Proposal for a European Digital Identity*; Renda, *Legal Framework for eID*.

Emerging technologies such as biometric authentication and artificial intelligence complement blockchain-based solutions, offering greater security and personalisation in digital identity management. Furthermore, they enhance the reliability of authentication processes by enabling continuous multi-factor verification and improving the detection of anomalies or fraudulent behaviour.[10]

IV. Biometric Authentication: Potential and Risks

Biometric authentication has been demonstrated to offer substantial advantages in the public sector context, particularly in relation to identity management. Its capacity to provide an effective and reliable verification method has led to its adoption across a wide range of domains, including financial services and law enforcement agencies.[11] It stands out for its ability to offer a higher level of security compared to traditional authentication methods, particularly due to the difficulty of falsifying or replicating biometric traits. Moreover, biometric systems have been shown to enhance user privacy by reducing reliance on repeatedly disclosed credentials and minimising the exposure of personal data during authentication processes.[12] This is due to the fact that biometric traits are considered to be more difficult to falsify or manipulate than other methods, such as passwords or personal identification numbers (PINs). A PIN is a personal identification number that is used as a password to access a system. Examples of the application of this technology include the use of facial recognition in airport environments and the implementation of voice authentication in public services. These systems contribute to the optimisation of administrative and operational processes by streamlining identity verification procedures and reducing the need for manual checks.[13]

Nevertheless, the pervasive implementation of biometric technologies gives rise to substantial concerns regarding privacy and data protection. Biometrics is a term that encompasses a range of distinct

---

[10] Hildebrandt, *Smart Technologies and the End(s) of Law*.

[11] Celik, Salama and Eltawil, "Internet of Bodies".

[12] *Ibid*.

[13] *Ibid*.

technologies that employ probabilistic matching methods to authenticate identities based on physiological or behavioural characteristics.[14] These characteristics can include physiognomic attributes (e.g., the face, iris, fingerprints and hand geometry) as well as behavioural characteristics (e.g. signature patterns, typing style and an individual's gait).

In one-to-one authentication systems, the biometric comparison is made between the individual's characteristics and data stored previously, as exemplified by the use of fingerprints to unlock mobile devices. Biometric authentication can be categorised into two distinct types: active and passive. Active authentication involves the direct participation of the individual in the authentication process, whereas passive authentication refers to the collection of data imperceptibly. An example of passive authentication can be found in voice biometrics during a telephone call, wherein the user's voice characteristics are analysed in an undetectable manner. Behavioural biometrics is the analysis of behavioural patterns. Such patterns may include the manner in which a device is held, the typing pattern (i.e., the manner in which the fingers come into contact with the screen and the force that is applied), movement, behaviour, and the language used (e.g., sentence structure, word choice, or grammar). These data are characterised as a passive form of authentication, which is increasingly used as an additional layer of security to confirm identity.

In mass identification systems, such as those employed in public security contexts and forensic investigations, the biometric characteristics of an unknown individual are compared with a pre-existing database. Examples of this include the use of facial recognition technology in public spaces and the comparison of DNA samples in criminal investigations. These systems can be integrated with other technologies for functions other than identification, such as surveillance and monitoring. This integration involves networks of surveillance cameras and artificial intelligence systems, which carry out continuous monitoring and identify individuals in real time. The integration of automation with artificial intelligence has been demonstrated to enhance the efficacy of biometric processes in authentication and identification, thus

[14] *Ibid*.

positioning it as a significant instrument within the domain of security.[15] Nevertheless, it must be acknowledged that such systems are not without their limitations.

Biometric systems are subject to two main types of error: false positives and false negatives. The former occurs when the system wrongly associates an input with an existing template, and the latter when a valid match is discarded. Given the probabilistic nature of biometric matching, errors may occur due to a number of factors, including similarities in biometric characteristics between individuals (e.g., monozygotic twins), variations caused by ageing, injuries or medical conditions, or even differences in interaction with the sensor between the registration and recognition phases (e.g., changes in body posture). Furthermore, environmental factors such as suboptimal lighting conditions have been shown to negatively affect the accuracy of biometric capture, compromising the reliability of recognition systems.[16] A further critical issue pertains to the biases inherent in the systems' training data, which have the potential to compromise the accuracy and precision of authentication or identification.

Furthermore, the security of biometric data is not guaranteed, as it is vulnerable to forgery attacks, such as spoofing, where false artefacts are used to deceive biometric sensors. Despite the implementation of techniques such as liveness detection to mitigate these risks, vulnerabilities inherent in biometric systems are not completely eliminated. Liveness detection refers to the set of methods used to verify whether the biometric input originates from a real, live individual rather than from an artefact such as a photograph, mask, synthetic fingerprint or recorded voice. These mechanisms analyse physiological cues (such as depth, texture, involuntary movements or temperature) to prevent spoofing attempts.[17]

In addition to the security-related issues that have been extensively documented, the implementation of biometric technologies gives rise to a number of ethical and privacy concerns. These challenges are inherently associated with the architecture and implementation of biometric systems. A significant challenge pertains to the phenomenon of function

---

[15] Celik, Salama and Eltawil, "Internet of Bodies".

[16] *Ibid*.

[17] Celik, Salama and Eltawil, "Internet of Bodies".

slippage, whereby biometric data collected for a specific purpose are subsequently utilised for secondary purposes without the explicit consent of the individual, for example when fingerprints or facial templates initially gathered for physical access control are later repurposed for employee monitoring or productivity assessment.[18] Another significant risk is the clandestine collection of biometric information, such as the unauthorised capture of facial features or fingerprints, which exacerbates privacy violations, especially as technology advances. Examples of this include the collection of facial biometric data from photographs taken at a distance or the taking of fingerprints after the individual has interacted with any surface.

Furthermore, the disclosure of sensitive, non-consented secondary information, including health-related data and genetic predispositions, is a potential consequence of biometric characteristics, thereby challenging the notion of informed consent. In many cases, the collection of biometric data occurs covertly or compulsorily, thereby restricting the individuals' autonomy over their personal information. This process can be regarded as dehumanising, insofar as it reduces the individuals' unique identity to mere computer data, with a potential adverse effect on their social relationships and sense of privacy.

V. Genetic Data and the Expansion of Biometrics

The utilisation of genetic data for identification purposes constitutes one of the most sensitive and controversial frontiers in contemporary biometrics. In contrast to other biometric characteristics, such as fingerprints, facial patterns or typing rhythm, DNA contains vast quantities of information relating to an individual's biological identity. This includes genetic predispositions, ancestry, potential diseases and kinship relations. This abundance of information renders DNA not merely a unique identifier, but a repository of highly sensitive data, the handling of which necessitates particularly stringent safeguards.

Historically, the use of DNA for identification has been confined to the domain of forensics. However, there has been an ongoing shift

---

[18] Maya Wang, China's Algorithms of Repression (Human Rights Watch 2019) https://www.hrw.org accessed 4 July 2025.

towards its application in civil and security contexts. The development of sophisticated genetic databases, the reduction in sequencing costs, and the creation of probabilistic algorithms for the inference of phenotype (skin colour, hair, ethnic ancestry, etc.) have encouraged governments and private companies to explore new forms of genetic identification. These include the use of Forensic DNA Phenotyping in criminal investigations, the large-scale collection of genetic material from minority populations in authoritarian regimes, and the expansion of commercial genetic testing for ancestry and health-related traits.[19] In countries with authoritarian regimes, such as China, the collection of DNA samples from minority populations, under the pretext of national security, has generated international warnings about the violation of human rights.[20] At the same time, concerns associated with disproportionate and non-consensual collection of genetic material have also been documented in democratic contexts, including Portugal, where the expansion of State genetic databases raises questions of proportionality, purpose limitation and the potential normalisation of genetic surveillance.[21]

The uniqueness and immutability of deoxyribonucleic acid (DNA) pose significant legal challenges. Unlike a password or authentication token, genetic material is irrevocable and non-replaceable; once exposed or misused, the harm to privacy is irreversible. The risk of secondary or unauthorised uses of genetic data is therefore particularly acute. Material collected initially for identification may subsequently be repurposed for medical analyses, statistical cross-referencing, criminal profiling or broader forms of social control.[22]

At the European level, the General Data Protection Regulation (GDPR) categorises genetic data as a special category of personal data, necessitating specific legal bases and augmented guarantees for its processing. Nevertheless, the practical effectiveness of these rules is contingent on the existence of supervisory mechanisms, institutional capacity, and a data protection culture: elements that are not always present, especially in security, border control or migration management.

---

[19] WANG, 'Algorithms of Repression'.

[20] *Ibid*.

[21] SÍLVIA DE CARVALHO HOMEM, 'A base de dados de perfis de ADN em Portugal – Deus ex Machina?' (Masters dissertation, Universidade do Minho 2023), P. 31-64.

[22] HOMEM, 'A base de dados de perfis de ADN'.

The introduction of Forensic DNA Phenotyping (FDP) represents a further advance that gives rise to significant ethical concerns. This technique, used to predict physical traits and ancestry from DNA samples, has increasingly been advocated as an investigative aid in criminal proceedings, particularly in cases where no suspect has been identified.[23] Yet its probabilistic nature, combined with the risk of misinterpretation as scientific certainty, creates the potential for miscarriages of justice, discriminatory policing and the reinforcement of structural racial biases.[24]

From the perspective of digital identity, the incorporation of genetic data within authentication or identification systems gives rise to significant inquiries concerning informational self-determination and the extent of State power over the biological body. The digitisation of human biology, along with its integration into interoperable databases, has the potential to precipitate a novel form of algorithmic biopower, insofar as predictive systems and automated infrastructures increasingly shape the ways in which individuals are classified, monitored and governed.[25] This dynamic aligns with what recent scholarship has termed the emergence of a "molecular panopticon," a genetic surveillance structure based in the continuous extraction, interpretation and circulation of biological information.[26]

In such circumstances, it becomes imperative to adopt a judicious legal approach that acknowledges the intrinsic dignity of genetic data, imposes clear limits on their utilisation, and ensures transparency, informed consent, and effective redress mechanisms. The protection of genetic identity is not merely a matter of privacy; it is also a question of justice, equality and freedom in a society increasingly shaped by invisible infrastructures of biological governance.

---

[23] Wang, 'Algorithms of Repression'.

[24] Homem, 'A base de dados de perfis de ADN'.

[25] Hildebrandt, *Smart Technologies and the End(s) of Law*.

[26] Sílvia de Carvalho Homem, 'The Molecular Panopticon: How Law Faces the Rise of Genetic Surveillance' (UMinho Research and Innovation Open Days, Braga, 11–14 November 2025).

VI. Wearable Technology

The integration of tracking and identification technologies through physical devices such as smart bracelets, Radio-Frequency IDentification (RFID) tags and subcutaneous integrated circuits (microchips) adds a tangible dimension to the debate on digital identity. These devices, frequently categorised as smart clothing (wearable technology), effectively transform the human body into a continuous point of emission, authentication and data collection, thereby blurring the conventional boundaries between identity, functionality and surveillance. A notable example of this phenomenon can be observed in the case of certain Swedish companies, where employees have voluntarily implanted subcutaneous microchips to facilitate access to buildings, log schedules and to perform basic transactional functions.[27] Despite being promoted as a practical and innovative solution, this practice gives rise to significant concerns regarding genuine consent, the control of personal data and the preservation of human dignity.

Innovations in technological clothing extend beyond subcutaneous implants to encompass garments and accessories with computer functions. The concept of Smart Clothing has already been realised in the form of a gesture-controlled jacket, for example the model developed by Google in partnership with Levi's, which incorporates Jacquard technology. This jacket enables the user to answer calls, control music and access digital services directly from the fabric. In addition to the aforementioned example, there is an observable emergence of various initiatives that combine technological clothing with the promotion of physical performance and well-being.

For instance, Under Armour has developed the UA RUSH™ line, which is designed with the specific intention of enhancing athletes' performance. This collection employs technologically advanced fabrics infused with minerals that are capable of reflecting body heat in the form of infrared radiation. The purpose of this function is to stimulate blood circulation and contribute to muscle recovery and reduced fatigue, even after physical exertion.

---

[27] Samuel Gibbs, 'Swedish Company Implants Microchips in Employees' *The Guardian* (London, 29 January 2015) https://www.theguardian.com accessed 4 July 2025.

In the context of running, Sensoria's smart socks are distinguished by their incorporation of pressure sensors that monitor the foot's support pattern in real time during the stride. These data facilitate a detailed analysis of running technique, enabling the correction of inappropriate postures and the prevention of injuries, particularly to the joints and heels. The socks function as a digital personal trainer, providing alerts and guidance based on the wearer's biomechanics.

Wearable X proposes an innovative approach to yoga by incorporating sensors into garments that provide tactile and audible feedback during exercises. The Nadi X trousers, a leading product of the brand, provide guided instructions to promote proper posture and precise movement, enabling wearers to practice yoga independently and customised to their individual needs, in any location. This integration signifies a technological advancement and a method of naturalising digital control through conventional clothing.

Wristbands and smart watches, including the Apple Watch, as well as devices utilising RFID (radio frequency identification) technology, have emerged as prevalent instruments for the monitoring of physiological parameters such as heart rate, sleep quality, geographical location and social interactions. The application of these technologies in corporate environments, events, gyms and even schools has given rise to new behavioural surveillance mechanisms. Concurrently, the Near Field Communication (NFC) technology embedded within these devices enables seamless transactions and access through a gesture, thereby rendering the user not only identifiable but also economically traceable.

The fusion of aesthetics and functionality is exemplified by smart jewellery, such as the Oura Ring. These accessories have been developed for the purpose of monitoring health parameters and functioning as continuous authentication tools. For instance, the ring gathers twenty biometric parameters containing sensitive health-related information on a continuous basis via the finger. This almost imperceptible incorporation of biometric identification into personal style gives rise to questions concerning the role of the consumer market in facilitating passive acceptance of the permanent collection of sensitive data.

The concept of computing in smart technological clothing has historical roots in experiments conducted by Steve Mann in the 1980s and 1990s. During this period, Mann developed devices capable of

transmitting video in real time and incorporating screens into glasses and helmets. The aim of these experiments was to achieve continuous interaction between the user and the digital environment. The proposal for personal imaging, which can be considered a form of perpetual, internal computer system, anticipated many of the functionalities offered by contemporary commercial wearable technology. However, the author himself warned of the risks of technological dependence and the possibility of oppressive use of these systems if they were controlled by central entities or applied in a coercive manner.[28]

The social normalisation of these devices, driven by miniaturisation and aesthetically pleasing design, contributes to body control becoming imperceptible, masked by convenience and aesthetics. Indeed, the more sophisticated and unobtrusive the smart technological clothing is, the more challenging it becomes to ascertain its potential impacts on the individuals' privacy, freedom of movement and informational autonomy.

In this context, it is essential to reflect on the ethical principles that should guide the development and implementation of smart technological clothing, ensuring that the human body does not become a continuous authentication terminal at the service of commercial or security interests. It is imperative to comprehend the implications of wearables not merely as a functional innovation, but as a new domain of contention between self-determination and surveillance.

VII. Surveillance Architectures and the Algorithmic Construction of Digital Identity

The increasing pervasiveness of digital technologies in contemporary society has precipitated the emergence of a novel surveillance paradigm, whereby individuals are subject to real-time monitoring through interconnected devices, biometric sensors, facial recognition systems and continuous geolocation mechanisms. This ecosystem of constant traceability, widely promoted as innovative and efficient, is at the basis

---

[28] S Mann, 'Wearable Computing: A First Step Toward Personal Imaging' (1997) 30(2) Computer 25 https://www.researchgate.net/publication/2954704_Wearable_computing_A_first_step_toward_personal_imaging

of what has come to be called digital surveillance architectures — a model of social organisation underpinned by the large-scale collection, analysis and commercialisation of personal data.[29]

In this context, digital identity is not only constructed from data consciously provided by the user, but also (and above all) by inferred data generated through behavioural patterns, preferences, mobility trajectories, interactions and response times to digital stimuli.[30] Consequently, users become only partially aware of the digital identity constructed about them, while remaining highly transparent to the technological platforms that analyse and classify their behaviour.

The concept of surveillance capitalism, as formulated by Shoshana Zuboff, offers a particularly useful lens through which to comprehend this dynamic.[31] Technology platforms function as extractive entities, the purpose of which is the collection of data with the objective of anticipating and modelling future behaviour. This is not solely for the purpose of advertising, but also for the automated decision-making processes that pertain to areas such as credit, public safety, health and employment. In this process, digital identity becomes a continuous algorithmic construction, subject to updates and reinterpretations that are beyond the control of the holder. This occurs, for example, when credit-scoring systems automatically adjust a person's risk profile based on behavioural and contextual data, when advertising platforms infer new preferences from patterns of navigation and attention, or when security and policing systems assign predictive classifications (such as "anomalous" or "high risk") on the basis of movement, sensor-derived information or interaction data.[32] [33] [34]

These practices imply a radical inversion of the principle of informational self-determination. Contrary to the conventional paradigm (understood less as an empirical reality and more as a regulatory aspiration grounded in the idea that individuals should be able to define and

---

[29] Zuboff, *Age of Surveillance Capitalism*.

[30] *Ibid*.

[31] *Ibid*.

[32] *Ibid*.

[33] Oscar H Gandy, *The Panoptic Sort: A Political Economy of Personal Information* (Westview Press 1993), pp. 436-457.

[34] Hildebrandt, *Smart Technologies and the End(s) of Law*.

control the contours of their digital identity) it is now the system that assumes this role. This shift is predicated on algorithmic inferences that are frequently inaccurate, biased or decontextualised. In corporate and institutional environments, the integration of wristbands with radio frequency identification, smart cameras and environmental sensors facilitates the unobtrusive monitoring of workers' movements, productivity and even emotional states, frequently without their full awareness or effective consent.[35] [36]

From a legal standpoint, these practices directly contradict the principles of proportionality, minimisation and purpose outlined in the General Data Protection Regulation (GDPR). The technical feasibility of data collection does not inherently guarantee its legality or legitimacy. Automated processing based on inferred data requires special attention, as it involves increased risks of discrimination, exclusion and error.

Furthermore, the opacity of algorithmic systems exacerbates the asymmetry between users and platforms, as individuals are unable to understand the criteria through which they are classified or the inferences drawn from their data. In many cases, these systems operate as 'black boxes', with limited or no explanation in the applicable terms and conditions — and sometimes without full interpretability even for the companies deploying them. As a result, users are asked to consent to processes whose functioning and implications they cannot meaningfully apprehend. This produces what is often described as conditional or functional consent: acceptance given not out of genuine agreement, but because refusal would result in exclusion from essential services or substantial limitations on access.[37] [38]

In a hyperconnected world, where bodies are permanent data mediators and decisions are delegated to computer systems, there is an urgent need to rethink the foundations of digital identity in the light of fundamental rights. Algorithmic transparency, the right to an explanation and effective control over personal data cannot be regarded as mere

---

[35] Zuboff, *Age of Surveillance Capitalism*.

[36] Gandy, *Panoptic Sort*.

[37] Hildebrandt, *Smart Technologies and the End(s) of Law*.

[38] Zuboff, *Age of Surveillance Capitalism*.

technical requirements, but rather as structural guarantees of a dignified and free digital citizenship.

VIII. Internet of Bodies (IoB): Body, Data and Control in the Age of Cybersurveillance

The increasing convergence between the human body and digital networks has led to the emergence of a new technological frontier known as the Internet of Bodies (IoB). The IoB constitutes a specialised subset of the broader Internet of Things (IoT), comprising interconnected devices located in, on or around the human body, with the capacity to collect, transmit or modify physiological, genetic or behavioural data. These technologies include wearable sensors, implantable microchips, ingestible digital pills and brain-machine interfaces. Wearable devices therefore represent only one segment of the IoB, which also encompasses implantable, ingestible and embedded systems. The expansion of the IoB marks a significant development in continuous monitoring and ubiquitous connectivity, with profound implications for health, safety, privacy and individual self-determination.

The concept of the Internet of Bodies encompasses devices with varying degrees of invasiveness and functionality, and can be classified into four main categories: (a) Smart wearables — fitness bracelets, smart watches, connected textiles, movement or heart rate sensors. b) Ingestible devices — capsules equipped with cameras or sensors for the purpose of gastrointestinal diagnosis or monitoring the ingestion of medications. c) Implantable devices — connected pacemakers, neurostimulators, and subcutaneous microchips utilised for authentication purposes. d) Embedded devices — systems integrated into the skin or organs, capable of interacting with the nervous or endocrine systems.

The primary function of these devices is the collection of biometric data or the modulation of bodily functions, thereby creating personal communication networks (Wireless Body Area Networks — WBANs) that integrate sensors, transmitters and interfaces modules. In typical IoB architectures, the data generated by these devices are relayed to external infrastructures (such as cloud platforms or edge-computing

nodes) where they may be processed in real time using artificial intelligence algorithms.[39]

The health sector is currently the primary beneficiary of IoB solutions. These technologies enable continuous remote monitoring of patients with chronic conditions, facilitate non-invasive diagnostics based on physiological parameters such as perspiration, heart rate, glucose levels or breathing patterns, and support personalised and predictive medicine through the integration of genetic and behavioural data into therapeutic plans. They also contribute to improving the quality of life of persons with disabilities by means of smart prostheses and neurofunctional implants.[40]

The global Covid-19 pandemic further highlighted the centrality of IoT infrastructures in public health, particularly through their use in contact-tracing systems, remote symptom monitoring and the management of healthcare facilities. These developments have contributed to consolidating the IoT as an essential component of contemporary health and emergency-response architectures.[41]

The expansion of the Internet of Bodies gives rise to a number of significant legal and ethical challenges, particularly with regard to privacy, security and informed consent. The operation of IoB devices is predicated on the collection of highly sensitive data, including cardiac parameters, sleep patterns, menstrual cycles and even brain signals. The secondary use of such data is possible in the absence of the individual concerned.

Case law demonstrates the utilisation of biometric data from body devices in legal proceedings. A notable example is the US case of Ross Compton, wherein information extracted from a pacemaker was employed as criminal evidence. The aforementioned precedents demonstrate the potential of the IoB to transform the human body into a digital witness. A widely cited example is the US case State v Ross Compton, in which data extracted from a pacemaker were admitted as criminal evidence to contradict the defendant's version of events.[42] Similar develop-

---

[39] Celik, Salama and Eltawil, 'Internet of Bodies'.

[40] *Ibid*.

[41] *Ibid*.

[42] BBC News, 'Pacemaker Data Used to Charge Man with Arson' (9 February 2017) https://www.bbc.com\ accessed 4 July 2025.

ments can also be observed in cases involving data from smartwatches and fitness trackers, illustrating a broader trend toward the forensic use of bodily-integrated or body-generated data. This evolution raises concerns for fundamental rights, particularly the privilege against self-incrimination.

Furthermore, algorithmic surveillance exercised by employers, insurers or public authorities through these devices has the potential to lead to subtle forms of discrimination or social exclusion, especially when applied to vulnerable populations with lower digital literacy or unequal access to technology.

From a technical standpoint, IoB devices are subject to similar risks to other connected objects, including software update failures, weak passwords, poor encryption, and known vulnerabilities in legacy systems. However, the critical nature of certain devices, including insulin pumps, defibrillators and brain implants, necessitates a heightened cybersecurity approach to mitigate the risk of compromising user physical integrity.

From a legal perspective, the regulation of the Internet of Bodies is dispersed and incomplete. Whilst medical devices are subject to oversight by the U.S. Food and Drug Administration (FDA) in the USA and the Medical Device Regulation (MDR) in the European Union, wellness or fitness devices frequently evade regulatory scrutiny, as they fall outside the scope of the Medical Device Regulation despite processing physiological and health-related data that qualify as "special categories of personal data" under the GDPR.[43] Their legal treatment therefore contrasts sharply with that of certified medical devices, which are subject to stringent safety, accuracy and oversight requirements.

Moreover, extant regulatory frameworks such as the European Union's GDPR were conceived before the widespread diffusion of smart clothing, implantable sensors and other IoB technologies. As a result, the GDPR regulates the data produced by these systems but does not address the broader techno-material context in which bodily data are generated: namely the continuous integration of sensing devices with the body, the real-time transmission of physiological information to external infrastructures, or the emergence of algorithmic inferences

[43] Regulation (EU) 2016/679.

derived from intimate biological signals. These gaps raise fundamental questions concerning consent, proportionality, bodily autonomy and the limits of technological intervention on the human body, illustrating the need for a more comprehensive legal framework capable of governing these novel forms of bodily interaction with digital systems.[44]

The Internet of Bodies marks a paradigm shift in the manner in which the law perceives the nexus between the body, technology and data. The IoB concept involves the transformation of the human body into a permanent node of the digital network. This development necessitates a re-evaluation of the limitations of technological control, the foundations of legal subjectivity, and the normative architecture that can safeguard human dignity in the era of total connectivity.

IX. Emotion Recognition and the Fragility of Automated Interpretation

The increased use of biometric technologies has given rise to concerns regarding the violation of both territorial and bodily privacy, especially in the context of mass surveillance. Facial Emotion Recognition (FER), a branch of biometrics, poses a number of specific and complex challenges, particularly due to the epistemic fragility of automated interpretations of human behaviour and the variability of emotional expression across individuals and groups.[45] Despite the potential of this technology to personalise services and enhance public safety, its precision is frequently contested, particularly in relation to the variability of facial expressions across individuals and cultural groups. Furthermore, the process of facial emotion recognition has the potential to inadvertently engender discrimination if the underlying algorithms are trained with biased data sets. This can result in erroneous interpretations that disproportionately affect specific demographic groups, particularly when the underlying datasets reflect structural biases or cultural assumptions embedded in the training process of such systems.[46] [47]

---

[44] Regulation (EU) 2016/679.

[45] Hildebrandt, *Smart Technologies and the End(s) of Law*.

[46] *Ibid*.

[47] Zuboff, *Age of Surveillance Capitalism*.

The following list provides a non-exhaustive overview of the potential applications of facial emotion recognition: marketing (e.g., on billboards); assessment of psychological disorders (e.g., psychosis, depression, autism, neurodegenerative diseases); identification of candidate disinterest during job interviews; monitoring of employee attention; detection of suspicious behaviour in retail environments to prevent theft or fraud; verification of statements in interrogation processes; intelligent border control; predictive screening of public spaces to identify emotions associated with potential terrorist threats; analysis of political attitudes.

The ability of such applications to demonstrate the capacity for facial emotion recognition, in conjunction with other forms of biometrics, to create automated profiles and make decisions based on emotional data is evident, particularly when such systems rely on inferred behavioural patterns to classify individuals and trigger automated responses.[48] [49] Nevertheless, this can result in undesirable or inequitable manipulation, which has the potential to undermine individual autonomy and erode public confidence in institutions, particularly when decisions based on emotional inference lack transparency or democratic accountability.[50] [51]

Moreover, the management of digital identities is undergoing a substantial transformation, shifting from centralised and often vulnerable architectures to decentralised or distributed models that increasingly rely on technologies such as blockchain and biometric authentication.[52] [53] While these innovations hold great promise in terms of enhancing security and privacy, they also give rise to significant ethical concerns, particularly with regard to the use of biometric data and the potential for invasion of privacy. The integration of emerging technologies, including artificial intelligence, biometric authentication and self-sovereign

---

[48] Hildebrandt, *Smart Technologies and the End(s) of Law*.

[49] Zuboff, *Age of Surveillance Capitalism*.

[50] Hildebrandt, *Smart Technologies and the End(s) of Law*.

[51] Zuboff, *Age of Surveillance Capitalism*.

[52] Renda, *Legal Framework for eID*.

[53] European Commission, Proposal for a Regulation amending Regulation (EU) No 910/2014 as regards establishing a framework for a European Digital Identity COM(2021) 281 final.

identity, holds the potential to enhance security and the autonomy of individuals in the management of their digital identities. However, it is crucial that the development and implementation of these solutions be carried out on the basis of robust ethical principles, ensuring that the privacy and fundamental rights of individuals are preserved. It is imperative to emphasise that establishing a digital identity management system that is both secure and reliable, while also respecting human dignity, necessitates a balanced approach.

The utilisation of facial emotion recognition to deduce emotional states from micro-expressions possesses legitimate applications, including supporting medical diagnoses and enhancing public security systems. Nevertheless, the indiscriminate use of this technology has the potential to generate biased outcomes and produce automated decision-making that is ethically questionable and inequitable. As an illustrative example, systems inspired by models such as Fogg Engagement Rating (FER) have been proposed for use in recruitment processes to evaluate candidates' motivation or emotional congruence during interviews. Such approaches risk excluding individuals with distinct neurological or cultural profiles, or those who do not conform to the normative behavioural parameters embedded in these systems.

X. Final considerations

The evolution of digital identity systems has been marked by a progressive shift from documentary identification to bodily and algorithmic identification. Technologies such as blockchain, biometrics, artificial intelligence and smart technological clothing have transformed the human body into a central element within authentication and control architectures. The advent of the Internet of Bodies intensifies this transformation, positioning the body as a permanent data interface subject to continuous monitoring, remote intervention and secondary uses of highly sensitive biological information.

This emerging paradigm raises profound challenges for the protection of privacy, informational self-determination and the physical and psychological integrity of individuals. The pervasive and often imperceptible nature of the Internet of Things obscures the boundaries of meaningful consent and autonomy, exposing individuals to subtle yet

increasingly effective forms of surveillance, profiling and exclusion. This is similar to what has been described as molecular panopticism: a configuration in which surveillance no longer relies on visible spatial control, but on the continuous extraction, circulation and algorithmic interpretation of biological, physiological and behavioural data. In this model, the body becomes the primary locus of monitoring and prediction, reinforcing the structural asymmetry between individuals and the systems that govern them.[54]

In light of these developments, it is imperative to articulate a robust legal and regulatory response capable of embedding the principles of proportionality, purpose limitation and data minimisation within emerging technological infrastructures. Ensuring the protection of digital identity in the IoT and IoB environments requires a multifaceted framework that combines technological standards, legal safeguards and an ethical and political reaffirmation of the foundational values of the democratic rule of law. Principles such as human dignity, autonomy and freedom must constitute the normative core of any system intended to guarantee the integrity, security and legitimacy of digital identities in an era increasingly marked by ubiquitous data extraction and algorithmic inference.

References

A. Celik, K. N. Salama and A. M. Eltawil, "The Internet of Bodies: A Systematic Survey on Propagation Characterisation and Channel Modeling" (2021) *IEEE Internet of Things Journal* https://www.researchgate.net/publication/353270324_The_Internet_of_Bodies_A_Systematic_Survey_on_Propagation_Characterization_and_Channel_Modeling

Andrea Renda, *The Legal Framework for eID and Trust Services in the EU* (CEPS Research Report 2018/01, 2018).

BBC News, 'Pacemaker Data Used to Charge Man with Arson' (9 February 2017) https://www.bbc.com\ accessed 4 July 2025.

European Commission, *Proposal for a Regulation amending Regulation (EU) No 910/2014 as regards establishing a framework for a European Digital Identity* COM(2021) 281 final. https://www.europeansources.info/record/proposal-for-a-regulation-amending-regulation-eu-no-910-2014-as-regards-establishing-a-framework-for-a-european-digital-identity/

[54] Homem, 'The Molecular Panopticon'.

MAYA WANG, "China's Algorithms of Repression" (Human Rights Watch, 2019) https://www.hrw.org accessed 4 July 2025.

S. MANN, "Wearable Computing: A First Step Toward Personal Imaging" (1997) 30(2) *Computer 25*. https://www.researchgate.net/publication/2954704_Wearable_computing_A_first_step_toward_personal_imaging

MIREILLE HILDEBRANDT, *Smart Technologies and the End(s) of Law: Novel Entanglements of Law and Technology* (Edward Elgar 2015).

OSCAR GANDY, *The Panoptic Sort: A Political Economy of Personal Information* (Westview Press 1993).

PAUL OHM, "Broken Promises of Privacy: Responding to the Surprising Failure of Anonymisation " (2009) 57 *UCLA Law Review* 1701.

Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 (General Data Protection Regulation) [2016] OJ L119/1.

SAMUEL GIBBS, "Swedish Company Implants Microchips in Employees" *The Guardian* (London, 29 January 2015) https://www.theguardian.com accessed 4 July 2025.

SHOSHANA ZUBOFF, *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power* (PublicAffairs 2019).

SÍLVIA DE CARVALHO HOMEM, 'A base de dados de perfis de ADN em Portugal — Deus ex Machina?' (Masters dissertation, Universidade do Minho 2023).

SÍLVIA DE CARVALHO HOMEM, 'The Molecular Panopticon: How Law Faces the Rise of Genetic Surveillance' (UMinho Research and Innovation Open Days, Braga, 11-14 November 2025).

Contents/Índice

Nota Prévia
Henrique Sousa Antunes, Luca Belli,
Filipa Urbano Calvão, Yasmin Curzi,
Walter Britto Gaspar, Eduardo Magrani,
Filipe Medon 5

Regular a Inteligência Artificial: da necessidade à dificuldade
A. Betâmio de Almeida 6

Data Protection Compliance in Messaging Apps in Brazil:
Measuring the Effectiveness of the Brazilian General Data
Protection Law
Luca Belli, Walter Britto Gaspar, Bianca Kremer,
Rodrigo Gomes, Beatriz Costa, Fernando Naegele,
Sofia Chang Nogueira and Daniel Dore Lage 21

Freedom of Expression and its Limits in Brazil:
A Historical Legislative Overview Between 1964 and 2021
Fernanda Carvalho Dias de Oliveira Silva,
Nicole de Barros Moreira Reis 59

A responsabilidade das plataformas pelo conhecimento
(presumido): subsídios para uma compreensão do artigo 16.º,
n.º 3, do Regulamento dos Serviços Digitais
Inês Neves 92

Estados privados e soberania digital: a regulação de plataformas
globais no Brasil e na América Latina
Rodrigo Ardissom de Souza 136

Digital Identity Management: Emerging Technologies and
the Future of Security and Privacy
Sílvia de Carvalho Homem 165

Authors

A. Betâmio de Almeida | Luca Belli |
Walter Britto Gaspar | Bianca Kremer |
Rodrigo Gomes | Beatriz Costa |
Fernando Naegele | Sofia Chang Nogueira |
Daniel Dore Lage | Fernanda Carvalho Dias de
Oliveira Silva | Nicole de Barros Moreira Reis |
Inês Neves | Rodrigo Ardissom de Souza |
Sílvia de Carvalho Homem

Organization

CATÓLICA FACULDADE DE DIREITO ESCOLA DE LISBOA

CATÓLICA RESEARCH CENTRE FOR THE FUTURE OF LAW LISBOA

FGV DIREITO RIO CENTRO DE TECNOLOGIA E SOCIEDADE

Sponsors

Abreu: advogados

Instituto de conhecimento: Powered by Abreu Advogados