

RICARDO SILVESTRE

**Liberdade de expressão política na era digital: ameaças
e soluções**

Orientador: Prof. Doutor Vasco Ferreira Rato

**Universidade Lusófona de Humanidades e Tecnologia
Faculdade de Ciências Sociais, Educação e Administração**

Lisboa

2021

RICARDO SILVESTRE

**Liberdade de expressão política na era digital: ameaças
e soluções**

Dissertação defendida em provas públicas para obtenção do Grau de Mestre em Diplomacia e Relações Internacionais, no Curso de Mestrado em Diplomacia e Relações Internacionais, conferido pela Universidade Lusófona de Humanidades e Tecnologias, no dia 14 de junho de 2021, com o Despacho Reitoral nº 127/2021, do dia 13 de abril de 2021, com a seguinte composição do júri:

Presidente: Prof. Doutor Sérgio Vieira da Silva

Arguente: Prof^a. Doutora Susana Salgado (ICS/UL)

Orientador: Prof. Doutor Vasco Fernando Ferreira Rato

Universidade Lusófona de Humanidades e Tecnologia

Faculdade de Ciências Sociais, Educação e Administração

Lisboa

2021

Epígrafe

Acredite na verdade. Abandonar factos é abandonar a liberdade.
Se nada é verdade, então ninguém pode criticar o poder,
porque não existe nenhuma base para o fazer.
Se nada é verdade, então tudo é espetáculo

Timothy Snyder

Agradecimentos

Ao Prof. Doutor Vasco Rato: pela orientação, pelos desafios mentais, e por sempre me dizer para analisar o fenómeno político com mais do que uma perspetiva.

Ao Prof. Doutor Sérgio Silva: pelos conselhos e sugestões, permanente disponibilidade e frontalidade.

Ao Pedro Castanheira (pela segunda vez em teses de mestrado) e Ana Tavares: pela ajuda incansável e eterna paciência em me ajudar a não cometer erros básicos, e complexos, na escrita.

Resumo

Pode a democracia sobreviver à internet é uma pergunta vital. Está o debate político *online* a deteriorar o processo democrático, com o aumento do extremismo e polarização? Democracias liberais assentam em liberdades como a de expressão. O usufruto desta é condição para outras formas de liberdade, e é crucial para uma sociedade informada e participativa em processos de deliberação política. Sempre existiram condicionalismos a liberdade de expressão e ao debate público, em forma de testes para determinar limites para censura de discurso e de conteúdos. Para isso foram determinados princípios como de ofensa, dano e discurso de ódio. Com a emergência de plataformas de redes sociais, em especial Facebook e Twitter, houve uma transferência do debate político da praça pública para o mundo digital. Isso levanta questões sobre normas, regulamentos e legislação, a necessidade de reforçar a arena de discussão e o mercado de ideias, e ter uma especial atenção ao funcionamento de algoritmos e arquiteturas digitais. Os objetivos desta Dissertação são de perceber quais as ameaças à liberdade de expressão e possibilidade em participar no debate político *online*, assim como quais são os passos necessários para reforçar a segurança, garantir acesso, e democratizar o processo.

Palavras-chave: liberdade de expressão, democracia, debate político, internet, redes sociais, Facebook, Twitter

Abstract

Can democracy survive the internet it's a vital question. Is the political debate online deteriorating the democratic process, with the rise of extremism and polarization? Liberal democracies are based on freedoms such as freedom of expression and it's a condition for other forms of freedom It's also crucial for an informed and participative society in political deliberation processes. There have always been constraints on freedom of expression and public debate, in the form of tests to determine limits for censoring discourse and content. For this, principles such as offense, harm and hate speech were determined. With the emergence of social media platforms, especially Facebook and Twitter, the political debate has shifted from the public square to the digital world. This raises questions about rules, regulations and legislation, the need to strengthen the discussion arena and the market for ideas, and to pay special attention to the operation of digital architectures and algorithms. The objectives of this Dissertation are to understand the threats to freedom of expression and the possibility of participating in the online political debate, as well as what are the necessary steps to strengthen security, guarantee access, and democratize the process.

Keyword: freedom of expression, democracy, political debate, internet, social networks, Facebook, Twitter

Lista de Abreviaturas e Siglas

CCPJ – Comissão da Carteira Profissional de Jornalistas

CEDH - Convenção Europeia dos Direitos Humanos

CEO – *Chief Executive Officer* (Diretor Executivo Empresarial)

ECRI - Relatório da Comissão Europeia contra o Racismo e Intolerância

ERC - Entidade Reguladora para a Comunicação Social

FTC - *Federal Trade Commission* (Comissão Federal Comércio Americana)

GSVimNetz - Lei Austríaca do Rigor e Responsabilidade na Internet

IP's - *Internet Provider* (Ponto de acesso à Internet)

LGBT - Comunidade de Lésbicas, Gays, Bissexuais e Transsexuais

NetzDG - Ato Alemão da Aplicação da Lei na Internet

SIN - Sociedade Civil para Iniciativa em Política de Drogas Leves (Organização Polaca)

TEDH - Tribunal Europeu dos Direitos Humanos

Índice

Epígrafe.....	2
Agradecimentos.....	3
Resumo.....	4
Abstract	5
Lista de Abreviatura e Siglas.....	6
Introdução	9
Capítulo 1 – Liberdade de Expressão	14
1.1 Como se caracteriza	14
1.2 Como se tornou um direito.....	17
1.3 Um direito <i>online</i>	21
Capítulo 2. Condicionalismos à liberdade de expressão.....	23
2.1 Como operar uma escala deslizando.....	25
2.2 O princípio da ofensa.....	27
2.3 Princípio do discurso de ódio.....	32
2.4 Princípio do dano.....	36
2.5 Discurso político e liberdade de expressão.....	39
Capítulo 3. A era digital e a liberdade de expressão.....	41
3.1 A evolução do discurso <i>online</i> . Da utopia à distopia.....	41
3.2 Quem censura o quê.....	43
Capítulo 4. As plataformas de redes sociais e o debate político.....	46
4.1. Liberdade de expressão nas plataformas digitais.....	46
4.2 Os novos gigantes sociais.....	48
4.2.1 Facebook	52
4.2.2 Twitter	57
4.3 A tribalização do debate político <i>online</i>	61
Capítulo 5. Discurso político <i>online</i> e soberania	64
5.1 A questão da União Europeia	66
5.2 Leis, normas e reguladores	68
5.3 Qual o melhor equilíbrio?	71
Capítulo 6. Criar condições para o debate político <i>online</i>.....	74
6.1 Competências digitais do cidadão	74
6.2 Melhorar a “arena de discussão” <i>online</i>	78
6.3 Alimentar o mercado de ideias	81
6.4 Qual a melhor regulação?.....	83

6.4.1 Regulação de algoritmos	86
6.4.2 Regulação da arquitetura digital	87
Capítulo 7. Sistematização dos mecanismos para proteção	
de discurso e/ou conteúdos políticos online.....	90
7.1 Leis e o sistema judicial	91
7.2 Plataformas digitais de redes sociais.....	93
7.3 Utilizadores e <i>stakeholders</i>	95
7.4 Quadros regulamentares	97
Conclusão.....	98
Bibliografia	102

Introdução

O cidadão comum, desde o advento das plataformas digitais que proporcionam um espaço público para a divulgação de opiniões, viu reunidas as condições para participar com muito maior facilidade, frequência e alcance em debates públicos de diferentes naturezas. Em particular, este cidadão pode ter opiniões políticas que acredita serem de benefício para outros participantes nos mesmos fóruns na *online*. Beneficiando da liberdade de expressão, coloca uma mensagem, ou vulgarmente conhecido um *post*, nas duas plataformas digitais de redes sociais mais utilizadas no mundo. Passadas algumas horas, é confrontado com mensagens em ambas as plataformas. No Facebook, a mensagem será: “[p]arece que colocou algo que não segue os nossos Padrões Comunitários. Removemos *posts* que atacam devidos à sua raça, etnicidade, origem nacional, afiliação religiosa, orientação sexual, género ou deficiência”. No Twitter, a mensagem é igualmente ominosa: “[a] sua conta foi suspensa por ter violado as Regras do Twitter. Especificamente violou as nossas regras por conduta odiosa. Você não pode promover violência, ameaçar ou assediar para com outras pessoas baseadas nas respetivas raças, etnicidades, origem nacionais, orientação sexual, género, identidade de género, afiliação religiosa, idade, deficiência ou doença séria”. Porém, para o cidadão, a mensagem não incitava a ódio. Não promovia violência ou fazia ameaças à integridade física de outros. Talvez o conteúdo pudesse ser considerado como rude, ou ofensivo, ou divergente do consenso generalizado. Contudo, parece faltar justificação para uma censura tão conclusiva, para além de causar as seguintes interrogações: a mensagem foi censurada por quem? E com que legitimidade? Chegámos a um ponto onde “guardiões digitais” são agora os juízes, júris e executores do que é permissível? Estaremos no ponto onde se condiciona o exercício da livre expressão para proteção daqueles que se sentem importunados com ideias diferentes? Estaremos numa situação onde se altera, ou suprime, discurso e/os conteúdos para que não sejam apresentadas queixas e pedidos de remoção de opiniões e argumentos?

As democracias liberais assentam em liberdades, como é a liberdade de expressão. Eis uma condição *sine qua non* para uma sociedade informada, participativa e com mecanismos robustos para fomentar o debate público. Juntamente com o direito de produzir, disseminar e receber informação, reúnem-se as condições para o desenvolvimento de conhecimento, procura da verdade e oposição a argumentos de autoridade. O debate de ideias é crucial para o progresso civilizacional, onde todas as opiniões, no plano de princípios, são importantes, pois proporcionam a oportunidade de reavaliar posições, ou melhor, as compreender e reforçar. Retirar estes direitos por causarem incómodo é condicionar, à

partida, o bom funcionamento da arena de discussão e do mercado de ideias¹. O *Teste de Voltaire*², “[e]u desaprovo aquilo que disse, mas defenderei até à morte o seu direito de o dizer” (Campese, 2015) tem servido bem a sociedade, mas precisa de ser reforçado na era digital. No caso *Palko v. Connecticut*, o Juiz Benjamim Cardoso escreveria que “dessa liberdade [de expressão], podemos dizer, é a matriz, a condição indispensável, de quase todas as outras formas de liberdade” (USSC, 1937).

No entanto, é também esperado que direito à liberdade de expressão não seja absoluto. Para uma saudável convivência em sociedade, e da salvaguarda do interesse público e privado, há limitações impostas a tipos de discurso e/ou conteúdos por serem considerados danosos. Essas limitações existem em cartas e acordos, com especial interesse para este trabalho no *Pacto Internacional sobre os Direitos Civis e Políticos*, tornado lei em Portugal através da Lei n.º 29/78 (Diário República, 1978). No artigo 19.º, ponto 3, salvaguarda-se o respeito e reputações de outros, e a proteção da segurança nacional, da ordem pública, e da saúde e moralidades públicas (UNHRC, 1966). Estes limites causam, inevitavelmente, uma tensão entre a defesa do princípio da igualdade e o usufruto da livre expressão.

O *Pacto Internacional sobre os Direitos Civis e Políticos* não garante o direito absoluto do exercício da liberdade de expressão. Requer que sejam criadas condições para se alcançarem os direitos incluídos na *Declaração Universal dos Direitos Humanos*. Os mecanismos legais descritos em leis internacionais exigem que testes para restrições da liberdade de expressão tenham padrões elevados de exigência, que sejam codificadas em leis, e que essas leis sejam acessíveis e “formuladas com precisão suficiente para permitir aos cidadãos regularem a sua conduta” (ECHR, 1979). Igualmente, as restrições devem ser comprovadas como “necessárias”, ou seja, com uma “necessidade social premente”, e que as razões para tal sejam “relevantes e suficientes” com proporcionalidade para os benefícios e os malefícios (ECHR, 1986). Contemplado nestes documentos encontram-se também o direito de se deter uma opinião, e de não se ser inquietado pela expressão da mesma, sendo necessário aplicar os testes anteriormente referidos para eventuais restrições (UNHRC, 2006). A pergunta a que é preciso responder é a seguinte: como podem normas, regras e leis ser adaptadas a uma realidade onde o discurso societal, e em particular o discurso político, passou das tertúlias em salões, de conversas em cafés, de ajuntamentos em avenidas, bairros

¹ Conceitos a serem melhor definidos posteriormente, mas resumidamente, “arena de discussão” são os espaços onde acontecem as trocas de argumentos e de opiniões, e o “mercado das ideias” é uma analogia entre o uso da liberdade de expressão e o conceito económico de regulação de mercados por oferta e procura.

² Na verdade, tais palavras são atribuídas a Evelyn Beatrice Hall, que descrevia assim qua era a convicção do filósofo francês.

e praças públicas para o “mundo digital”. A Internet foi idealizada para ampliar os direitos humanos. Como sugerido num relatório das Nações Unidas, a Internet “facilita a concretização de um grupo de outros direitos humanos” (La Rue, 2011: 7). É isso que se verifica? Existe uma expansão dos direitos, ou, pelo contrário, uma contração dos mesmos?

Um estudo realizado em Portugal, no ano de 2018, pelo Reuters Institute, em colaboração com a OberCom para o estudo do jornalismo, mostrou que num universo de 2008 inquiridos, 33.9% dos portugueses dizem ter como fonte noticiosa preferencial a Internet, na forma de jornais on-line e motores de busca, enquanto 15.3% encontra essa informação em redes sociais. Uma percentagem de 42.6% dos inquiridos disseram que consultam a Internet para notícias duas a cinco vezes por dia, 14.1% 6 a 10 vezes ao dia e 12% mais de 10 vezes ao dia. O estudo também mostra que 29.9% produzem comentários às notícias que observam, e 36% mostram-se preocupados com eventuais problemas no exercício da liberdade de expressão na Internet (Reuters Institute, 2018). Inevitavelmente, com a progressiva penetração da Internet na sociedade portuguesa, começam a observar-se os primeiros resultados da tensão entre o uso do mundo digital, e em especial das plataformas digitais de redes sociais, e a liberdade de expressão e da produção de discurso e conteúdos. Foi assim no caso da vereadora da Câmara do Porto, Odete Patrício, julgada por crimes de ofensa à pessoa coletiva do Município por causa de um *post* no Facebook, ou da ex-deputada Portuguesa no Parlamento Europeu, Ana Gomes, ameaçada de processo judicial por um clube de futebol por causa de um comentário deixado no Twitter³.

A existência *online* de discurso e/ou conteúdos de natureza ilegal tem consequências negativas para a sociedade. Na arena de discussão e no mercado de ideias, onde existe a troca de argumentos, opiniões e informação, ação essencial numa democracia participativa, e à medida que a Internet se transforma na “Internet das coisas”, ubíqua e inescapável, é necessário assegurar que direitos e normas que estão plasmados em cartas e atos, ou que são aceites maioritariamente, continuam a aplicar-se. E, se julgado necessário, como podem ser criados mecanismos para a sua defesa. Em particular para este trabalho, como se o pode fazer com plataformas digitais que proporcionam o serviço de “fórum público” através de redes sociais. As empresas responsáveis por estes fóruns social precisam de tomar a iniciativa, ou

³ No segundo caso, o clube em questão ficou-se pela ameaça (<https://www.slbenfica.pt/pt-pt/agora/noticias/2019/07/05/clube-benfica-comunicado-processo-judicial-contr-ana-gomes>), já no primeiro, a queixa foi apresentada em tribunal, resultando numa sentença que deu provimento ao recurso da arguida (<http://www.dgsi.pt/jtrp.nsf/56a6e7121657f91e80257cda00381fdf/1434e0c05305aa67802585a7003a2f58?OpenDocument>).

ver-se-ão forçadas a isso. Afinal, essas são as liberdades que garantem que os utilizadores usem as plataformas para contribuir para o que acham ser o bem comum.

Em novembro de 2019, Singapura mostrou como criar, e aplicar, uma lei que deve preocupar todos os defensores de liberdade de expressão, e da apresentação de ideias e opiniões no debate político. O governo dessa cidade-estado decretou que certas opiniões contrárias a versões oficiais governamentais fossem rotuladas com “declarações falsas”, afirmado como tal com um “carimbo virtual”, colocado sobre esses *posts* no Facebook. A lei, denominada de *Ato de Proteção Contra Falsidades e Manipulações*, foi aprovada pelo governo pela necessidade de “parar desinformação perigosa e conteúdos odiosos” (Griffiths, 2019). Bem mais próximo, países que fazem parte da União Europeia já aprovaram, ou estão em vias de o fazer, leis que não sendo tão draconianas como o exemplo apresentado anteriormente, mostram sinais preocupantes. Na Áustria, a lei do *Rigor e Responsabilidade na Internet* força as plataformas digitais a armazenar os dados dos utilizadores na eventualidade de ter de os entregar às autoridades policiais, ou a utilizadores que desejam fazer uma queixa criminal sobre uma possível ofensa. Na Alemanha, encontra-se em vigência a Lei NetzDG, que obriga plataformas digitais com mais de dois milhões de utilizadores a rever, e apagar, conteúdos que sejam considerados como discurso de ódio, ou que contenham suposta informação falsa (Die Bundesregierung, 2019).

Em França, uma proposta legislativa parecida, conhecida como a *Lei Avia*, propõe combater discurso e conteúdos que sejam “manifestamente ilegais” em questões de raça, religião, sexo ou orientação sexual (France Press, 2019). De acordo com esta proposta de lei, plataformas digitais de redes sociais serão consideradas como “aceleradores de conteúdo”, alterando a moldura legal da responsabilidade destas. O Governo britânico apresentou um *White Paper* onde propõe um regulador central capaz de estabelecer, e impor, regras com proibições na liberdade de expressão que seja ilegal ou socialmente prejudicial (UK Government, 2019). Até mesmo a Comissão Europeia reviu a diretiva em vigor que regula o comércio eletrónico, e com isso a ação de plataformas digitais, para criar o *Ato de Serviços Digitais* (European Commission, 2020). Este novo ato reflete quais as preocupações de legislação na União Europeia para o funcionamento de plataformas digitais, privacidade, serviços de armazenamento, tratamento de dados, e liberdade de expressão na Internet (de Vries et al, 2019).

Pode, no entanto, ser-se, ainda, um absolutista da livre expressão? E o que quer dizer isso dizer sobre o direito a defender a expressão de ideias e opiniões? E absolutista em relação a quê? Necessariamente em ligação a proteções constitucionais ou legais, tanto

nacionais como internacionais. Será possível conceber, no mundo digital, um debate político público com diferenças de opiniões apresentadas de uma forma construtiva, processo fundamental para garantir o funcionamento do processo democrático? Ou poderão as plataformas digitais, principalmente as de redes sociais, terem alterado definitivamente o conceito de liberdade de expressão para outro tipo de absolutismo: da proteção a qualquer preço de sensibilidade, do acautelar toda e qualquer ofensa, de resguardar todos aqueles que se queixam de uma forma continuada ou estridente?

Os objetivos desta dissertação são produzir um corpo de conhecimento, em Português, de como entender liberdade de expressão mediante a sua evolução histórica: como se definiram os condicionalismos e como esses se estenderam para a arena de discussão política. Com o nascimento da era digital, como é que esses condicionalismos se tornam em perdas de direitos *online*. E finalmente, como lidar com discurso iliberal e com as ameaças à liberdade de expressão política no mundo digital, principalmente no que se relaciona com plataformas digitais de redes sociais. A metodologia utilizada é de uma *estratégia de investigação qualitativa*, com a recolha de informação com o objetivo de desenvolver um significado associado ao fenómeno de debate político no ecossistema digital, e dos seus participantes (Unikaitė-Jakuntavičienė & Rakutienė, 2013). Esta estratégia permite criar uma narrativa construtivista, que visa desenvolver uma teoria de uma forma dedutiva, começando por factos específicos, observação empírica e avançando para uma generalização teórica dos factos relacionados com a teoria. Igualmente, será aplicada uma *investigação científica qualitativa*, com a análise do comportamento dos diferentes agentes envolvidos na construção da teoria, assim como valores, crenças e emoções. Tal será feito através da observação, análise de discursos, de documentos e pareceres de organizações governamentais e da sociedade civil e de artigos noticiosos. A lógica da investigação é assim indutiva, com ponto de partida de cognição da realidade, conceitos flexíveis e generalizações analíticas com a ajuda de exemplos (Unikaitė-Jakuntavičienė & Rakutienė, 2013).

Capítulo 1. Liberdade de expressão

1.1 Como se caracteriza

Para John Stuart Mill, na obra *On Liberty*, uma opinião censurada é uma oportunidade perdida da procura da verdade, se não absoluta, perto daquilo que possa ser aceite como uma verdade universal, que serve de enquadramento para a tomada de decisões e de produções de ações. Para Mill, um dos problemas mais sérios a impedir o progresso é a eventualidade de um “sono profundo de uma opinião decidida” (Mill, 1859: 41). As consequências positivas para o indivíduo, e para a sociedade, do exercício da sua liberdade de expressão são conhecidos e consensualmente aceites, pelo menos em democracias participativas. Por exemplo, é uma forma de concretizar a nossa humanidade individual (Ash, 2016: 73), aumentando a soberania individual (Scalon, 2015: 3), contribuindo para tomadas de decisão informadas. Este desenvolvimento pessoal, ajudado pela capacidade de o indivíduo poder expressar-se livremente, torna-o num membro na sociedade com pleno poder das suas funções (Hudson, 2018). Por corolário, assume-se que a censura de opiniões e argumentos, apresentados na forma de discurso e/ou conteúdos, retira matéria a essas funções, diminuindo o desenvolvimento e a realização pessoal.

O exercício da liberdade de expressão, associado à liberdade de receber e divulgar informação, é também primordial para uma boa governação (Sunstein, 1995), sendo uma forma de controlo da governação e da função dos governantes. Dito de outra forma, é uma liberdade essencial de ser defendida para o funcionamento democrático de uma sociedade. James Madison, no documento *Virginia Resolution*, e numa resposta à passagem do *Ato dos Estrangeiros e da Sedição*⁴ de 1798 (US Congress, 1798), propunha que o poder de delimitar a liberdade de expressão devia “produzir um alarme universal, porque é direcionado contra o direito de livremente examinar pessoas públicas e medidas, e de comunicar livremente entre as pessoas” (Madison, 1798). Democracias liberais regem-se por princípios democráticos, por exemplo, o escrutínio da governação e o acesso à informação por parte dos cidadãos, a capacidade de tornar pública uma opinião, e de apresentar uma reclamação ou uma queixa. Isso permite a correção de injustiças e a criação de consensos. Essas são condições para a criação, e manutenção, de uma sociedade aberta à diversidade de ideias e argumentos, evitando um deslizar para a tirania, violência e barbárie. John Locke fazia um argumento similar, em 1689, aquando da publicação de *Uma Carta Relativa a Tolerância*⁵ onde apelava a uma abertura à diversidade religiosa de maneira a evitar conflitos (Locke, 1689), o que mais

⁴ *Alien and Sedition Act*, no original.

⁵ *A Letter Concerning Toleration*, no original.

tarde seria ecoado por Bollinger defendendo que a liberdade de expressão “testa a nossa capacidade de viver numa sociedade que é, necessariamente, definida por conflitos e controvérsias” (Bollinger, 2010: 48).

Na *Declaração Universal dos Direitos Humanos* existem os artigos 12.º e 19.º de particular interesse no que se refere à liberdade de expressão, produção e disseminação, e receção de discurso e/ou conteúdos. Estes dois artigos foram originalmente apresentados em 1948, no mesmo ano em que a IBM desenvolvia o *Selective Sequence Eletronic Calculator*, o primeiro computador eletromecânico, e um ano antes de serem testados, pela primeira vez, discos magnéticos capazes de armazenar informação. É razoável pensar que John Peters Humphrey, autor da primeira versão da *Declaração Universal* não antecipava a revolução digital que aconteceria em menos de 50 anos. Estes direitos sobreviveram, pelo menos até agora, ao nascimento, e à quase universalidade da Internet nas sociedades desenvolvidas. Isso não obsta, no entanto, que haja cada vez mais pedidos para que a declaração seja atualizada para os tempos modernos. Há quem queira criar uma nova declaração, ou proclamar a independência do mundo digital das regras e normas criadas no mundo real⁶. Este “admirável mundo novo”, no dizer de Aldous Huxley, trouxe ameaças e desafios que reforçam a necessidade destas liberdades continuarem a ser defendidas com renovado vigor. Afinal, estes direitos são o garante para se obterem outros direitos: de pensamento; de inovação; de interação social e cultural; de associação e reivindicação de justiça.

Com direitos e benefícios bem definidos no mundo real, existe cada vez mais a exigência, a nível de leis internacionais e particularmente de leis internacionais sobre direitos humanos, que se apliquem os mesmos direitos ao ciberespaço. Os artigos 12.º e 19.º da *Declaração Universal* têm objeto suficiente para fazerem uma ligação entre o exercício da liberdade de expressão, vista de uma forma tradicional, e o uso da mesma liberdade *online*. No artigo 12.º, defende-se que “[n]inguém sofrerá intromissões arbitrárias na sua vida privada, na sua família, no seu domicílio ou na sua correspondência, nem ataques à sua honra e reputação. Contra tais intromissões ou ataques toda a pessoa tem direito à proteção da lei” (Rodrigues & Carneiro, 2011: 24). Já no artigo 19.º pode ler-se que “[t]odo o indivíduo tem direito à liberdade de opinião e de expressão, o que implica o direito de não ser inquietado pelas suas opiniões e o de procurar, receber e difundir, sem consideração de fronteiras, informações e ideias por qualquer meio de expressão” (Rodrigues & Carneiro, 2011: 24).

⁶ Ver, como exemplos, as propostas por parte do grupo *Internet Freedom* (<https://www.eff.org/cyberspace-independence>) ou de John Perry Barlow (<https://declarationofinternetfreedom.org/>).

Também a *Convenção Europeia para os Direitos Humanos* apresenta no artigo 10.º, ponto 1, que: “[t]odos têm o direito da liberdade de expressão. Este direito inclui a liberdade de ter opiniões e de receber e divulgar informação, sem interferências por parte de autoridades públicas e independente de fronteiras”. No ponto 2 do mesmo artigo “[o] exercício de tais liberdades, uma vez que carregam consigo deveres e responsabilidades, pode estar sujeito a formalidades, condições, restrições e penalizações como aquelas prescritas por lei e que são necessárias para uma sociedade democrática, no interesse de segurança nacional, integridade territorial ou segurança pública para a prevenção de desordem ou de crime, para a proteção da saúde e das morais, para a proteção da reputação de outros, para prevenir a publicação de informação recebida confidencialmente, ou para manter a autoridade e imparcialidade do sistema judicial” (ECHR, 1950).

No âmbito deste trabalho, o *Pacto Internacional dos Direitos Civis e Políticos* é de especial importância. Aprovado em 16 de dezembro de 1966 pela Assembleia Geral das Nações Unidas, nos termos do seu artigo 49.º, entrou em vigor em 23 de março de 1976. Na procura da defesa do indivíduo no exercício dos deveres para com os concidadãos e comunidade, as Nações Unidas criaram um conjunto de “direitos de primeira geração”, que asseguram um acesso à justiça e à participação política (UNHRC, 1966). De particular importância no pacto é o parágrafo 3, alíneas a) e b) do artigo 19.º, “[o] exercício dos direitos providenciados no Parágrafo 2 deste artigo carrega consigo deveres especiais e responsabilidades. Poderá, por isso, ser sujeito a certas restrições, mas essas só podem ser providenciadas pela lei e sendo necessárias: a) pelo respeito ao direito da reputação dos outros; b) para a proteção da segurança nacional ou da ordem pública (*ordre public*), ou da saúde e morais públicas”, conforme testes rigorosos de necessidade e proporcionalidade (UNHRC, 2011: 6). Estas restrições têm sido a base para enquadramentos legais e institucionais no que se trata do exercício da liberdade de expressões e seus limites.

Porém, não existe, ainda, uma caracterização consensualmente aceite de como se deve compreender, e aceitar, estas liberdades no mundo digital. Alguns passos começaram a ser dados nessa direção. Em 2013, a Assembleia-Geral das Nações Unidas aceitou a *Resolução 68/167*, onde se afirma que os direitos humanos, incluindo o da liberdade de expressão, se aplicam *online* (UNHROHC, 2018). Esta iniciativa nasceu da necessidade de se mitigar o impacto de uma vigilância generalizada e da interceção de comunicações pessoais, da liberdade de expressão e de associação, e do funcionamento de uma sociedade civil livre e vibrante. Nessa resolução, a Assembleia-Geral chama todos os Estados-membro a respeitar a proteger o direito à privacidade em comunicações digitais, assim como a reverem procedimentos, práticas e legislação com interferência na perda, ou ganho, de liberdades.

Também nesse documento é importante salientar que remete para o *Pacto Internacional dos Direitos Civis e Políticos*, agora já ratificado por 167 Estados-membro, como exemplo para a criação de pontos de ligação entre os direitos no mundo real e no digital.

1.2 Como se tornou um direito

O direito à liberdade de expressão encontra-se codificado em documentos que servem como instrumentos de natureza legal, tanto a nível nacional como supranacional. Exemplos destes documentos que contêm provisões para a defesa desta liberdade são a *Declaração Universal de Direitos Humanos* (artigo 19.º), o *Pacto Internacional dos Direitos Civis e Políticos* (artigo 19.º), a *Convenção Internacional para a Eliminação de todas as Formas de Discriminação Racial* (artigo 5.º(d)(viii)), a *Convenção dos Direitos das Crianças* (artigo 10.º), a *Convenção Europeia dos Direitos Humanos* (artigo 10.º), a *Carta Africana dos Direitos Humanos e dos Povos* (artigo 9.º), a *Convenção Americana dos Direitos Humanos* (artigo 13.º) e a *Declaração dos Defensores dos Direitos Humanos* (artigo 6.º).

O *gold standard*, no que se trata da definição operacional para regras, normas e atos legislativos, está expresso no artigo 19.º da *Declaração Universal dos Direitos Humanos*, que afirma que todo os indivíduos têm direito à liberdade de opinião e de expressão, com também o direito de não serem inquietados pelas suas opiniões, e de procurar, receber e difundir, sem consideração de fronteiras, informações e ideias por qualquer meio de expressão (United Nations, 1948). Com um maior enfoque no que se trata de discurso político, o artigo 19.º do *Pacto Internacional dos Direitos Cívicos e Políticos* (UNHRC, 1966), é aquele mais vezes apresentado como o melhor equilíbrio entre a defesa da liberdade de expressão e quais os testes a que pode ser sujeita no caso de restrições. Particularmente, o artigo 19.º define, na mesma linha da *Declaração Universal* que ninguém pode ser discriminado por causa das suas opiniões, e que toda a pessoa tem direito à liberdade de expressão; este direito compreende a liberdade de procurar, receber e divulgar informações e ideias de toda a índole sem consideração de fronteiras, seja oralmente, por escrito, de forma impressa ou artística, ou por qualquer outro processo escolhido. Porém, é também neste artigo que estão definidos “deveres e responsabilidades especiais”, e que o exercício do direito “pode estar sujeito a certas restrições, expressamente previstas na lei, e que sejam necessárias para: a) assegurar o respeito pelos direitos e a reputação de outrem; b) a proteção da segurança nacional, a ordem pública ou a saúde ou a moral pública” (UNHRC, 1966). Apesar de tanto a *Declaração Universal* como o *Pacto Internacional* não serem documentos vinculativos, têm beneficiado de estatuto de lei internacional regulamentar.

Existe uma declaração das Nações Unidas que reforça a universalidade do direito à liberdade de expressão, recomendando aos Estados-membro o reforço da obrigação de órgãos sociais servirem como um veículo de informação, e de proporcionarem um espaço para a livre expressão na sociedade (UNHRC, 2014)⁷. Igualmente, os Estados-membro não devem impor restrições na liberdade de expressão, a não ser que estas restrições sejam resultado de aplicações de testes com limites elevados para a censura ou sonegação de discurso e/ou conteúdos, no âmbito da lei internacional. Este poder deve ser extremamente limitado a gerar restrições à produção de discurso e/ou conteúdos, mais, quando estes são interpretados à luz de tradições locais, de enquadramentos culturais e sistemas de valores particulares. Nesta latitude de defesa da liberdade de expressão inclui-se o discurso político, uma vez que este é intrínseco à democracia.

Historicamente, o caso *Schenck v. United States*, ouvido pelo Supremo Tribunal dos Estados Unidos em 1919, é um dos primeiros momentos onde é produzida uma consideração crucial sobre liberdade de expressão, e quais os seus limites, tendo como base a Constituição e leis americanas. Em 1917, o Congresso americano aprovou a medida conhecida como o *Ato de Espionagem* (US Congress, 1917). Este pacote legislativo visava proibir interferências pela sociedade civil ao processo de recrutamento de soldados para o exército americano. Charles Schenck, na altura o Secretário-geral do Partido Socialista Americano, liderava a resistência à entrada dos Estados Unidos na guerra. Para publicitar essa opinião, Schenck encomendou 15 000 panfletos a opor o recrutamento pelo exército para serem distribuídos pelas ruas. Schenck e seus associados foram presos por sedição, e cumpriram seis meses numa prisão federal. O recorrer da sentença chegou ao Supremo Tribunal, com o argumento que existira uma violação do direito de livre expressão de Schenck e seus correligionários. Numa decisão unânime, o Supremo Tribunal não daria razão aos queixosos. O acórdão escrito pelo Juiz Oliver Wendell Holmes, tornou-se num dos pareceres mais conhecidos, e discutidos, na apreciação do direito à liberdade de expressão. Numa analogia a uma situação de perigo, o Juiz Holmes escreveu que “[a] mais estrita proteção da liberdade de expressão não protegerá um homem de gritar fogo num teatro, causando pânico (...). A questão em todos os casos é se as palavras que são usadas em tais circunstâncias e são de tal natureza a criarem um perigo claro e eminente e que possam trazer males substantivos que o Congresso tenha um direito de prevenir” (USSC, 1919). Criou-se assim jurisprudência sobre a importância do contexto do discurso proferido. O acórdão acrescentaria que “em outros

⁷ Essa declaração teve contribuições do Representante da Liberdade da Imprensa da Organização para a Segurança e Cooperação na Europa, do Relator Especial para a Liberdade de Expressão da Organização dos Estados Americanos, e do Relator Especial da Liberdade de Expressão e Acesso à Informação da Comissão Africana nos Direitos Humanos e das Pessoas.

lugares e noutras alturas”, o conteúdo dos folhetos estaria protegido pela Primeira Emenda da Constituição. No entanto, em período de guerra, esse direito estava condicionado. O “perigo claro e iminente” passou a ser o teste para avaliar qual o dano que o discurso e/ou conteúdos pode causar, e criou escola que “cada caso é um caso” (Waimberg, 2015).

Um ano mais tarde, o próprio Juiz Wendell Holmes tentaria refinar este teste, usando o caso *Abrams v. United States*, igualmente com o enquadramento do *Ato de Espionagem* de 1917 (USSC, 1919a). Folhetos, desta vez lançados de janelas de edifícios em Nova Iorque, denunciavam o envio de tropas americanas para a Rússia. Noutros folhetos, impressos em yiddish, os esforços americanos para impedir a Revolução Russa eram criticados, assim como a produção de armas para serem usadas na Rússia. Noutro momento central para a definição do direito da liberdade de expressão o Juiz Holmes reavaliaria a questão do contexto associado à produção de discurso e conteúdos, e alterou, mesmo que subtilmente, o conceito apresentado em *Schenck v. United States*. Foi proposto então um limite mais restritivo para regulação por parte do Estado dependente da intenção da criação de um perigo claro e iminente, e qual a real produção desse perigo.

Já o Juiz Louis Brandeis, também do Supremo Tribunal dos Estados Unidos, no acórdão do caso *Whitney v. California*, sugeriu que não deve ser atribuído perigo à produção e disseminação de discurso e/ou conteúdos se houver ampla oportunidade de serem discutidos previamente, ou no momento. Esta consideração antecipou os conceitos da arena de discussão e do mercado de ideias. O caso *Whitney v. California* refere-se a Charlotte Whitney, uma das fundadoras do Partido dos Trabalhadores Comunistas. Numa convenção do Partido, Whitney apresentou uma moção onde apelava à ação política. No entanto, a moção final aprovada seria francamente mais radical do que a proposta por Charlotte. O Estado da Califórnia indiciaria a autora por crime de organização com objetivo de causar mudanças políticas e económicas com recurso ao uso ilegal de violência (USSC, 1927). O Supremo confirmaria a decisão aquando do recurso, argumentando que o Estado, no exercício do seu poder policial, tem o direito de punir aqueles que abusam do seu direito de expressão, se incitarem à prática de crimes, distúrbio da paz pública, ou colocarem em perigo os pilares da função governativa com ameaças de golpes de estado.

A opinião da minoria, da autoria do Juiz Louis Brandeis, argumentava que liberdade de expressão pode ser regulada ou limitada, mas precisa de ficar bem definido o perigo associado à mensagem e não apenas se é uma ideia impopular, ou que pode causar ofensa. Na apreciação do Juiz Brandeis, o “[m]edo de uma lesão séria, só por si, não pode justificar supressão da liberdade de expressão e de associação (...) para justificar a supressão da

liberdade de expressão devem existir bases razoáveis para recear que um mal sério resulte se a liberdade de expressão for praticada. Deve haver bases razoáveis para acreditar que o perigo é iminente. Deve existir bases razoáveis para acreditar que o mal a ser prevenido é sério” (USSC, 1927).

É também nesta decisão, e na opinião da maioria, que é tornada pública, pela primeira vez, a ideia de elevada importância expressa na frase que se tornou uma máxima para os defensores da liberdade de expressão e da troca aberta de argumentos e ideias: “[s]e existe um tempo para expor, através da discussão de falsidades e falácias, para evitar ao mal por um processo de educação, o remédio a ser aplicado é mais tipos de expressão, não silêncio forçado” (USSC, 1927a). Somente em 1969, ainda nos Estados Unidos, haveria outro acréscimo importante na definição de direito de livre expressão via o caso apresentado no Supremo Tribunal envolvendo o conceito de discurso de ódio. O caso *Brandenburg v. Ohio*, refere-se a Clarence Brandenburg, membro do Ku Klux Klan, que, numa reportagem televisiva sobre um encontro do Klan no Condado de Hamilton no Ohio, incorreu em discurso considerado como de ódio, referindo ser necessário realizar uma vingança contra negros (usando a expressão *niggers*), contra judeus e todos aqueles que os apoiam. Brandenburg seria condenado no Tribunal Distrital de Ohio, mas veria a sentença ser revertida no Supremo, com a justificação que o governo não pode punir a advocacia de uso de força, ou promoção de atividades ilegais, desde que estas sejam no abstrato. Dessa forma, ficou definido que a liberdade de expressão será diminuída se for usada para “direcionar, incentivar ou produzir atos ilegais iminentes, e provável de incitar ou produzir ações similares” (USSC, 1969).

A União Europeia, principalmente aquando da redação dos tratados que a regulam, assume a proteção de direitos humanos fundamentais como base para o entendimento entre povos. Por exemplo, o artigo 2.º da *Tratado da União Europeia* apresenta que “[a] União é fundada nos valores de respeito pela dignidade humana, liberdade, democracia, igualdade, estado de direito e respeito pelos direitos humanos” (European Union, 2012). No artigo 11.º da *Carta Fundamental dos Direitos da União Europeia*, está definido o direito à liberdade de expressão e de informação, seguindo a linha do artigo 19.º da *Declaração Universal de Direitos Humanos*. Esse artigo afirma igualmente que a liberdade de pluralismo na imprensa deve ser respeitada (European Union, 2000). Porém, e como é recorrente aquando de codificações desta natureza, a Carta também refere que estas liberdades estão sujeitas a restrições. Aquando dessa necessidade, tem força de regulação a *Convenção Europeia dos Direitos do Homem*, em particular o artigo 10.º, que contempla restrições ou sanções, previstas pela lei, contemplando necessidades de segurança nacional, integridade territorial ou segurança pública, defesa da ordem e prevenção do crime, proteção da saúde ou da moral,

proteção da honra e dos direitos de outrem, impedir a divulgação de informações confidenciais ou garantir a autoridade e imparcialidade do poder judicial (ECHR, 1950).

A convenção estabelecida pelo Tribunal Europeu dos Direitos Humanos contempla a apresentação de queixas por quem ache que os seus direitos foram violados por um dos Estados-membro e após todas as outras formas de resolução terem sido esgotadas. De acrescentar que um dos primeiros casos sobre liberdade de expressão na Europa aconteceu em 1960, no âmbito dos direitos atribuídos pelo artigo 10.º da Convenção. Raymond De Becker⁸ apresentou uma queixa contra o Governo do Reino da Bélgica depois de ter sido condenado à morte pelo Conselho de Guerra de Bruxelas, por ter colaborado com as autoridades sob ocupação Nazi durante a Segunda Grande Guerra (ECHR, 1962). De Becker foi considerado como culpado pela transformação de organizações a mando do inimigo, assim como da produção e promoção de propaganda contra a resistência, incentivando belgas a cometer o crime de lutar contra o seu próprio país (ECHR, 1962: 4). A pena acabaria por ser comutada para 17 anos de cadeia e a perda dos direitos de ensinar, trabalhar em jornais ou televisão e até de ser proprietário de um negócio.

Em 1979, no caso *Handyside v. Reino Unido*, o Tribunal Europeu não deu razão ao queixoso quando este argumentou que não havia razão legal para a sua liberdade de transmissão de conteúdos ser retirada. Em causa estava a posse, para venda, do livro *The Little Red Book For School Pupils* dos dinamarqueses Søren Hansen e Jesper Jensen, que era considerado como obsceno. Apesar da decisão, no acórdão da maioria encontra-se o importante contributo que “[a] liberdade de expressão (...) é aplicável não apenas quando a ‘informação’ ou as ‘ideias’ que são favoravelmente recebidas ou consideradas como inofensivas (...) mas também quando ofendem, chocam ou causam distúrbio para o Estado ou qualquer setor da população” (ECHR, 1976).

1.3 Um direito *online*

Com o advento da era digital, a utilização da Internet e de plataformas digitais de redes sociais, tornou-se evidente a necessidade de extrapolar as definições de liberdade de expressão e da produção de discurso e/ou conteúdos do mundo real para o virtual. Principalmente quando estas foram concebidas para sociedades com menor acesso a opiniões e argumentos, ou capacidade os produzir. Ao se definir como as liberdades se adaptam ao mundo *online*, é necessário compreender como se aplicam os limites. As Nações Unidas, na sua trigésima-segunda sessão plenária, produziram o *Relatório do Conselho dos*

⁸ De Becker era jornalista, escritor e editor do jornal *Le Soir* na Bélgica.

Direitos Humanos, que inclui pontos de partida para a “promoção, proteção e usufruto dos direitos humanos na Internet”. Algumas destas prescrições servem como um guia para a construção de instrumentos legais, normas e regras, que se adaptem a uma nova realidade de ausência de “guardiões de conteúdos”, da criação do “jornalismo unipessoal”, e da “editorialização pessoal de conteúdos” (UNHRC, 2016).

No documento são também tecidas considerações sobre a facilidade e rapidez de acesso a discurso e conteúdos na Internet. O Conselho para os Direitos Humanos das Nações Unidas determina que o direito da liberdade de expressão na Internet é um assunto de crescente importância, à medida que o desenvolvimento tecnológico permite a indivíduos, a nível global, acederem a estas novas formas de receber e difundir informação, assim como a usarem tecnologias de comunicação (UNHRC, 2016: 2). Esta determinação é enquadrada por recomendações para um trabalho comum na criação de medidas para impedir, ou condicionar, o acesso a informação *online*; diminuir a liberdade de expressão na Internet; manipular ou controlar o discurso e/ou conteúdos; invasões de privacidade e outras violações de leis de direitos humanos internacionais. A necessidade de se criarem regras e normas para o exercício do direito à liberdade de expressão prende-se também com o conceito da universalidade desse direito. O brilhantismo de desenvolver direitos humanos, em 1948, “sem consideração de fronteiras” ganha uma nova dimensão quando não existem barreiras físicas para partilha de informação, e com isso haver custos menores associados à propagação de discurso e/ou conteúdos com ideias e argumentos. É necessário fazer a ressalva que barreiras físicas podem ser compensadas por “barreiras virtuais”, como é caso das *Great Firewall* na República Popular da China (Economy, 2018) ou Federação Russa (Roth, 2019).

Previamente, situações de tensão sobre o uso da liberdade de expressão podiam ser resolvidas localmente. Com a Internet é necessário que os direitos adquiridos possam ser exercidos por todos. Tal como no mundo real, o caminho a percorrer para a manutenção de direitos na Internet irá envolver a sociedade civil, as organizações governamentais e não-governamentais e as empresas digitais, de maneira a que todos os *stakeholders* se encontrem incluídos neste processo.

Capítulo 2. Condicionalismos à liberdade de expressão

Historicamente, sempre houve aqueles que se apresentam como absolutistas da livre expressão. Exemplos são os Juízes Hugo L. Black e William Douglas do Supremo Tribunal dos Estados Unidos (O’Neil, 2019), ou Alexander Meiklejohn, um reformador educativo, filósofo e Presidente do Amherst College (Meiklejohn, 1961). Contemporaneamente, colonistas como Caleb McClain ou Tim Black também defendem que não deve haver limites para a liberdade de expressão, mesmo quando confrontados com os resultados negativos do seu exercício abusivo. Para Cass Sunstein, a defesa de um absolutismo “puro”, com imunidade completa a qualquer teste aplicado a qualquer forma de discurso e/ou conteúdo, é algo onde “ninguém realmente pensa assim” (Sunstein, 1995: 121) ou que é “no máximo uma teologia, e não tem, realmente, nenhum seguidor” (Sunstein, 1995: 122). Em contrapartida, Caleb McCain argumenta que “[a] pesar de causar satisfação podermos calar um oponente, continua a ser a mais imatura das formas de discurso social. A verdadeira força moral a política vem de confiarmos que o nosso lado está correto e que será vindicado pela história” (McCain, 2017). Já Tim Black argumenta que “[n]ão é que alguém como Thomas Jefferson, ou Tom Paine, ou Voltaire, para não mencionar muitos outros que respiravam o ar radical da altura, não estivessem conscientes que havia discurso malicioso, ou abusivo, ou simplesmente idiótico. É que naquela altura, liberdade e liberdade de expressão eram vistos como aspetos positivos, como algo que beneficiava a humanidade, um princípio para prevenir que um governo resvalasse para a tirania. Que permitia os ‘melhores anjos da nossa natureza’ prosperarem, e que ajudavam à procura da verdade” (Black, 2014).

Liberdade de expressão não deve conter, à partida, prescrições para o seu controlo. Afinal, quando prescrições começam a ser aplicadas, perde-se o significado primordial da palavra “livre” (Starr, 2019). É necessário aceitar a responsabilidade associada ao exercício deste direito. Qual o valor do discurso e/ou os conteúdos, e qual o efeito que produzem no resto da sociedade, cria um contínuo entre os absolutistas das liberdades até aos que pedem a imposição de regulações aquando da verificação da primeira ofensa, passando pelos que acreditam em soluções de compromisso. Os defensores da perspetiva absolutista acreditam que o ponto de partida é qual o “dano” que o discurso e/ou o conteúdo causa, e não o seu valor. Dito de outra forma, existem consequências negativas que sejam suficientes para haver condicionalismos?

Cass Sunstein apresenta uma classificação de conteúdos, e de qual o seu valor, para servir de guia para a aplicação de testes à liberdade de expressão. Por exemplo, restrições podem ser propostas quando o discurso e/ou conteúdos são de “baixo valor”, significando que

não trazem uma mais-valia para a arena de discussão. Um exemplo é quando é produzida uma mensagem com propósito de ameaçar outra pessoa, ou quando causa a sua difamação. Outras restrições seguem outro tipo de exigência: produção de informação falsa; conselhos médicos ou jurídicos não fundamentados; produção e disseminação de mensagens políticas quando se faz parte de forças militares ou de segurança. Quando os conteúdos são “neutros”, por exemplo no mercado de ideias de natureza política, estes devem ser avaliados respeitando o enquadramento onde são produzidos, como é o caso de saber quem é responsável pelo discurso e/ou conteúdo, e quais as formas disponíveis para a pessoa o fazer (Sunstein, 1995: 12-13). Sunstein coloca o repto de saber qual o tipo de proteção dada a discurso e/ou de conteúdos considerados como de “valor inferior” (Sunstein, 1995: 123). Para o autor, deve ser apreciado qual o valor da mensagem, e se houver uma necessidade de haver um teste aplicado, ou uma regulação passada, é necessário existir uma razão suficiente para isso acontecer. Uma razão proposta é quando o discurso de baixo valor, que não de teor político, é prejudicial, como é o caso de informação falsa ou publicidade enganosa, uma vez que a este tipo de informação e/ou conteúdos está associada a eventualidade de causar dano.

Igualmente, a autonomia de quem emite, e de quem recebe, discurso e/ou conteúdos, é outro fator importante (Sunstein, 1995: 139). A eventualidade de organizações, sejam governamentais, supragovernamentais ou privadas, retirarem a autonomia de serem expressas opiniões, ou de se ouvir um argumento, é resvalar para um paternalismo contrário ao livre exercício de liberdades adquiridas. É óbvio que certos condicionalismos são aceitáveis à partida. É compreensível a tendência para um controlo do discurso e/ou conteúdos para o “bem” da sociedade. Alguns dos casos são ameaças de violência, suborno, assédio sexual, *bullying*, formas de expressão que podem ser descritas como de baixo valor. Nessa situação, é defensável determinar que a autonomia de se receber estas mensagens é secundária relativamente à necessidade de proteção. Este tipo de defesa choca com a perspetiva absolutista, onde o emissor deve ter a possibilidade de receber qualquer discurso e/ou conteúdos, desde que munido de mecanismos legais de proteção. Igualmente, para os absolutistas, deve-se poder aceder livremente à arena de discussão e confrontar discurso e/ou conteúdos de baixo valor.

Como é inevitável, há variáveis independentes que tornam a inevitabilidade deste argumento insustentável: nem toda a gente tem a possibilidade, engenho, destreza ou poder de argumentação. Ou tempo disponível para estudar opiniões contrárias, ver quais as falhas que essas argumentações contêm e construir uma refutação capaz. Defender a posição de absolutismo presume que todos os intervenientes têm instrumentos e capacidades iguais. Porém, a insustentabilidade da posição alternativa também é evidente. Assumir a posição

paternalista que Sunstein propõe como discurso e/ou conteúdos de baixo valor, é postular que existe uma qualquer percentagem de indivíduos que têm de ser tratados com “cuidados especiais” por não estarem preparados, ou serem incapazes, de uma resposta robusta, a mensagens, opiniões e argumentos.

2.1 Como operar uma escala deslizando

Se existe a necessidade de testes e mecanismos de regulação sobre o valor e classificação do discurso e/ou conteúdos, quais os critérios? Desde Thomas Paine a Timothy Ash, passando por John Stuart Mill e Cass Sunstein, que essa pergunta inquieta os defensores da liberdade de expressão, conscientes que esta também é alvo de limites. Existe uma segunda questão: até onde podem ir estas proibições ou condicionalismos? Zechariah Chafee, em 1946, argumentou ser necessário o mínimo possível de regulação, via aplicação de leis, para quem deseja expressar-se livremente. Isso acarreta que, quem beneficia da liberdade não queira explorar limites. Ao fazê-lo, criará uma tensão indesejada onde a lei se terá de expandir de modo a responder a cada vez mais infrações (Chafee, 1946). Esta competição entre exercício de liberdade e necessidade de regulação introduz o conceito de “escala deslizando”. Desenvolvido por Frederick Schauer em 1985, este princípio alerta para a eventualidade de alterações de direitos adquiridos, que sendo aceitáveis no presente (casos instantâneos) podem tornar-se intoleráveis no futuro (tornando-se casos perigosos).

O princípio da escala deslizando defende que um caso instantâneo, por pequenas que sejam as alterações traga no presente, pode ter consequências drásticas no usufruto de direitos no futuro (Schauer, 1985). Na extensão, instintiva, deste cenário, um dos receios sobre condicionalismos, mesmo que pequenos, à liberdade de expressão, é a criação de censura, primeiro específica, depois generalizada, e em última análise, completa. Por outro lado, falta de controlo poderá inclinar a escala deslizando na direção oposta, para um “estado natural”, como descrito por Thomas Hobbes ou John Locke com discurso de baixo valor a causar dano por causa da ausência de restrições. A forma de lidar com o princípio da escala deslizando é ser-se preciso na regulação do discurso e nas condições para a produção de conteúdos, para se encontrar o melhor equilíbrio entre exercício de liberdades e convivência salutar em sociedade, ou como apresentado na *Convenção Europeia dos Direitos Humanos*, condições “necessárias numa sociedade democrática” (ECHR, 1950).

Quando expressos, pela primeira vez, os direitos fundamentais, procurou-se uma codificação sensata, de forma a não desequilibrar a escala deslizando. O artigo 19.º do *Pacto Internacional de Direitos Cívicos e Políticos*, mencionado anteriormente (UNHRC, 1966), é um

desses exemplos. Porém, no pacto são também definidas regras para usufruto das liberdades. Aliás, a primeira frase do número 3 é reveladora disso mesmo, “[o] exercício dos direitos providenciados pelo parágrafo 2 deste artigo carrega consigo deveres e responsabilidades especiais. E pode assim ser sujeito a certas restrições, mas que podem ser unicamente providenciadas pela lei e quando é necessário”. A Comissão dos Direitos Humanos das Nações Unidas interpreta a definição de liberdade de expressão de maneira a incluir discurso e/ou conteúdos que sejam considerados, por exemplo, “muito ofensivos” (UNHRC, 2011: 3). Este teste à liberdade de expressão remete para o parágrafo 3 do artigo 19.º do pacto, onde se torna legítimo impor limitações se o discurso e/ou conteúdo for avaliado à luz da proteção dos direitos e reputação de outros, e para a proteção da ordem pública, saúde pública e moral pública. Estas limitações devem ser no âmbito da aplicação das leis, respondendo a regulamentos que sejam claros e precisos num “princípio da certeza legal” (UNESCO, 2016: 57).

No que se trata de conceitos mais abrangentes, tal como ordem pública, o artigo 19.º permite restrições a atividades ilegais que criem dano e que sejam uma ameaça à ordem pública, desde que seja aplicado o princípio da necessidade e da proporcionalidade. Um exemplo deste tipo de proibição é dado pela Comissão de Direitos Humanos das Nações Unidas no que se trata de “elogiar”, “glorificar” ou “justificar” terrorismo (UNHRC, 2011: 11). Já o conceito de moral pública é mais fluído, uma vez que envolve uma definição alargada, incluindo conceitos como tradições sociais, religiosas e filosóficas. Naturalmente, isso causa que apreciações de questões de moralidade sejam díspares. A Comissão de Direitos Humanos das Nações Unidas define que a moral deve ser baseada em princípios que não sejam exclusivos de apenas uma das tradições, de uma única doutrina, ou de um sistema, mesmo que seja consensualmente aceite nesses. Condicionismos à liberdade de expressão desta natureza devem ser entendidos à luz da universalidade de direitos humanos e de princípios de não-discriminação (UNHRC, 2011: 8).

Outra forma de regular o exercício da liberdade de expressão é a ameaça de sanções. Tradicionalmente, a forma mais comum para esse tipo de coerção é a intervenção do Estado, que historicamente se traduz em aplicação de multas, ou, em situações extremas, encarceramento. Porém, existe um outro sistema de condicionalismo, que não sendo coercivo, pode ser igualmente, ou mais, opressivo. John Stuart Mill, na sua obra de referência, *On Liberty*, avisa que a vontade do povo, ou daqueles que são aceites como seus representantes, pode oprimir quem não partilha ideias generalizadas ou consensualmente aceites (Mill, 1859: 8). Mill avisa que, quando a sociedade é tirânica, esse absolutismo e autoritarismo não é só produzida por dirigentes, forças de autoridade e funcionários públicos,

mas também por indivíduos que acreditam ter um mandato para tais níveis de controle: “[a] tirania social é mais formidável que muitos outros tipos de opressão política, desde que, habitualmente não aplicando usualmente tais penalidades extremas, providencia menos meios de escape, penetra muito mais profundamente nos detalhes da vida, e escraviza a alma” (Mill, 1859: 9).

A proteção contra este tipo de condicionalismo passa não só por uma proteção contra intervenções injustificadas de governos, mas também contra a tirania da opinião prevalente. Haverá, naturalmente, um limite para uma interferência legítima da opinião coletiva sobre a individual. No entanto, encontrar esse limite, e mantê-lo, é “tão indispensável à boa condição dos assuntos humanos como é a proteção contra o despotismo público” (Mill, 1859: 9). A outra consequência de sentimentos prevalentes é o receio de emitir-se uma opinião que é alvo de uma ridicularização generalizada, ou da acusação de desaforo moral. No entanto, estas consequências dependem da capacidade de quem emite essa opinião ser capaz de a defender, de suportar os abusos daqueles que não concordam com a mesma, e de defender a sua liberdade quando lhe querem retirar esse direito.

Com o desenvolvimento da era digital, este sentimento prevalente sugerido por Mill, de uma possível autocensura por causa da tirania da maioria, ganhou uma dimensão que seria impossível para o filósofo imaginar. Da “superioridade moral”, à “cultura do cancelamento”, à “sinalização de virtude” e às “microagressões”, os avisos que Mill fazia na segunda metade do século XIX são cada vez mais prementes para compreender as tentativas de condicionalismos por “exércitos” de indivíduos, que “patrulham” a Internet à procura de “infratores”, substituindo-se assim aos legisladores e autoridades.

2.2 O princípio da ofensa

Joel Feinberg é um dos autores mais importantes para entender o princípio da ofensa, caracterizando-o como discurso que causa “[a]borrecimento passageiro, desapontamento, nojo, embaraço, e outras condições desagradáveis como medo, ansiedade e outras menores (inofensivas) dores e esforços, que não são, neles próprios, necessariamente perigosos. Consequentemente, independente da forma como o princípio do dano é mediado, não certificará como legítimas quaisquer interferências com a liberdade de alguns cidadãos que são feitas com o único propósito de prevenir estados desagradáveis em outros” (Feinberg, 1988: 1). Feinberg propõe que ofensa e dano tenham definições normativas, gerais e específicas: ofensa relaciona-se com estados de desconforto mental, enquanto dano acontece numa situação onde existe uma ação ilegal. Especificamente, uma

pessoa sente-se ofendida quando se encontra num estado de desconforto mental, e atribui esse estado à ação prejudicial por outro, com ressentimento por se sentir nesse estado.

O princípio da ofensa previne que pessoas possam ofender outras, quando de uma forma ilegal ou prejudicial. Isto é contrário, por exemplo, à ofensa que se percebe quando se usa linguagem ordinária (Feinberg, 1988: 2). Feinberg acrescenta que, para se aplicar o princípio, tem de existir algo de errado, não simplesmente que a pessoa que se diz ofendida ache que algo está errado (Feinberg, 1988: 3). Sentir-se ofensa não é uma condição de dano. Moralmente, sentir-se ofendido é algo relativamente trivial. O princípio da ofensa surge então como uma forma de expandir conceitos relacionados com princípios reguladores de discurso que não de âmbito legal. O princípio de dano é específico e cria uma fasquia elevada para o controlo, porém há outros tipos de liberdade de expressão consideradas lesivas que não são ilegais. O autor propõe então um sistema de controlo: “há sempre uma boa razão no suporte de uma proibição criminal proposta que provavelmente seja uma maneira efetiva de prevenir ofensas sérias (...) este princípio afirma, efetivamente, que a prevenção de condutas ofensivas é apropriadamente um assunto do Estado” (Feinberg, 1988: 1).

Porém, o que é considerado como ofensivo, depende em muito de qual a sensibilidade da pessoa, de qual o enquadramento social e cultural, da propensão para se ser mais ou menos tolerante com certos temas. O que pode ser considerado como ofensivo para uns é divertido para outros. A representação de um assunto político, religioso, ou social, na forma de um *cartoon*, por exemplo, pode facilmente causar duas reações opostas: por um lado uma ofensa, de tal forma sentida que leva à violência, ou por outro lado uma diversão passageira e inconsequente. Isso pode se aplicar a um *cartoon* com o Papa com um preservativo no nariz, ou o Profeta Maomé com uma bomba como turbante. Feinberg sugere variáveis de interesse para os limites como: qual a extensão; duração; valor; a facilidade com que se pode evitar o discurso; quais as motivações de quem profere o discurso; qual o número de pessoas que se dizem sentir ofendidas; qual a intensidade da ofensa e o interesse da comunidade no seu controlo. No entanto, estas variáveis não são isentas de subjetividade ou problemas de interpretação. É difícil conhecer quais as reais motivações de quem expressa uma opinião ou faz uma apologia, ou qual o número “de corte” para se atingir um limite ou qual o grau de intensidade. Existe assim um trabalho contínuo para se avaliar, corretamente, quais as razões para se sentir ofensa.

Para Feinberg, “[p]ode ser legítimo, de uma perspetiva moral, que o Estado, através de leis criminais, proíba certas ações que não causam ofensa, com a justificação que essas ações podem constituir males de outra natureza” (Feinberg, 1988: 3). Neste caso, punições

devem ser exercidas quando se esgotam as soluções de bom senso, tais como injunções legais, avisos de contenção por parte das autoridades, e/ou desaprovação social. Se isso não for suficiente são então necessários outros mecanismos para evitar comportamentos de persistência ou reincidência. Esses casos aplicam-se particularmente quando o discurso ofensivo se torna incômodo, o alvo da ofensa se sente “encurralado”, e o esforço para evitar o incômodo é elevado e descompensado (Feinberg, 1988: 5). Esta é uma componente crucial: é evitável a exposição a discurso e/ou conteúdos ofensivo? Para melhor entender este dilema, um dos exemplos recorrente na literatura é o caso da iniciativa do Partido Nacional Socialista da América de realizar uma marcha em Skokie, Illinois, uma área residencial nos arredores de Chicago, principalmente composta por judeus, alguns sobreviventes do Holocausto. Em 1977, Frank Collin, o então líder do partido, anunciou que iria realizar uma marcha com os simpatizantes da causa nacional-socialista nas ruas de Skokie. Os residentes argumentaram que a que simbologia Nazi seria o equivalente a um apelo público para o assassinato de judeus, e conseguiram uma providência cautelar de forma a proibir a marcha.

Em janeiro de 1978, o Supremo Tribunal de Illinois decidiu a favor de Frank Collin, recorrendo à regra da neutralidade dos conteúdos, que postula que o discurso político não deve ser censurado por causa da mensagem que transmite, mesmo que essa pareça ser verbalmente abusiva (SCI, 1978). Assim, a liberdade de expressão deve ser restringida quando causa uma interferência física, é imposta a uma audiência refém, ou quando apela diretamente a atos imediatos e causadores de danos. No caso de Skokie, o Tribunal não deu razão ao argumento dos residentes, afirmando que simbologia é discurso, com a missão de transmitir uma ideia. Tal deve ser protegido, mesmo quando a ideia é ofensiva. Isso abre caminho para o precedente onde liberdade de expressão não deve ser suprimida mesmo que quem esteja exposto ao discurso e/ou conteúdo não seja tolerante com às ideias transmitidas.

A ideia que, ser-se exposto a discurso e/ou conteúdos que podem ser interpretados, subjetivamente, como incomodativos, pode causar dano mental, é um ponto de partida perigoso. Jeremy Waldon propõem que “[a] lei deve proteger a dignidade das pessoas, mas não as proteger de ofensas. A dignidade é um aspeto objetivo da posição de uma pessoa na sociedade. Por sua vez, a ofensa relaciona-se com questões subjetivas: mágoa, choque ou ira, que são reações subjetivas” (Waldron, 2012: 105-108). Confiar em mecanismos legais para proteger pessoas que se sentem incomodadas é entrar em exigências de regulação que rapidamente se tornam in comportáveis no princípio da escala deslizante. Qualquer teste que leve a este processo terá de ter limites muito elevados. Terá de conter discurso e/ou conteúdos muito para além do que é inconveniente, irritante ou aborrecido, na linha de Feinberg. O desconforto deve traduzir-se em dano mental, que não especulativo, e, quando

possível, deve ser suportado por evidência, por exemplo de natureza médica. A área da psicologia pode ser de ajuda, criando quadros clínicos para atestar qual o efeito nefasto da exposição à liberdade de expressão. Em Skokie, equipas de psicólogos deram apoio aos sobreviventes do Holocausto que se descreveram como mentalmente traumatizados pela iminência da marcha Nazi (Downs, 1985: 640).

Timothy Ash sugere existir três formas concorrentes de encontrar um equilíbrio entre o exercício da expressão, e a resolução do princípio da ofensa. Instituições governamentais definem leis que impõem limites ao que é considerado ofensivo, dentro das suas competências como legisladores na procura de um bem-comum. Ash sugere uma grelha de avaliação com três eixos: qual o contexto onde o discurso/conteúdo é divulgado; qual a justificação que tem; e se foi alvo de uma contenção justificável. Estes três eixos “rodam” dentro de uma esfera mais alargada, onde se inclui o tempo, a forma, o local, a audiência para que é destinado, qual o meio utilizado, e os enquadramentos históricos e culturais (Ash, 2016: 94). Feinberg também sugere três critérios que podem ser aplicados para conferir este equilíbrio: qual a intensidade, e duração, do desconforto mental, e o quanto esse desconforto pode ser antecipado aquando da exposição de discurso e/ou conteúdos de teor ofensivo. Dito de outra forma, qual a capacidade do queixoso de evitar o discurso e/ou conteúdo ofensivos. Quando o queixoso assumiu, de uma forma voluntária, o risco de se sentir ofendido ao se expor ao discurso e/ou conteúdo, nestes casos, não apreciação do autor, aplica-se a máxima que *volenti non fit injuria*⁹ (Feinberg, 1988: 27).

Outro aspeto importante na aplicação do princípio da ofensa é saber qual é a extensão da mesma. Discurso e/ou conteúdos que possam ser considerados como “extremamente ofensivos” por um grupo de pessoas, podem não causar nenhuma reação em grupos muito maiores. Como referido anteriormente, um *cartoon* com o Profeta Maomé, ou com o Papa, ou com um Rabi pode causar desconforto para um grupo alargado de pessoas, mas apesar de esse ser um número avolumado, pode ser residual quando comparado com um universo de pessoas que não sentiram qualquer ofensa. O mesmo acontece com discurso político. Um comentário sarcástico a um partido, ou jocoso a um político, pode causar aos simpatizantes desse partido, ou do político, o sentimento de que foi ultrapassado o limite da ofensa. No entanto, para todas os outros não adeptos da ideologia ou da causa, nem sequer ficou perto disso. Até mesmo o exercício de livre expressão, através da representação pública de símbolos de regimes totalitários e genocidas, não significa, necessariamente, que todas as pessoas que sejam expostas a esses símbolos se sintam incomodadas. O caso de Skokie

⁹ *Vontade não é um insulto*, em Latim

continua a ser um dos mais modelares nesse particular. Qual é então a extensão de uma ação criminal para um discurso e/ou conteúdo que seja ofensivo, mas que não represente qualquer perigo, ou que não cause risco de dano? Na opinião de Feinberg, depende da força do efeito que produz (Feinberg, 1988: 25).

Ao contrário de outro tipo de problemas, ofensas e danos são produzidos a seres humanos. Como tal, existe sempre a possibilidade de haver alguém com queixas genuínas e que reclame que certas expressões sejam censuradas, ou pelo menos modificadas. Um outro teste de ofensa é proposto por John Stuart Mill. Na opinião do filósofo, sente-se ofensa quando a opinião contrária a que fomos expostos é notável e poderosa (Mill, 1859: 50). Similar à “força produzida” como sugerido por Feinberg, para Mill a moderação de linguagem pode ser o suficiente para se evitar ofensa desnecessária, e daí resultar uma apreciação, sem distrações, das ideias e argumentos transmitidos. Já uma ofensa intencional, mesmo quando associado a uma opinião popular, pode fazer com que a importância do argumento seja perdida. Mill defendia, no entanto, a necessidade de um controlo mais apertado quando as ofensas resvassem para a indecência.

Se o princípio da ofensa começar a ser balizado com limites muito restritivos, corre-se o risco da livre expressão ser sistematicamente colocada em causa, entrando-se numa escala deslizante que leva à censura, autocensura e tirania pela maioria. Isto é particularmente importante no que se trata de discurso e/ou conteúdos de natureza política. Nessa arena de discussão pode existir à partida, e é bom se existir, a intenção de causar desconforto, ou embaraço, ou até mesmo de ofensa aos outros participantes no debate político. Da forma como, recentemente, o discurso político tem sido envolto numa camada de politicamente correto, é preciso que exista uma alteração na cultura contemporânea, que se tem tornado insidiosa de ver ofensa em todo o lado e a acontecer com qualquer coisa. Teremos de aceitar que é necessário saber interagir com outros membros da sociedade usando limites mínimos para discurso ofensivo. Isso não invalida que seja desenvolvido um conjunto de normas para lidar com situações limite, que não ilegais (Ash, 2016: 93). O princípio da ofensa tem de ser visto como secundário para a liberdade individual, da contribuição para uma sociedade respeitadora da diferença e para um debate político robusto. Queremos viver numa sociedade onde qualquer cidadão está pronto a ofender-se “por tudo e por nada”? Que precedentes estamos a criar para futuras gerações, onde ofensa serve como uma forma de interromper uma troca de argumentos, encerrar o mercado de ideias, censurar opiniões? (Ash, 2016: 91). É necessário desenvolver-se um ecossistema onde haja mais capacidade de não-ofensa, e principalmente perder-se o “direito” de se estar “perpetuamente

ofendido”. Afinal, também há o “direito” de não nos ofendermos “tão” facilmente (Barron, 2005: 266).

2.3 Princípio do discurso de ódio

Uma sociedade democrática beneficia dos princípios de igualdade e dignidade, e ambos são diminuídos pela existência de discurso de ódio. Autores como Waldron até se interrogam se é útil para a sociedade haver um debate sobre eventuais benefícios que estejam associados à existência de discurso de ódio (Waldron, 2012). Com uma perspetiva diferente temos novamente John Stuart Mill que argumenta que opiniões, mesmo odiosas, devem ser permitidas, nem que seja para evitar que ideias não sejam acomodadas no “sono das opiniões decididas” (Mill, 1859: 41). Não só esta é uma preocupação legítima, como outras são facilmente dedutíveis se o discurso de ódio não for um assunto abordado frontalmente. O facto de não se reconhecer, ou discutir, este tipo de discurso e/ou conteúdos, não fará que quem tenha tal tipo de opiniões, ou queira fazer tal tipo de argumentos, o deixe de fazer. Pode-se inclusive argumentar que é melhor para a sociedade que se saiba quem tem este tipo de crenças, numa lógica de “a luz do sol é o melhor desinfetante”, do que tentar reprimir tais opiniões e argumentos.

Conteúdo, a questão mais importante que se coloca é qual o ponto de transição entre discurso de ódio e o provocar de dano. É um ataque à dignidade pessoal uma ação danosa? É preciso passar-se a fronteira da argumentação verbal para a ameaça de ataques físicos? Em países como os Estados Unidos, que não tem leis específicas para discurso de ódio, não se observa que exista um maior dano causado por discurso do que em outros países com democracias liberais (SEP, 2019). David Bonin, na linha de Mill, acredita que não é necessária legislação específica para discurso de ódio. Para este autor, discurso que contenha ameaças já se encontra numa categoria para a qual existem proibições. Discurso de ódio não tem de ser, necessariamente, um discurso que implique uma ameaça física de causar dano (Bonin, 2011). Um grupo de homens de cor não se sentirá ameaçado por um discurso racista de uma mulher idosa branca, ou um grupo de homens brancos por parte de uma criança negra. Uma solução proposta é que o discurso de ódio não seja banido por ser odioso, mas por incitar à violência que resulta num ato físico. Mill defende o mesmo argumento quando apresenta o caso dos vendedores de milho e a iminência de violência por causa de discurso de ódio (Mill, 1859).

Também é aceitável que se assuma, como ponto de partida, que o discurso de ódio cause ofensa, e que o desconforto causado não pode ser facilmente ignorado (SEP, 2019).

Apesar da extensão do desconforto depender da possibilidade de evitar esse tipo de discurso, a intensidade da ofensa, e a quem se destina, quando é, por exemplo, dirigido a um número reduzido de pessoas e de uma forma precisa, pode justificar a limitação de discurso de ódio como sugerido por Feinberg. Fish concorda com a abordagem de Feinberg, sugerindo a regulação do discurso de ódio, uma vez que a não regulação aumenta o risco de ser causado dano, mesmo que isso implique perderem-se contribuições para o debate societal, ou até na escala deslizante se caminhar para um caso perigoso (Fish, 1994). Também para Mill, mesmo que o discurso de ódio não cause dano, pode se tornar um risco para o funcionamento de uma democracia robusta (Mill, 1978). Quando se trata de desenvolver enquadramentos legais, baseado também em qual a natureza do grupo alvo, discurso de ódio é aquele que incita ao dano, discriminação, hostilidade, violência.

O discurso de ódio procura desumanizar o outro, enquanto desenvolve mecanismos de identificação e solidariedade entre membros de grupos onde grassa intolerância. Há assim a criação de uma mentalidade tribal, com efeitos negativos a nível societal. Num ecossistema de troca de opiniões e argumentos, com uma arena de discussão robusta no que se trata do discurso político, tem que se assegurar que cada cidadão pode participar com as suas opiniões, queixas e sugestões, e sem estar sujeito a discurso de ódio, que lhe retira o direito a poder expressar opiniões sem ser inquietado por isso, como expresso no artigo 19.º da *Declaração Universal dos Direitos Humanos* (Rodrigues & Carneiro, 2011: 24). Também o artigo 20.º do *Pacto Internacional dos Direitos Cívicos e Políticos* cria condições para a aplicação de testes à liberdade de expressão no que se refere ao discurso de ódio: “[t]oda a apologia ao ódio nacional, racial ou religioso que constitua incitação à discriminação, à hostilidade ou à violência estará proibida por lei” (UNHRC, 1966). Discurso de ódio encontra-se assim num contínuo entre liberdade de expressão, direitos individuais e de grupos, e condições como dignidade e igualdade (UNESCO, 2015: 10).

Presentemente não existe uma única categoria de tipo de discurso e/ou conteúdos que sejam consensualmente considerados como discurso de ódio, não existindo assim uma base sólida para a limitação da liberdade de expressão por esse pressuposto (Wermiel, 2018). Expressões odiosas com ameaças de violência, ou que incitem a comportamentos criminosos, são alvo de legislação própria. Já discurso e/ou conteúdos ofensivos que criem ambientes hostis têm uma outra classificação, e são proibidos em casos extremos. O Teste de Brandenburg, inspirado no caso apresentado anteriormente de *Brandenburg v. Ohio*, é apontado como um modelo para avaliar quando o discurso pode levar a uma ação violenta. Relembrando a decisão judicial, o Supremo Tribunal Americano considerou que discurso de ódio é “legal e moralmente” permissível desde que, provavelmente, não cause dano iminente.

Visto de forma oposta, discurso de ódio pode ser punido criminalmente se “direcionar, incentivar ou produzir atos ilegais iminentes, e provável de incitar ou produzir ações similares” (USSC, 1969). Susan Benesch sugere cinco critérios onde se torna legítimo invocar o princípio de dano durante discurso de ódio: quando o orador é detentor de uma influência marcada numa audiência alargada; quando o tipo de discurso é claramente compreendido por essa audiência como um apelo a realizar atos de violência; quando essa audiência manifesta um estado de impressionabilidade e vulnerabilidade a tipos de discurso onde existe um claro discriminar a queixas e receios; quando o contexto sócio histórico onde o discurso é enquadrado é propício a existência de atos de violência; e quando o meio de disseminação é uma influência nele próprio.

Este último ponto, menos intuitivo, pode ser melhor entendido quando se pensa no meio de disseminação como sendo primário ou secundário, por exemplo um discurso versus um panfleto, ou se é um momento que, por exemplo, pode ser único e não reproduzível (Benesch, 2013). É importante notar que alguns dos critérios de Benesch se aplicam, de uma forma incontornável, ao discurso político: um orador que detém influência sobre uma audiência, quando existe impressionabilidade e vulnerabilidade por parte dessa audiência e quando o meio de disseminação é uma influência nele próprio. Políticos com dotes de oratória, alguns deles com uma influência extremamente nefasta na história do Século XX, como foram Adolf Hitler ou Benito Mussolini, usaram muito bem esses dotes para introduzir no discurso o apelo a atos de violência, usando habilmente contextos sócio históricos. Esse perigo continua a existir nos dias de hoje e é preciso estar vigilante para o detetar precocemente. O ataque ao Capitólio Americano no dia 6 de janeiro de 2021, será outro possível exemplo, ainda sujeito a uma interpretação histórica.

A falta de definição de conceitos chave tem causado alguns dos “pânicos morais” observados no século XXI, desde defensores do politicamente correto até aos absolutistas que acreditam que quaisquer sistemas políticos e/ou religiosos podem ser criticados sem limitações. Uma das posições é que ódio nem deve fazer parte da base de um sistema normativo regulador, ao contrário de racismo. Ódio é uma emoção, como tal não obedece a sistemas de controlo como é, por exemplo, o processo reflexivo que (deve) estar associado à expressão de ideias ou opiniões. O Conselho da Europa na União Europeia, por exemplo, no documento *Relatório Explicativo Associado ao Protocolo Adicional para a Convenção em Cibercrime*, caracteriza ódio como “rejeição e inimizade” (CoE, 2003: 3). Tal caracterização cria a possibilidade de múltiplas interpretações que, por sua vez, geram dificuldades na elaboração de uma definição consensual para um quadro legal. Um indivíduo, racional, educado e sofisticado pode ter uma rejeição e inimizade, quiçá até extrema, para com outras

peças ou comportamentos. É esse estado emocional suficiente para ser investigado criminalmente? A ideia de que estados emocionais possam ser alvo de controlo é biologicamente, e moralmente, uma solução impraticável e inadequada. O desejo de avaliar discurso de ódio baseado na presunção de um conhecimento integral das intenções de quem o profere não tem proteção legal. É necessário saber se as intenções se transformam em atos. George Orwell, no seminal livro 1984, apesar de ser uma obra de ficção, inspirada nas sociedades comunistas da altura, apresentou qual o crime a ser codificado para o tipo de preocupação acima explicada: o *crimeideia*¹⁰.

Rotular como discurso de ódio algo com que não se concorda, no lugar de se conseguir argumentar com uma opinião válida ou aceitar o mérito da outra posição, é desequilibrar a escala deslizando para uma eventual censura generalizada, e pior, levar à pré-censura de ideias, de argumentos e ultimamente de pessoas. O que se descreve recentemente como “cultura de cancelamento”¹¹ tem sido produto dessa lógica. No entanto, esforços têm sido realizados para haver melhores definições para a análise de discurso e/ou conteúdos, e seus efeitos, quando se trata de expressões de ódio. Por exemplo, conceitos como “discurso perigoso” ou “discurso de medo” têm sido incluídos em investigação académica, assim como posições de instituições supranacionais (UNESCO, 2015: 12). Discurso perigoso pode ser avaliado segundo os pontos propostos por Susan Benesch, já apresentados, sobre carácter e popularidade de quem produz o discurso, o estado emocional da audiência, conteúdos do discurso e apelos à ação, contexto e meios usados. Já a ideia de discurso de medo relaciona-se com linguagem que, quando utilizada repetidamente, cria uma “mentalidade de cerco” que pode levar à legitimação de atos de violência¹².

No que se trata de discurso de ódio em Portugal, o *Relatório da Comissão Europeia contra o Racismo e Intolerância*, de 2 de outubro de 2018, apresenta que “o discurso de ódio e o racismo estão presentes no discurso público e são dirigidos em particular à minoria cigana e aos negros” (ECRI, 2018: 19). Algumas queixas apresentadas visam membros da classe política, nomeadamente o professor Pedro Cosme Vieira, André Ventura, na altura candidato do PSD às eleições autárquicas, e até mesmo o Primeiro-ministro do XX do Governo Constitucional de Portugal, Pedro Passos Coelho, que foi acusado de racismo após ter referido que não queria “qualquer um a viver em Portugal”, quando criticava uma proposta-de-lei que garantiria autorização de residência a qualquer emigrante que conseguisse

¹⁰ “*Thoughtcrime*”, traduzido do original. Orwell definiu *thoughtcrime* como pensamentos que são contrários aos dogmas políticos do sistema político dominante, o Ingsoc, ou “Socialismo Inglês”. Contemporaneamente, a expressão é utilizada para descrever crenças contrárias a normas sociais dominantes.

¹¹ Sobre esta questão ver Greenspan (2020)

¹² Sobre “discurso perigoso” ver Benesch (2012), e sobre “discurso de medo” ver Buyse (2014).

demonstrar que tinha uma promessa de trabalho (ECRI, 2018: 19). Porém, outro tipo de interpretação defende que a razão para os comentários do então líder do Partido Social Democrata prendia-se mais com questões financeiras do que raciais (Tadeu, 2017). O único caso apresentado no relatório ECRI que relaciona o que é definido como discurso de ódio, discurso político e utilização de plataformas digitais de redes sociais foi o caso da ex-membro do Parlamento Europeu, Manuel dos Santos, do Partido Socialista, que em 2017 publicou um *tweet* com teor “anticigano”, numa mensagem dirigida, curiosamente, a Luísa Salgueiro, uma deputada da Assembleia da República do mesmo partido (ECRI, 2018: 19).

No relatório da ECRI é interessante constatar a afirmação de não existirem estatísticas sobre o número de incidentes de discurso de ódio na Internet, mas que “a sociedade civil informou a ECRI da existência de centenas de mensagens carregadas de ódio, particularmente em fóruns de extrema-direita” e que “[e]ste discurso de ódio visa (...) especialmente os ciganos, os negros, os muçulmanos e as pessoas LGBT” (ECRI, 2018: 19-20). O facto de ser a “sociedade civil” a informar destas mensagens “carregadas de ódio” levanta a questão delicada que é qual garantia que estas “centenas” de mensagens sejam realmente discurso de ódio, e não possam ser confundidas com o princípio da ofensa. Porém, é também legítimo crer que, muitas destas poderão ser eventualmente caracterizáveis como discurso de ódio. Levanta-se a questão interessante de saber o porquê do número de queixas formais com as autoridades e sistema judicial continuar a ser francamente reduzido. Será que os testes aplicados ao discurso e conteúdos não estão bem calibrados, ou as potenciais vítimas preferem não passar pelo processo de proteção legal?

2.4 Princípio do dano

Torna-se assim, depois da apresentação do princípio de ofensa e de discurso de ódio, sistematizar o que se entende por princípio de dano. Uma das opiniões de John Stuart Mill tornou-se um marco orientador para os defensores de uma liberdade absoluta no que se refere ao direito à expressão. Na opinião do filósofo, é necessário explorar os limites da liberdade e levar um argumento à sua conclusão lógica, não devendo ser esse processo interrompido quando se atinge uma qualquer condição de desaprovação social. Se a liberdade de expressão é condicionada, o preço que se paga é uma “pacificação intelectual que sacrifica a coragem moral da mente humana” (Mill, 1859: 31). Apesar de manter a força do argumento, Mill também concede que existem testes aplicáveis quando explorar limites incorre na produção de dano. Nessas alturas deve usar-se um princípio simples, e que se pode caracterizar por “o único propósito pelo qual o poder pode ser aplicado de uma forma leal sobre qualquer outro membro de uma comunidade civilizada, contra a sua vontade, é de

prevenir o dano aos outros” (Mill, 1859: 9). Desta forma, os limites para o controlo da liberdade de expressão devem ser, necessariamente, mínimos, uma vez que é insustentável defender que a maior parte do discurso cause dano, e que todos estejam expostos a ele.

Jacobson expande o teste proposto por Mill da não-sanção à produção de discurso e conteúdos simplesmente porque alguém pode sofrer dano. Para o autor, o poder da censura justifica-se apenas quando legítimo: quando o discurso causa uma clara e direta violação dos direitos dos outros (Jacobson, 2000). Exemplos são: a difamação; a chantagem; a publicidade enganosa ou que cause danos a menores; e a pornografia infantil. Assim, pode estabelecer-se o corolário que a liberdade de expressão deve ser restrita, quando o contexto onde esta liberdade é exercida é de molde a instigar atos desumanos. No entanto, precisa ser respeitado o princípio já apresentado que a advocacia de que, uma ideia ou argumento, mesmo que aberrantes, não é, necessariamente indutora de comportamentos de dano. Há a necessidade de ter-se presente a diferença entre instigação e advocacia. Instigação aplica-se na situação onde há um apelo a cometer uma ação, no momento ou no futuro imediato, do que, simplesmente, o pedido para acreditar numa ideia ou aceitar um argumento. Por exemplo, e como definido anteriormente, o discurso ofensivo torna-se um risco para o causar dano quando este está, intimamente, ligado a realização de uma ação nociva.

Seria novamente o Juiz Wendell Holmes a abrir um espaço de debate sobre a ligação entre intenção e ato, no que se trata do princípio de dano, com parecer no acórdão da decisão do Supremo Tribunal dos Estados Unidos no caso *Schenck v. United States*. Ao criar uma analogia entre o efeito do discurso de dano com a ação de alguém que grita “fogo” num teatro cheio de pessoas (USSC, 1919), criou-se o racional, repetidamente utilizado, para justificar que em certas circunstâncias, quando discurso é diretamente relacionado com ações que produzem resultados danosos, este deve ser restringido. Esta ideia é contrária à perspetiva Milliana onde discurso, para ser censurado, tem de causar dano físico e não ficar pelo abstrato de poder causar dano. Também é preciso ter em conta a que população se dirige o discurso e/ou conteúdos, e qual o efeito real que estes têm, que não apenas os emocionais. Importante também é qual a facilidade que as pessoas para que se dirige o discurso o possam evitar, como expresso no caso de *Skokie*. Sendo assim, se o discurso e/ou conteúdos leva a uma ação com consequências imediatas, e com a produção de dano, este deve ser proibido. Se o discurso e/ou conteúdos não causa essa resposta, apesar de poder causar, hipoteticamente, no futuro, não deve ser alvo de restrições. Existe então a tensão entre saber se o princípio de dano se aplica apenas quando existe a iminência da produção de dano, ou se os testes a aplicar são mais restritos e se avalia qual o risco de haver um aumento da probabilidade da produção de dano. Novamente Benesch apresenta uma forma de resolver esse dilema,

definindo que o discurso pode ser controlado quando aumenta o risco de existir atos violentos via técnicas retóricas com ameaças explícitas ou incitamento à violência (Benesch, Buerger, Glavinic, & Manion, 2018).

No entanto, violência física não é a única forma de causar dano. Existem outros atos que podem ser incluídos neste tipo de classificação. Exemplos conhecidos são o *bullying*, que de tal forma extremo pode levar ao suicídio, a devassa da vida privada, roubo de informação, campanhas de descridibilização, boicotes a indivíduos, empresas e instituições. Estes casos não são incluídos aquando da descrição clássica de proteções legais, como é o caso da difamação, proteção da segurança nacional, ordem pública, ou saúde ou moral públicas no artigo 19.º do *Pacto Internacional dos Direitos Civis e Políticos* (UNHRC, 1966), e da difamação via artigo 180.º do *Código Penal Português* (Estado Português, 2019). Por outro lado, e em particular com as leis de difamação, estas podem também ser uma forma de diminuição da livre expressão (Article XIX, 2009: 1). Igualmente as proteções contra possível difamação interagem com o direito de receber informação factual, opinativa ou conceptual. A utilização de leis de difamação, por parte de instituições com poder para o realizar, e com o intuito de proteger os seus interesses políticos e/ou económicos, pode silenciar vozes dissidentes (Article XIX, 2009: 3). No *Relatório do Relator Especial Frank La Rue*, para o Conselho dos Direitos Humanos das Nações Unidas (2011), sobre a promoção e proteção dos direitos de liberdade de opinião e de expressão, sugere-se descriminalizar o ato de difamação (La Rue, 2011: 11) uma vez que este não é causador de dano.

Diferenças na caracterização do princípio de dano continuará a ser uma área de trabalho para académicos, legisladores e instituições da sociedade civil, na procura de um quadro legal com o melhor equilíbrio entre usufruto de liberdades. Feinberg sugere, inclusive, que sejam elaborados um conjunto de testes à liberdade de expressão de maneira a categorizar quatro efeitos do exercício dessa liberdade no que se relaciona com o princípio de dano: dano a outros, ofensa a outros, dano ao próprio, infração sem dano (Feinberg, 1988). Estes testes devem incluir variáveis, agravantes ou atenuantes conforme o caso, dependendo do valor do discurso, do mérito do conteúdo, da maneira como é dispensado, das intenções e reais efeitos, na linha do que é proposto por Susan Benesch relativamente a discurso de ódio poder ser caracterizado como de dano (Benesch, 2013), sendo, novamente, acautelar para a subjetividade de tais variáveis.

2.5 Discurso político e liberdade de expressão

Discurso político espelha como o cidadão espera ser governado: expressando desgostos, elogios, preferências e expectativas. Essa é a forma de exercer o que pode ser considerado como “administração pública”, não no sentido comum do termo, mas sobre quais as formas do cidadão influenciar a decisão política. Ao contrário de interações pessoais com regras relativas, o discurso político precisa de regras, e normas, que estabelecem qual o debate entre participantes com diferentes pontos de vista, e daí resultar uma negociação com vista a um consenso. Este processo não só constrói um corpo deliberativo público, que deve ser aproveitado por governantes e legisladores, mas também, ao acontecer, legitima a governação, pois corresponde à vontade do povo, ou pelo menos daquele que participa. Uma democracia forte, como apresentada por Benjamin Barber é “definida por política no modo participativo: literalmente, a autogovernar pelos cidadãos no lugar do governo representativo em nome dos cidadãos” (Barber, 2003: 151). Esta conceção de democracia forte depende da participação permanente da comunidade na resolução de problemas que cria fins “neles próprios” de benefício público. Isto resulta de uma procura de soluções de uma forma conjunta. O sentido de comunidade cresce “resultado da participação, sendo que ao mesmo tempo, é a existência da comunidade que permite a participação: atividade cívica educa o indivíduo em como pensar publicamente como cidadão, assim como a cidadania informa a atividade cívica, com o requerido sentido de praça pública e de justiça” (Barber, 2003: 152). No entanto, para isto acontecer é necessário garantir que existe um igual acesso à informação, e que esta seja de valor. A criação de assimetrias no acesso à informação, ou qual o tipo de informação a que se acede, condiciona a participação, a atividade cívica, em última análise o exercício da cidadania.

O discurso e o debate político, se não condicionado, gera relações entre representantes e representados que resulta na resolução de problemas e reforço das boas decisões. Cass Sunstein defende que o discurso político está na base de um sistema deliberativo e que merece uma proteção mais extensa do que outros tipos de discursos e/ou conteúdos não garantidos constitucionalmente. Matérias com uma parte de falsidade ou de engano devem ser analisadas com testes mais liberais do que aquelas que resultam de outras arenas de discussão que não política. É irrealista pensar que o discurso político seja constituído por apenas verdades, ou produzido com completa honestidade. Sunstein defende que discurso e/ou conteúdos devem ser tratados como políticos, quando direcionados, e recebidos por uma comunidade, contendo informação técnica e num assunto de interesse que requer um processo de deliberação pública. Porém, fica a ressalva que, quando é reconhecido que o exercício do discurso e a produção de conteúdo são de natureza política, não é

necessário que todos entendam do que se trata, bastando que apenas um subgrupo o faça (Sunstein, 1995: 130). Isso acautelará questões de ideologia, onde certas opiniões, argumentos e mensagens poderão ser interpretadas como políticas, mas não isso poderá não acontecer, necessariamente, com outros participantes no debate público. Um exemplo recorrente são questões de moralidade, onde certas posições poderão ser defendidas como apolíticas para uma parte da sociedade, mas tal isenção não é acolhida por outras.

Sunstein partilha um bom exemplo de como a liberdade de expressão alimenta o discurso político, e vice-versa. Quando se lê no preâmbulo da Constituição Americana “[n]ós, o povo”¹³, isso retrata a substituição de um monarca por uma República, com um sistema de governação por discussão através de uma deliberação pública inclusiva (Sunstein, 1995: xvi). Isto confere uma maior importância ao valor do discurso, e ao mérito da argumentação do que às decisões por autoridade ou privilégio. O autor introduz a ideia de uma concepção Madisoniana da liberdade de expressão, referida dessa forma por ser inspirada no trabalho de James Madison. Nesta concepção reside a tradição da liberdade estar imbuída de uma luta constante para “entender as relações entre estas novas concepções de soberania, o compromisso para a igualdade política, a crença na deliberação democrática, e o sistema da livre expressão” (Sunstein, 1995: xvii). Sunstein refere-se às ideias de governação apresentadas por James Madison nos *Federalist Papers*¹⁴, em particular no *Federalist N° 10*: existe um fator principal numa República que é a delegação de governação em representantes, e onde um “refinamento e aumento das opiniões do público passa por um filtro de um corpo de cidadãos selecionados, cuja sua sabedoria, melhor definirá o verdadeiro interesse do país” (Madison, 1787). É necessário, nesta concepção, que exista uma participação efetiva dos membros da sociedade, mesmo que filtradas por representantes por si escolhidos, na procura de encontrar funções deliberativas justas. Porém, essa participação, e paridade num primeiro nível, só existe com a garantia do exercício da liberdade de expressão no discurso político na procura de uma sociedade mais justa e igualitária.

¹³ *We, the people*, no original.

¹⁴ O site na Internet da *Library of Congress* dos Estados Unidos disponibiliza os 85 *Papers*, sendo que alguns textos de Hamilton e Madison focam-se em questões de liberdade e governação.

Capítulo 3. A era digital e a liberdade de expressão

3.1 A evolução do discurso *online*. Da utopia à distopia

O início da era digital pode ser comparado com o aparecimento da imprensa escrita, pela forma como facilita o usufruto de liberdades. Particularmente de produzir, divulgar e receber informação, de superar barreiras físicas e temporais, do desenvolvimento de novos tipos de discursos e/ou conteúdos, e pela facilidade da participação na arena de discussão e no mercado de ideias. Porém, e como é de esperar que aconteça, quando um sistema de partilha de informação, opiniões e argumentos, se torna tão onnipresente e intrusivo como é a Internet, também há o risco que aconteça o contrário, ou seja, condicionalismos a essas liberdades, resultante da ação de instituições governamentais ou privadas (Ash, 2016: 20).

Em 1996 John Perry Barlow apresentava a *Declaração da Independência do Ciberespaço*, aproveitando a palavra ciberespaço para codificar um novo domínio de funcionamento, liberdade e responsabilidades. O termo ciberespaço foi utilizado pela primeira vez em 1982, numa novela com o título *Burning Chrome* da autoria de William Gibson, para retratar um novo espaço, com as suas regras e leis. De facto, e lendo parte da declaração de Barlow somos avisados que “[g]overnos do mundo industrializado (...) vocês não têm soberania onde nos encontramos” (Barlow, 1996). Esta tentativa de expressar, pela primeira vez, um conjunto de expectativas e de exigências deste novo espaço, é claramente inspirada na leitura da *Declaração de Independência dos Estados Unidos* e no apelo expresso nesse documento de uma independência perante uma força maior e opressiva. Porém, ao contrário da liberdade do povo para com um rei, o apelo à emancipação do cidadão digital dos governos terrestres parece ser muito mais difícil, se não impossível de alcançar. Como escreve Ash: “[a] Internet nunca foi, ou será, totalmente independente da influência de governos, empresas e outros poderes terrestres” (Ash, 2016: 22). É de esperar que existirão sempre condicionalismos aquando da utilização de plataformas digitais para o exercício da liberdade de expressão, tanto fora como dentro de limites legais.

Em vinte anos, Silicon Valley deixou de ser um lugar onde um conjunto de jovens, distribuídos por garagens e casas alugadas imbuídas no espírito de repúblicas estudantis, movimentados por ideias utópicas sobre inovação, conectividade e livre expressão, para outro onde esses pequenos projetos se tornam em conglomerados capitalistas, empresas multinacionais com regras e normas internas que se assemelham a aparelhos governativos. A outra diferença é ver a mudança das motivações iniciais para outras onde se deteta indiferença, se não convívência, com perdas de direitos e liberdades, de pouca proteção a

ameaças à privacidade e propriedade de informação, à diminuição das escolhas devido a monopólios comerciais, à abdicação para com exigências por parte de regimes autoritários. Nos Estados Unidos, o crescimento orgânico e autónomo de empresas privadas do setor digital foi particularmente facilitado, principalmente por causa da ausência de políticas de regulação, a partir do meio da década de 90, durante a gestão da Administração Clinton. Essa teria sido a altura para diretivas estruturais que condicionassem o modelo de desenvolvimento das plataformas, principalmente a nível de incentivo económico como fator dominante. No entanto, a escolha, consciente, dos legisladores americanos de intervenção estatal mínima na iniciativa privada e na regulação do mercado, aplicou-se ao desenvolvimento do mundo digital, sem se ter antecipado que os bens, serviços, transações de produtos, inovação e expansão não seriam a mesma forma de mercado como as mais tradicionais.

Desde 1996, as empresas que proporcionam plataformas digitais, e mais recentemente do que se denomina de “redes sociais”, têm beneficiado de uma isenção de responsabilidade sobre qual o tipo de discurso e/ou conteúdos que nelas são colocados via a Secção do Código Penal dos Estados Unidos com o título “Proteção do bloqueio privado e do controlo de material ofensivo”¹⁵. A principal medida, e a que tem sido mais acerrimamente defendida pelas empresas digitais, é a que declara que as plataformas, ou os seus utilizadores, não são considerados como editores de conteúdos de informação que seja providenciada por outras fontes. Ou seja, difundir discurso e/ou conteúdos nas plataformas digitais tem proteções legais que não se aplicam noutros canais de informação. Se por um lado essa imunidade permitiu a expansão de arenas de discussão, incluindo apolítica, tornando-as mais vibrante e participativa, também permitiu que fosse criado um ecossistema onde “vale (quase) tudo”, com consequências óbvias: o aumento da polarização, ferocidade, crueldade e até desumanidade. Na altura da elaboração deste trabalho, assiste-se a uma tentativa de sobre-correção desse início mais permissivo como resposta a casos de gravidade elevada sobre utilização indevida de dados pessoais, criação de ecossistemas insustentáveis de interações *online*, de ataques a processos democráticos, disseminação de notícias falsas e produção de polarização política, algo que será abordado mais à frente. Estamos assim a passar uma fase onde existem legítimas preocupações de que a democracia possa não sobreviver à Internet (Persily, 2017).

¹⁵ *Protection for private blocking and screening of offensive material*, no original. Para saber mais sobre esta medida ver USC (2018).

3.2 Quem censura o quê

John Milton, no discurso ao Parlamento da Inglaterra sobre licença de impressão e censura, levantou uma das questões cruciais no que se trata de regular o exercício da livre expressão: “[q]uem deve regular as conversas da nossa juventude, homens e mulheres juntos, como é costume neste país: quem poderá decidir o que deve ser dito, ou presumido, e não mais além? Finalmente, quem poderá proibir e separar todos os momentos ociosos, todas as más companhias? Essas coisas serão, e devem ser; mas como podem ser o menos danosas, o menos estimulantes, aqui consiste em a séria e sabedora governação pelo Estado” (Milton, 1644: 19). O filósofo aconselhou então que: “verdade e conhecimento não são mercadorias para serem monopolizadas e trocadas em bilhetes e estatutos e padrões” (Milton, 1644: 24), e fez o aviso aos Membros do Parlamento, “[a]creditem em mim, Lordes e Comuns, aqueles que aconselham para tal supressão [da impressão sem licença], é tão bom como uma licitação feita por nós próprios. Se for desejado de se conhecer a causa imediata de todos esta liberdade na escrita e na liberdade da expressão, não pode ser atribuída uma mais verdadeira que a nossa suave, e livre, e de governação humana: que é a liberdade” (Milton, 1644: 24). Milton fazia assim o argumento de qual a relação do Estado com a liberdade de expressão. Porém, como se aplicam estas ideias à era digital?

Quando o vídeo, de qualidade e gosto questionáveis, *A Inocência dos Muçulmanos* foi publicado no YouTube, o Google recusou pedidos para que o vídeo fosse banido por parte dos governos dos Estados Unidos, Índia, Indonésia, Malásia, Singapura, Jordânia, Arábia Saudita. Porém a mesma empresa bloqueou o acesso ao vídeo no Egito e na Líbia quando a Embaixada Americana em Cairo e o Consulado na Líbia em Bengasi se encontravam sob cerco. É inevitável pensar que este tipo de decisão por parte de uma empresa privada, arbitrária como parece, gera precedentes preocupantes, onde decisões tomadas por decreto podem decidir quem vê o quê, onde, e quando. Igualmente, este tipo de decisão levanta a questão incômoda de o quanto a ameaça de violência se torna uma “moeda negocial” para controlo da liberdade de expressão. Precedentes existem no mundo real suficientemente preocupantes, como foi o caso da publicação de *cartoons* com o Profeta Maomé pelo jornal Dinamarquês Jyllands-Posten. Ameaças de violência fizeram com que jornais e canais de televisão não mostrassem os *cartoons*, condição essencial para se formar uma opinião informada.

Outra consequência da Internet “omnipresente”, ou “a Internet das coisas”, é a facilidade, através de sistemas de vigilância em larga escala, de saber muito mais sobre o utilizador de plataformas digitais do que este terá, sequer, noção. Desde sistemas de

localização, gravação de conversas, rastreio de escolhas e hábitos, monitorização de discurso/conteúdos em redes sociais, uma nova distopia está a caminho, com já primeiros indícios preocupantes a virem da República Popular da China. Este país definiu, desde cedo na utilização da Internet pelos seus cidadãos, que o controlo do acesso a conteúdos *online* era uma necessidade premente por parte de estruturas governamentais (Swinger & Brooking, 2018). Políticas de “limpeza da *web*” foram instituídas e estas fazem com que milhões de *posts* sejam apagadas da Internet, principalmente aquelas que destoem da “perspetiva harmoniosa” da história desse país. Um conteúdo sobre um acontecimento tão importante como foram os protestos na Praça de Tiananmen, em 1989, não será encontrado em servidores de Internet chineses. Outras entradas são eliminadas por terem palavras ou frases “perigosas”, e essas chegam às trezentas (Swinger & Brooking, 2018). Governos de países como Cuba, Tailândia, Vietname, Zimbábue têm mostrado interesse em criar sistemas de controlo do estilo desenvolvido na China, e o Presidente Russo, Vladimir Putin, pediu ajuda a Pequim para que a experiência do governo Chinês sirva como modelo para engenheiros informáticos, a mando do Kremlin, desenvolvam mecanismos similares a serem aplicados na Rússia (Swinger & Brooking, 2018).

O sucesso de tal sistema de controlo na Internet precisa não só de uma grande capacidade de processamento de dados, como de controladores humanos para descortinar, com precisão, se o que é colocado *online* pode ser passível de regulação. Na Alemanha e Áustria, a negação do Holocausto é alvo de ação criminal. Se os governos Alemão ou Austríaco quisessem suprimir todas as formas de discurso e/ou conteúdos que se relacionem com esse tipo de opinião, isso seria, eventualmente, possível. Porém teriam de se tornar um país como a República Popular da China, com a necessidade de criar “aparelhos de estado” capazes de tal intervenção. Assim, têm de ser desenvolvidas outras soluções. No caso da Alemanha, uma das soluções encontradas é impedir acesso a esse tipo de informação nacionalmente, mas não noutros servidores de Internet fora do país.

Recentemente, também tem sido alvo de atenção a ação dos moderadores humanos de conteúdos em plataformas digitais de redes sociais, particularmente no Facebook. Desde os Estados Unidos até Portugal¹⁶, o cenário é de um filme de horror em contínuo. *Posts*, marcados pelos algoritmos para vistoria humana incluem violência contra mulheres, abusos de crianças, assassinatos, suicídios, terrorismo, pornografia infantil, para dar alguns exemplos. Um momento que fez despertar ainda mais interesse sobre moderação humana em plataformas digitais, e que foi um fator para autoridades governativas e legisladores

¹⁶ Para saber mais sobre histórias medonhas sobre esta função ver Newton (2019) e Neto (2019).

entrarem em ação, foi o ataque à Mesquita de Christchurch na Nova Zelândia. Por 17 minutos Brenton Tarrant transmitiu em direto no Facebook um assalto à mão armada que terminaria com 42 vítimas. Em menos de 24 horas existiam aproximadamente 1.5 milhões de cópias disponíveis na plataforma (Gunia, 2019) sem ser possível aos moderadores acompanharem tal proliferação. Como resultado, vários países, entre elas a Nova Zelândia, a Austrália, o Reino Unido e a França, criaram sistemas de multas e de responsabilização criminal das empresas que proporcionam o serviço onde tais imagens ou vídeos sejam difundidos. Isso obrigará a que cada vez mais humanos tenham de visionar atos de extrema crueldade, ou então os “guardiões de conteúdos” começarão a ser soluções digitais com critérios extremos que criaram uma sociedade esterilizada.

Esperava-se que a Internet assegurasse uma forma equitativa de discussão cívica e política. Porém, observa-se, progressivamente, apelos a medidas iliberais e até mesmo repressivas para controlo da liberdade de expressão. Talvez por causa da ingenuidade sobre o potencial da Internet, ou à atitude *laissez-faire* de legisladores, imprensa e cidadão comum, enfrentamos agora um desafio societal: a era digital é compatível com a liberdade de expressão, mesmo que não absoluta? Existirá um ponto de equilíbrio, num contínuo entre os que defendem que o protecionismo deve ser mínimo com a capacidade de um cidadão livre poder ver, ouvir e dizer aquilo que quer, dentro dos limites legais, e os que querem suprimir a história, a indecência, a crueldade, e em última análise, as ideias e argumentos. A pergunta mais importante é então: quem tem a legitimidade para encontrar esse equilíbrio?

Capítulo 4. As plataformas de redes sociais e o debate político

4.1 Liberdade de expressão nas plataformas digitais

De acordo com o relatório da Freedom House sobre o estado da liberdade da Internet, pelo nono ano consecutivo observa-se um crescimento no número de países onde é detetada uma diminuição dessa liberdade. Em 2019, o número de países onde se observou um retrocesso, ou falta de progresso, foi de trinta e três no total dos sessenta e cinco que proporcionam a informação para a produção do relatório. Foi encontrado também um novo recorde de países, quarenta e sete, onde se verificou perdas de liberdade devido a detenções de utilizadores por razões políticas, sociais e religiosas. Violência física foi também observada em trinta e um países, por razões similares (Freedom House, 2019). Dos mecanismos mais vezes utilizados para controlo de discurso e/ou conteúdos incluem-se leis repressivas, novos sistemas regulatórios, aumento na capacidade de tratamento de dados. Igualmente tem havido uma intensificação na pressão por entidades governativas para empresas privadas disponibilizarem informação pessoal dos utilizadores. Outras medidas, de natureza preventiva, são censura arbitrária, filtros digitais por palavras ou padrões de imagens e supressão de conteúdos em resposta a queixas por outros utilizadores (Freedom House, 2019).

Noutro tipo de intervenção, medidas legais foram aplicadas para o controlo da liberdade de expressão, particularmente durante períodos eleitorais em doze países dos sessenta e cinco. Dois exemplos merecem ser referidos com mais detalhe. Na Turquia, em 2018, a polícia apreendeu um conjunto de cidadãos por “insultarem” o Presidente Recep Erdoğan nas redes sociais, isto nas vésperas de eleições parlamentares. Igualmente, noutra das maiores democracias do mundo, a Índia, um jornalista foi preso por desrespeitar o *Ato de Segurança Nacional* ao criticar o Primeiro-ministro Modi no Facebook (Freedom House, 2019). Aparentemente, quando as medidas existentes não satisfazem os líderes dos países, novas são criadas para aumentar a extensão de poderes e do controlo. Por exemplo, no Bangladesh, o Governo do Primeiro-ministro Sheikh Hasina passou uma lei com o objetivo combater “certos tipos de propaganda *online*” (Freedom House, 2019). Já em 2014, a Freedom House chamava a atenção para as medidas aplicadas com o objetivo de suprimir opiniões dissidentes *online*, com instrumentos como crime por difamação, expansão dos poderes do Estado para aumentar o alcance de leis segurança, bloqueio de discurso e/ou conteúdos sem permissão legal, aumento da vigilância não autorizada, até à redução ou supressão do acesso à Internet (Freedom House, 2014: 4-7). A outra forma de restringir a liberdade de expressão é a criação de condições para a autocensura por parte dos utilizadores das plataformas, ou forçar isso

externamente. Quando isso acontece, existe a tendência de recorrer a outros fóruns, que não os mais estritamente regulados. Sítios na Internet para colocação de opiniões marginais são faltam. O Reddit tem sido uma dessas plataformas por excelência, mas tem tido o mérito de se ter tornado mais exigente ao nível da qualidade das interações e debates, não hesitando em fechar *chat groups* onde acontecem mensagens com material que incorre no discurso de ódio ou princípio de dano. Porém, já outras não têm o mesmo sentido de responsabilidade, como é o caso de sítios na Internet como o 4chan, o 8chan (na altura da produção deste trabalho renomeado 8kun e acessível apenas por navegador Tor) e o GAB¹⁷. O debate continua em aberto para saber qual é a extensão da regulação da liberdade de expressão: se se vai chegar ao ponto de fecharem-se sítios da Internet, ou se as soluções menos extremas apresentadas acima servem para resolver os problemas atuais.

O sucesso das plataformas digitais, principalmente das chamadas redes sociais, relaciona-se com a simplicidade de utilização e com a arquitetura digital, que permite disseminar e aceder a discurso e/ou conteúdos com um preço reduzido ao nível do tempo e do espaço. A juntar a isso existe a capacidade de interagir com outros utilizadores com uma facilidade que nunca antes existiu. Desde a revolução tecnológica, mas principalmente desde o aparecimento deste tipo de plataformas digitais que os fóruns de discussão informais passaram, na quase totalidade, de parques, ruas, cafés e restaurantes para o mundo virtual. Em 2017, isso foi reforçado no acórdão do Supremo Tribunal Americano no caso *Packingham v. North Carolina*, que incidiu sobre o direito de agressores sexuais poderem aceder às redes sociais. No parecer da maioria do Tribunal avisa-se que, “[s]e no passado poderia haver alguma dificuldade em identificar os lugares mais importantes (num sentido espacial) para a troca de ideias, hoje a resposta é clara. É o ciberespaço – o ‘vasto fórum democrático da Internet’, em geral, e as redes sociais em particular” (USSC, 2017: 4-5). Na regulação de discurso e conteúdos, as grandes empresas digitais têm tido uma de duas posições: ou afirmam que não devem ser consideradas como responsáveis, uma vez que não há mecanismos legais para responsabilidade intermediária, ou então que não querem determinar o que é liberdade de expressão, ou a infringir nesse direito. Nos Estados Unidos, uma das primeiras medidas a ser adotada, e que continua a determinar uma das proteções legais mais preponderante no mundo digital, é a Secção 230 do *Ato da Decência nas Comunicações*¹⁸. Nessa secção pode-se ler que “[n]enhum intermediário digital¹⁹, ou utilizador de um serviço interativo digital, deve ser tratado como um publicador, ou emissor, de qualquer informação

¹⁷ Recorrentemente, estes *sites* mudam o seu nome e acesso como resposta a um episódio de violência societal que se associa com materiais que podem ser encontrados nestas plataformas. No momento da produção deste documento os dois *sites* que estão ativos e podem ser visitados são o <https://gab.com/> e <https://4chan.com/>

¹⁸ *Communications Decency Act*, no original.

¹⁹ *Provider*, no original.

que seja proporcionada por outro intermediário digital produtor de conteúdos” (USC, 2018). Na União Europeia, a lógica de responsabilidade legal é semelhante. A *Diretiva e-Comércio de 2000/31/EC* do Parlamento Europeu e Conselho, de 8 de junho de 2000, contem exceções a princípios de responsabilidade legal que, mediante certas condições, iliba as empresas que funcionam como intermediários digitais da responsabilidade do tipo de discurso e/ou conteúdos que os seus utilizadores decidem tornar públicos. No entanto, está também expresso nessa diretiva que as empresas têm de agir de uma forma expediente para remover a informação ilegal colocada nas respetivas plataformas, desde que tenham conhecimento do mesmo, ou relativamente a reclamações por dano o conhecimento dos fatores ou circunstâncias nos quais a atividade ou informação ilegal é aparente (European Commission, 2018: 2).

Estas proteções legais permitem que milhões de pessoas possam usar a Internet no usufruto dos seus direitos fundamentais, tanto no acesso e disseminação de discurso e/ou conteúdos até à procura de informação em motores de busca, sem as empresas que servem de intermediário digital serem de tal modo avassaladas por queixas criminais que as levariam a deixar de proporcionar esses serviços. Porém, o contrário também se aplica e esta desresponsabilização das plataformas digitais também levam a uma imponderação, e experimentalismo, que tem tido consequências graves.

4.2 Os novos gigantes sociais

Em 2010, Eric Schmidt, na altura diretor-executivo do Google, discursava perante um grupo de jornalistas no World Economic Forum em Davos. A meio da sua palestra sofreu, o que pode ser considerado de uma forma jocosa, um deslize Freudiano. Schmidt, no meio de um argumento quis justificar que “[o] Google não é um país, não produz leis ou exerce diplomacia, mas precisa de defender as nossas fronteiras”. O CEO teve consciência do lapso e rapidamente emendou para “[c]orreção, as nossas redes” (Ash, 2016: 47). No início de 2020, a Microsoft anunciou que iria abrir uma representação nas Nações Unidas e um diplomata dinamarquês, antigo embaixador, seria o responsável pelo “estreitar das relações com instituições da União Europeia, assim como com os governos através da Europa” (Microsoft, 2020), abrindo assim um canal de relações institucionais que não existia até recentemente. Noutro exemplo, na primavera do mesmo ano, o Facebook criou um Conselho de Supervisão²⁰, que terá como função “reverter alguma decisão que o Facebook possa sobre permitir ou retirar *posts* da plataforma (Kaye, 2020) numa lógica de “supremo tribunal”, que

²⁰ *Oversight Board*, no original. Para saber mais visitar a página, <https://oversightboard.com/>

avaliará, em última instância, qual o mérito de uma decisão de remover, ou não, um conteúdo na sua plataforma. Porém, tirando os sistemas de reclamação, de pressão através da mobilização de outros utilizadores, de boicotes comerciais, estas companhias não têm de responder diretamente aos utilizadores. Porém, poderão ter de o fazer perante governos. Empresas digitais como o Yahoo, o Google, o Facebook e o Twitter, não são Estados com a capacidade de desenvolver e aplicar leis, podendo, no entanto, exercer influência em processos legislativos, seja de uma forma direta através de *lobbying*, ou indireta, com mudanças nas condições de acesso da utilização das plataformas. Mudanças no serviço podem causar, no entanto, que as “comunidades” de utilizadores reclamem ações por parte de governantes, que por sua vez se podem traduzir em audições parlamentares, medidas anti monopólio, legislação com aplicação de regulações por parte do estado em empresas privadas e nos mercados onde se posicionam.

Da forma como o serviço de plataformas digitais está montado neste momento, oferecendo um espaço para fóruns de discussão em redes sociais, empresas como o Facebook e o Twitter tornaram-se *stakeholders* no que se trata de regular que tipo de testes são aplicados para a utilização destes espaços pelos cidadãos e eleitores. Esta dinâmica causa tensão no mundo ocidental, onde existe menos tendências totalitaristas de controlo da liberdade de expressão, ou da produção e disseminação de discurso e/ou conteúdos por parte de entidades governamentais. Em sociedades de democracia liberal, espera-se que o Estado tenha a atitude de assegurar, e não interferir, nos direitos dos cidadãos. Porém, para defender direitos fundamentais precisam de controlar abusos, tanto por indivíduos, como por empresas que servem de intermediários digitais (Hudson, 2018). Com o Google a processar 40.000 mil pedidos de acesso a informação por segundo (em média), que resulta, também em média, de 3.5 mil milhões de pedidos por dia (ILS, 2019), com o Facebook com 2.4 mil milhões de utilizadores (Statista, 2019) e o Twitter com 330 milhões de utilizadores ativos por mês (dados do terceiro quarto de 2019) (Statista, 2019a), há a necessidade de criarem regras de funcionamento destes gigantes digitais, no que se trata da proteção da arena de discussão e do reforço do mercado de ideias.

Quando, ou se, companhias digitais instalarem sistemas de controlo sobre o que é colocado nas suas plataformas, seja por decisão interna, regulação por legisladores, ou receio de represálias por parte de governos, pode acontecer ficar diminuída a possibilidade do utilizador envolver-se em debates sociais naquele que é o espaço privilegiado para o fazer atualmente. Isso causa uma diminuição da capacidade de desenvolvimento e realização pessoal, restringindo um contributo que possa dar para o debate público (Hudson, 2018). E isto não é uma realidade improvável ou inesperada. Nos últimos anos, Twitter e Facebook têm

retirado cada vez mais utilizadores das plataformas, alguns de uma forma definitiva e irrevocável. Argumentos legais para a defesa de direitos fundamentais em plataformas digitais são suportados pela *Declaração Universal de Direitos Humanos* (Jackson, 2014), e no *Comentário Geral Nº 34 ao Artigo 19.º do Pacto Internacional dos Direitos Cívicos e Políticos*. Nesse documento é expressa a preocupação de que os sistemas regulatórios devem contemplar diferenças entre os setores de impressão, televisivos e de Internet, e que, devido ao desenvolvimento dos *mass media*, torna-se necessário que a regulação, ou dito de outra forma, o controlo sobre discurso e/ou conteúdos, não seja impeditivo do exercício da liberdade de expressão (UNHRC, 2011). Igualmente, qualquer restrição na operação de sítios na Internet, blogues ou outras disseminações eletrónicas de informação, incluindo sistemas de suporte como são os intermediários digitais ou motores de busca, são permitidos no espírito do que é apresentado no ponto 3 do artigo 19.º: respeito pelos direitos e reputações de terceiros e proteção da segurança nacional, da ordem pública, ou da saúde e morais públicas. Restrições permissíveis geralmente devem restringir-se a conteúdos específicos. Também não é aceitável que, usando as razões apresentadas no ponto 3, artigo 19.º, que se proíba discurso e/ou conteúdo por causa de uma percebida crítica a ação de um governo ou a um sistema político aplicado por um executivo (UNHRC, 2011: 11).

Tim Wu, no livro *Master Switch*, apresenta outros condicionalismos à livre expressão que resultam da forma como as plataformas digitais foram originalmente concebidas, e pelo modelo de funcionamento que adotaram. Estes modelos baseiam-se na criação de um monopólio, via adesão maciça de utilizadores, e facilidade, diversidade, robustez e estabilidade da plataforma. Desta forma, foram criadas, quase à partida, assimetrias intransponíveis para projetos similares concorrentes se instalarem no mercado. Este monopólio provoca também um poderio financeiro que permite assimilar *startups*, ainda nos primeiros estágios de desenvolvimento, como foi o caso do Facebook adquirir o Instagram e o WhatsApp, para dar dois exemplos de plataformas com milhões de utilizadores. Quando os utilizadores recorrem sempre às mesmas plataformas digitais, isso envia a procura e a partilha de informação o que faz com que sejam criados subgrupos *online* para a expressão de opiniões, incluindo as menos populares ou mais subversivas. Finalmente, outro ponto importante na apreciação da consolidação de empresas, é a vulnerabilidade à censura. Se no início de uma *startup* existe liberdade de ação na divulgação de mensagens e/ou conteúdos, à medida que a plataforma digital cresce, começa a sofrer pressões para maior controlo de opiniões e argumentos por parte de organizações governamentais com interesses em áreas comerciais ou de segurança (Wu, 2011).

Apesar ser um processo em desenvolvimento, as plataformas digitais de redes sociais têm trabalhado para a criação de espaços de debate público incluindo sistemas de proteção dos utilizadores, por vezes mesmo contra a vontade destes. Existem regras para a participação *online*, foram definidos termos e condições que obrigaram à criação de perfis, incluída moderação por humanos ou por máquinas, foram feitas parcerias com instituições que verificam factos, e alterações de algoritmos para suprimir discurso e conteúdos que incorram no princípio de dano ou discurso de ódio. Porém, existe uma atenuante que pode ser atribuída aos responsáveis por estes processos, que é a velocidade da mudança, tanto tecnológica como cultural. Os responsáveis por plataformas digitais, desde conselhos de direção até ao *staff* técnico, têm tido dificuldades em criar sistemas de controlo e de segurança capazes de antecipar tendências no mundo real que se refletem nas redes sociais. Um dos exemplos paradigmáticos é a ubiquidade de telemóveis com câmaras de filmar, a grande maioria com câmaras de alta definição e largura de banda de dados 4G. Hoje, um utilizador pode fazer o *upload* de um conteúdo em minutos, o que antes exigira meios e recursos pouco acessíveis. Realidades como estas deixam as plataformas digitais vulneráveis a comportamentos aberrantes que precisam de ser incluídos em sistemas de segurança. Desta forma, a forma mais fácil de prevenir discurso e/ou conteúdos indesejados é, no caso das plataformas digitais, a pré-censura. Filtragem de conteúdos, através de palavras-chave ou combinações dessas palavras, eliminação ou oclusão de comentários de forma a ficarem invisíveis aos utilizadores, algoritmos que detetam imagens indesejáveis, são ferramentas já existentes, e ainda mais importante, o início do processo. Como iremos abordar seguidamente, poderá haver o momento onde as proteções legais existentes no que se trata da responsabilidade legal por material colocado *online*, possam ser modificadas, ou mesmo anuladas. Isto aumenta o risco de serem censurados discursos e/ou conteúdos que não passam os testes para detetar o princípio de dano ou de discurso de ódio. E mesmo com a ajuda da componente humana, o receio é que censura à “bem” da comunidade não seja sensível o suficiente para distinguir o que é criminoso no lugar de simplesmente indesejável (EDRi, 2019).

Outro ponto de contenção é a responsabilidade social que estas empresas digitais têm, e similarmente, qual a participação dos membros da sociedade, ou como gostam de referir os responsáveis pelas plataformas, da “comunidade”, na decisão das políticas empresariais. Rebecca MacKinnon refere isso no livro *Consent of the Networked*, onde a autora alerta para um poder “não democraticamente atribuído” a um conjunto de jovens em Silicon Valley, que assumem a função de advogados de defesa e de acusação, de juizes e júris, e de polícias e legisladores, operando no ciberespaço (McKinnon, 2012: 154), quase

numa extensão da independência exigida por John Perry Barlow (Barlow, 1996). Existe uma estrutura que se assemelha à de governos, com simulacros de: políticos; legisladores; grupos de pressão; autoridades; diplomatas; juristas e quadros técnicos. Porém, estas estruturas não respondem, pelo menos de uma forma “eleitoral”, ao universo de pessoas que passam grande parte do seu tempo nas suas plataformas. No entanto, também não tem de o fazer. As empresas têm de responder aos acionistas através dos seus Presidentes ou CEO's e à sua personalidade jurídica com garantias e obrigações. Em última análise, para uma empresa, o que deve ser o princípio orientador é a obtenção de lucro. Os utilizadores podem usar o seu poder, que é passar tempo, ou não, na plataforma, ou através de campanhas de boicote, ou pressão social via imprensa tradicional ou por petições a governos.

4.2.1 Facebook

Apesar da popularidade a nível mundial, e do número e atividade dos utilizadores, é bastante notável o desconforto com políticas da empresa. Brian Acton, o cofundador do aplicativo WhatsApp, um serviço de troca de mensagens com 500 milhões de utilizadores ativos por dia²¹, que o Facebook adquiriu por 18 mil milhões de Euros em 2014, defende a ideia de cada vez mais utilizadores cancelarem a conta na plataforma digital (Shieber, 2018). Igualmente Chris Hughes, um dos fundadores do Facebook, afirmou que o poder detido numa só pessoa, Marck Zuckerberg, é “sem precedente e muito pouco americano”, defendendo que a empresa devia ser dividida noutras mais pequenas. Igualmente, Hughes defende a ideia de ser criado um regulador, em particular nos Estados Unidos sob jurisdição do Congresso, para condicionar o funcionamento da plataforma (Hughes, 2019).

O facto de milhões de pessoas obterem notícias, entretenimento e educação no Facebook faz com que estes utilizadores permaneçam na plataforma uma quantidade de tempo considerável²². Isso cria as condições ideais para a venda de espaço para anúncios. A venda de publicidade gera um poderio económico que permite ao Facebook adquirir qualquer eventual competidor enquanto este ainda está no processo de crescimento. Isso tem despertado atenção, tanto nos Estados Unidos, como na Europa, para possíveis ações anti-monopólio. Nos Estados Unidos essa ação tem sido liderada pelo Departamento de Justiça em conjugação com os Procuradores-Gerais dos Estados, e na União Europeia com a Comissão para a Concorrência. De facto, nos Estados Unidos, a Federal Trade Commission²³ (FTC), em dezembro de 2020, iniciou um processo no Tribunal Distrital para o Distrito de

²¹ Dados do primeiro trimestre de 2019. Para dados mais atualizados visitar Statista (2019b).

²² Na altura da realização deste trabalho o tempo, em média, dos utilizadores no Facebook era de 58 minutos por dia. Para saber mais ver O'Connor (2019).

²³ Para conhecer melhor o trabalho da FCT, pode-se visitar o seu website em <https://www.ftc.gov/>

Columbia, onde se encontra a cidade de Washington, e por sua vez a sede da FTC, para “prevenir a sua conduta anticompetitiva e métodos injustos de competição em ou afetar o comércio, em violação da Secção 5(a) do FTC Act, 15 U.S.C. § 45(a)”, isto após uma investigação em cooperação com Procuradores Estaduais de 46 estados, onde se alega que Facebook eliminou ameaças à sua hegemonia no mercado adquirindo a rival Instagram e WhatsApp, assim como pela imposição de condições anticompetitivas para programadores de *software* (FTC, 2020). A intenção de “quebrar” Facebook foi, naturalmente, mal recebida na empresa, com a Vice-Presidente Jennifer Newstead a garantir a “defesa vigorosa” da liberdade de escolha dos seus utilizadores, assim como a referir que a aquisição do Instagram e WhatsApp foi autorizada pela própria FCT quando ocorreram, e que a vontade do governo federal de “voltar atrás na decisão envia um aviso preocupante para as empresas Americanas que aquisições podem não ser finais” (Fung, 2020).

O Facebook tem-se mostrado, compreensivelmente mediante a dimensão da empresa e o alcance que tem, um feroz opositor a medidas de regulação por parte de instituições governamentais. Em 2016, na Europa, o Facebook entrou em conflito com a Comissão Europeia quando a empresa foi chamada para reuniões com intuito de desenvolver mecanismos de prevenção e combate ao discurso de ódio na sua plataforma. O Facebook propôs que qualquer mecanismo, ou regra de avaliação, a ser aplicado para remoção de conteúdos ofensivos ou de discurso de ódio, ficassem submetidos aos termos e condições da plataforma, no lugar de serem enquadrados pela legislação da União Europeia. Após um texto final ter sido proposto pela Comissão Europeia, o Facebook continuou a tentar convencer os membros da Comissão que as regras da empresa deviam ter precedência relativamente a legislação supranacional ou leis nacionais (Kayali, 2019). Na minuta de uma das reuniões pode-se ler que o Facebook considera haver dois tipos de leis: a privada (as normas, termos e condições que são aceite pela comunidade de utilizadores), e a lei pública (como definida por legisladores). O Facebook chegou a argumentar perante a Comissão Europeia que não havia “necessidade de regulação” (CEO, 2018), ou que a resolução de problemas no controlo de discurso e/ou conteúdos sejam de ordem tecnológica, no lugar de legislativa. Esta última, na opinião dos negociadores do Facebook, causa mais complexidade, assim como é alvo de interpretações diversas e com diferentes aplicações nos Estados-membro da União (Kayali, 2019). Não é de estranhar que o Facebook tivesse contratado para vice-presidente para Assuntos Globais e Comunicação²⁴, Nick Clegg, o antigo vice-primeiro ministro do Reino Unido, ex-Membro do Parlamento Europeu e também ex-membro da Comissão Europeia para

²⁴ *Vice-President for Global Affairs and Communications*, no original.

o comércio. É fácil assumir que Clegg se encontra numa boa posição para conseguir o melhor acesso aos centros de decisão Europeus, tanto em Londres como Bruxelas (Sweeney, 2018).

O Facebook, como qualquer outra empresa que esteja no ramo da disseminação de discurso e/ou conteúdos, sofre pressões para censurar o exercício da livre expressão. Isto acontece quando existem preocupações de natureza legal ou quando são contrárias ao *status quo* desejado por líderes políticos. Exemplos do primeiro caso são justificados pela prevenção de atividades criminais, assédio ou intimidação, publicação de pornografia e para combate ao cibercrime. Igualmente, existe a preocupação de proteger direitos de autor e não ser dado espaço a publicidade enganosa. O Facebook tem, no entanto, outras preocupações. Como uma empresa com fins lucrativos e registada em bolsas de valores, existem interesses que se relacionam com a imagem da empresa e com a legitimidade do uso da plataforma. O Facebook incorre então em “operações de chame” para defender o seu modelo de operação, evitando críticas à arquitetura da plataforma, algoritmos ou sistema de controlo de publicações. Para além do Vice-Presidente Clegg, já apresentado, junta-se a Diretora-executiva, Sheryl Sandberg, e, claro, o dono da empresa. Em TED Talks, especiais televisivos e em audiências no Senado dos Estados Unidos, Marck Zuckerberg, revela sempre uma expressão de incredulidade pelos problemas que a arquitetura da plataforma causa, com a defesa das intenções por detrás da ideia original para a criação da mesma e a utilização recorrente da expressão “comunidade”. Este é um esforço para assegurar a políticos e legisladores que os problemas são resolvidos entre a empresa e a “comunidade” e que não há necessidade de intervenções estatais, ou alterações ao funcionamento das plataformas (Wong, 2018).

Em 2019, uma decisão judicial sobre liberdade de expressão criou um precedente que será importante a nível de futuros casos de censura nesta plataforma digital. O Tribunal Distrital de Varsóvia proibiu o Facebook de remover a página, perfis e grupos relacionados com uma organização não-governamental polaca chamada Sociedade Civil para Iniciativa em Política de Drogas Leves (SIN)²⁵. SIN apresentou a queixa no tribunal por causa de o Facebook ter retirado da plataforma, de uma forma injustificada, os conteúdos que se encontravam na página da SIN, impossibilitando assim a disseminação de informação e criação de tópicos de debate com os seguidores da página. SIN alegou que, desta forma, não só não conseguiria desempenhar a sua função educativa, como a supressão dos seus conteúdos na plataforma fazia um julgamento de valor sobre os conteúdos divulgados, assim como diminuía a credibilidade da organização. Esta decisão judicial foi apresentada, por uma

²⁵ *Civil Society Drug Policy Initiative*, da tradução inglês do original em Polaco.

associação para os direitos humanos e civis e direitos digitais, como “a primeira vitória contra censura privatizada” (Panoptykon Foundation, 2019).

Porém, Facebook continua com o processo de unilateralmente banir utilizadores e conteúdos mediante alterações recentes relativamente ao que é considerado como discurso de ódio (Burch, 2020) e desinformação relativamente à pandemia Covid-19 (Burch, 2020a). O zénite desta política culminou com a decisão, aparentemente desde o topo da empresa, de remover do Facebook e do Instagram a conta associada ao então Presidente dos Estados Unidos, Donald Trump, a ser reatribuída, possivelmente, após a tomada de posse do 46º Presidente (Isaac & Conger, 2021). A justificação do CEO Mark Zuckerberg, colocado num *post* na sua conta pessoal, foi de que “o contexto atual (...) envolvendo o uso da nossa plataforma para o incitamento de uma insurreição violenta (...) o risco de permitir o Presidente de continuar a usar o nosso serviço é simplesmente demasiado grande” (Zuckerberg, 2021), o que, naturalmente, levantou muitas críticas em círculos próximos de movimentos políticos conservadores e naqueles que apontaram a falta de ação da empresa quando membros do exército e líderes budistas em Myanmar fomentaram violência contra a minoria muçulmana Rohingya (Bengali, 2021). Facebook remeteu para o recém-formado Conselho de Supervisão, apresentado anteriormente, um parecer sobre a decisão tomada, e o Conselho vê-se assim com o dilema de: defender a ordem pública ou hostilizar um número significativo de seguidores, aumentando, igualmente, a perceção pública de parcialidade e de tentativa, a tudo custo, de manter uma posição dominante como rede social.

No que se relaciona à gestão de fóruns públicos para o debate político, o Facebook tem tido um historial, que se pode dizer, de uma forma extraordinariamente generosa, problemático. Para além do caso Cambridge Analytica, a falta de controlo sobre a identidade dos compradores de publicidades políticas, sendo um dos exemplos mais emblemáticos a compra, em rublos, de publicidade para as eleições presidenciais americanas de 2016, para serem vistos por utilizadores americanos nos Estados Unidos, será algo que continuará a ser apontado ao Facebook como uma falha imperdoável. Ainda mais preocupante, e em contracorrente com a maioria das opiniões do setor político, digital e académico, o Facebook, e mais ostensivamente, Mark Zuckerberg, recusa-se a mudar a posição relativamente a avaliar qual a veracidade da publicidade política na plataforma. Insistindo que publicidades políticas devem ser vistas como neutras em conteúdo, na linha do proposto por Cass Sunstein (Sunstein, 1995: 12-13), mesmo que essa neutralidade esteja comprometida pela falta de veracidade nas mesmas, o Facebook continuará a ter uma ação negativa no que se trata da arena de discussão e mercado de ideias de natureza política. Porém, Zuckeberg

insiste em afirmar que “numa democracia, eu não penso que seja direito de empresas privadas de censurar políticos ou notícias” (Scola & Overly, 2019).

Esta política mais liberal e menos paternalista da empresa, próxima de uma posição de absolutismo da liberdade de expressão seria mais compreensível se o eleitorado fosse mais informado, participante e sofisticado. Porém, não é essa a realidade atual, nem no mundo real nem no digital. É de um otimismo irrealista assumir que o utilizador do Facebook terá a capacidade, vontade e tempo para estudar, com detalhe, questões políticas e saber se uma posição é totalmente verdadeira, qual a que, como avisa Sunstein, inclui falsidade ou engano, comparada com aquela que é totalmente errónea para criar confusão e polarização. Porém, sob pressão por parte de utilizadores e de políticos, o Facebook tentou encontrar um “meio-termo” banindo publicidade política da plataforma uma semana antes da eleição presidencial Americana de 2020, assim como se comprometeu a remover *posts* com mensagens com intuito de suprimir ou desencorajar o ato de votar (Byers, 2020).

Segundo Zuckerberg, o retorno financeiro da venda de publicidades políticas, relativamente ao total encaixado pela empresa em publicidade é de 0.5% (Scola & Overly, 2019). Isso causa ainda mais estranheza a obstinação da decisão do CEO da empresa, uma vez que parece que o incentivo financeiro não é um fator determinante. O outro efeito positivo de uma mudança desta política seria o de obrigar os agentes políticos a encontrarem melhores argumentos para tentar trazer a si votantes, ou a convencer outros a não votarem nos opositores. Tudo isto é agravado pelo facto de o Facebook se recusar a providenciar informação a legisladores, investigadores, académicos e jornalistas sobre quem compra publicidades, quais as populações-alvo a que se destinam e quais as mensagens que contêm. Esta atitude estende-se também ao funcionamento de algoritmos e de tratamento de meta dados, criando uma política digital de “caixa negra” quase inacessível. Por outro lado, o Facebook tem criado parcerias com instituições responsáveis por verificação de factos²⁶ de modo a realizar algum controlo sobre a veracidade dos conteúdos publicados, aplicando-se a conteúdos colocados em páginas de utilizadores, ou em páginas de grupos políticos, políticos individuais, fundações políticas e grupos de pressão. Porém, essas parcerias têm sido criticadas por políticos de “esquerda” e de “direita”, assim como *fact-checkers* liberais e conservadores (Newton, 2019a) que acusam o Facebook de usar essas instituições como uma operação de relações públicas, e assim evitar muitas críticas num momento de crise para a empresa (Rodriguez, 2019).

²⁶ *Fact-checkers*, no original.

4.2.2 Twitter

O Twitter tem uma arquitetura e filosofia de serviço diferente do Facebook. Originalmente pensada como um local para colocar “breves descargas de informação inconsequente” (Dorsey, 2016) a plataforma oferecia um sistema de *microblogging* para interesses, *hobbies* e curiosidades. Porém, em 2008, e com o epicentro nos Estados Unidos, o Twitter começou a ser utilizado para mensagens políticas por ativistas conservadores após a emergência de Barack Obama, e quando este se tornou num potencial candidato à presidência. Em 2009, com controlo do poder legislativo e executivo pelo Partido Democrata, um movimento popular usou o Twitter para agregar cidadãos e ajudar a formar o que seria mais tarde o Tea Party (Nguyen, 2017). O crescimento do interesse na plataforma por parte de políticos foi progressivo, por causa da forma rápida e barata para comunicar com os eleitores, atingindo o expoente com o 44º Presidente dos Estados Unidos, Barack Obama, com 111 milhões de seguidores, e o seu sucessor, Donald J. Trump, com 68 milhões²⁷. Hoje, o Twitter tornou-se um local para jornalistas, políticos, cronistas e comentadores terem uma ligação direta com os seus seguidores, em áreas tão diversas como desporto, cultura e ciência. Passado uma década, e depois de momentos tão importantes como os protestos nas eleições Iranianas de 2010, a revolução Tunisina no mesmo ano, a revolução Egípcia de 2011, os protestos de 2013 na Maidan Nezalezhnosti em Kiev, e em Hong Kong em 2019, o Twitter é uma arena de discussão política vibrante, apesar do limite de 280 caracteres por mensagem.

Ao contrário do Facebook, o Twitter tem mostrado mais agilidade em se adaptar a sinais de insatisfação por parte dos utilizadores no que se trata do funcionamento da plataforma. Por exemplo, a empresa mudou a política que vigorava anterior a 2005 de que os “*tweets* deviam fluir”, com um controlo mínimo por parte da empresa. Foram colocados sistemas de denúncia de *tweets* que contenham mensagens abusivas, misóginas e ameaças de ataques sexuais ou de divulgação de informação pessoal (Lomas, 2015). Outra resposta construtiva foi quando se tornou óbvio que a plataforma estava a ser utilizada para a ampliação de mensagens políticas através de *bots*²⁸. Em 2018 foi acrescentada à possibilidade de denunciar *bots*, ou de reportar uma conta falsa de alguém a fazer-se passar por um outro utilizador. Em julho do mesmo ano, o Twitter anunciava que tinha removido 70 milhões de perfis por violarem as políticas da plataforma digital (Lecher, 2018), apesar de isso ser uma desvantagem económica para a empresa (Kastrenakes, 2018). O mesmo observou-

²⁷ Números referentes ao final do ano de 2019.

²⁸ *Bot* é o diminutivo usado para descrever robots na Internet, na verdade programas informáticos que simulam utilizadores humanos, usando inteligência artificial para descortinarem quais os temas em tendência e criar “tráfego” na plataforma para o tópico de interesse de acordo com o programa do *bot*.

se durante a pandemia Covid-19, com a atualização dos termos e condições para assegurar que a negação de informação de peritos, o incentivo à utilização de tratamentos ineficazes, ou a transmissão de conteúdos falsos fosse razão suficiente para um utilizador ser banido da plataforma (Hatmaker, 2020).

No seguimento do *Relatório do Procurador-Geral Robert Mueller* sobre a influência nas eleições presidenciais nos Estados Unidos, 2 752 perfis com origem na Federação Russa foram permanentemente banidos por estarem associados à empresa Internet Research Agency, uma das peças centrais no relatório Mueller (Rocheleau, 2017). No final do ano 2019 seria a vez de cancelar 88 000 perfis, com origem na Arábia Saudita, que foram detetadas a criar manipulações na plataforma, principalmente com a replicação e ampliação de mensagens pró-governo Saudita (Holmes, 2019), e de 200 000 com origem na China, parte de uma operação do governo chines para influenciar a opinião pública mundial contra os protestos em Hong Kong (Hannon, 2019). A nível de política interna dos Estados Unidos, a empresa tem sido alvo de críticas, tal como o Facebook. Estas críticas vão desde a falta de *factchecking*, pelas decisões de suspensão de contas individuais, por conteúdos percebidos como antissemitas, anti-islâmicos, anti-transgénero, por ser um veículo para a promoção de teorias da conspiração, notícias falsas, ou propaganda nacionalista, supremacista e racista. Alguns dos casos mais emblemáticos são o Partido Nazi Americano, o Movimento de Resistência Nórdico, o nacionalista Britânico Tommy Robinson, Alex Jones do grupo InfoWars, a colunista Canadiana Lindsay Shepherd, entre outros.

O facto do 45º Presidente dos Estados Unidos ter usado esta plataforma como forma de comunicar com os seus seguidores, entusiastas, adversários e antagonistas, faz com que o Twitter estivesse sempre sob grande vigilância, tanto de um lado, como do outro, do espectro político. Designadamente, daqueles que acusavam a plataforma de permitir a Donald Trump *tweets* que podem facilmente ser considerados como de promoção de violência ou de discurso de ódio em testes mais liberais sobre o tipo e o valor do seu conteúdo (Grewal, 2018; Yglesias, 2019).

A escolha do Twitter de não interferir com a liberdade de expressão por parte do então Presidente dos Estados Unidos, até à sua expulsão definitiva, era justificada pelo “teor noticioso” dos *tweets* (Sandler, 2019). No entanto, o valor associado a este tipo de discurso e/ou conteúdos foi posto em causa repetidamente com pedidos que o Twitter suspendesse a conta pessoal do detentor do Ofício de Presidente dos Estados Unidos (Swisher, 2019; Sandler, 2019). Por sua vez, o aparelho político tentou também condicionar a empresa de Jack Dorsey, com a Administração Trump a ameaçar ordens executivas presidenciais para

precaver “tendências anti conservadoras” (McGill & Lippman, 2019) e legisladores a ameaçar suspender o privilégio de imunidade das plataformas de conteúdos colocados pelos utilizadores, dado pela secção 230 do Ato da Decência nas Comunicações (Tracy, 2020). Quando Twitter decidiu colocar rótulos nos *tweets* do Presidente, cada vez que este mencionava que resultado das eleições tinha sido a favor da sua reeleição, ou que existira uma fraude generalizada, o então chefe do executivo ameaçou vetar o Ato da Autorização da Defesa Nacional²⁹, por causa da “[s]ecção 230 (...) ser uma séria ameaça para a segurança nacional e integridade das eleições” (Romm, 2020). Porém, o facto de tal ameaça nunca se ter concretizado, tanto a nível diretivo como legislativo, é fácil entender, uma vez que uma remoção da proteção dada pela secção 230 seria tão prejudicial para conservadores como para liberais, porque a resposta esperada pelas plataformas digitais será de retirar discurso e/ou conteúdos de uma forma sistemática para evitar processos judiciais,

Por causa de existir vontade política, tanto nos Estados Unidos como na Europa, de alterar mecanismos de proteção das plataformas digitais de responsabilidade legal por conteúdos colocados pelos utilizadores, o Twitter defende as suas posições afirmando que: “[e]stes são constantes desafios que evoluem, e nós continuaremos os as nossas políticas e decisões sob escrutínio, particularmente à medida que vamos tendo mais conhecimento sobre a relação entre *tweets* de líderes mundiais e potencial para a produção de dano *offline*” (Twitter, 2019). Isso teve a sua expressão máxima quando Twitter banuiu da plataforma o Presidente Trump em final de mandato. Depois do ataque ao Capitólio Americano no dia 6 de janeiro, a empresa definiu, que existira uma relação direta entre os *tweets* do então presidente e o incitamento à violência, e que isso poderia continuar a acontecer se a conta se mantivesse ativa. Para o Twitter: “[o] nosso quadro de interesse público³⁰ existe para permitir que o público oiça diretamente os seus representantes eleitos. É construído num princípio que as pessoas têm o direito de interrogar os seus líderes. Porém, fizemos claro desde há alguns anos que estas contas não estão acima das nossas regras, e não podem usar Twitter para incitar violência, entre outras coisas” (Twitter, 2021). Porém, tal como exposto no caso do Facebook, também as críticas não se fizeram esperar à plataforma que permite ter contas ativas para líderes como Vladimir Putin ou Hassan Rouhani. Um desenvolvimento interessante da suspensão permanente de Donald Trump foi uma diminuição notória da disseminação de desinformação. De facto, de acordo com um estudo da empresa Signal Labs, *tweets* sobre fraude eleitoral diminuíram em 73% na semana a seguir à suspensão da conta (Dwoskin & Timberg, 2021). Porém, mais informação é necessário, quando se faz este tipo de análise,

²⁹ *National Defense Authorization Act*, no original, um ato que permite a alocação de 740 mil milhões de dólares para despesas do exército Americano.

³⁰ *Public Interest Framework*, no original.

para incluir outras variáveis dependentes e independentes que possam ajudar a produzir tais tendências.

Uma crítica positiva que se pode fazer à empresa é que esta compreendeu o problema para o processo democrático e debate político *online* da compra de publicidade política. Em outubro de 2019, poucos dias depois de se saber que o Reino Unido iria realizar eleições para o Parlamento em dezembro, e antecipando já o início da campanha eleitoral nos Estados Unidos para as Presidenciais de 2020, Jack Dorsey publicou uma sequência de *tweets* na sua conta pessoal onde fazia o argumento, louvável, que “o alcance da mensagem política deve ser ganho, não comprado” (Scola & Overly, 2019). A decisão de não permitir publicidades políticas foi tomada quando projeções das empresas Advertising Analytics, e da Cross Screen Media, previam um investimento de cerca de cinco mil milhões de Euros em publicidades políticas em 2020. Porém, é necessário acrescentar, que se estima que a maior parte dessa verba seja potencialmente gasta em publicidades na plataforma do Facebook e no motor de busca do Google, principalmente na plataforma do YouTube (Scola & Overly, 2019). No entanto, Dorsey preferiu encaminhar o futuro da plataforma para um ambiente menos controverso, principalmente depois de, durante os anos de 2017 e 2018, o Twitter ter mostrado um controlo menos conseguido a nível de contas falsas e *bots*, e de campanhas de desinformação, assédio e polarização.

O Twitter tem exercido também uma política de abertura a grupos externos para compreender melhor qual o efeito da utilização da plataforma. Alguns exemplos desses trabalhos colaborativos são os realizados pelos grupos de investigação da Universidade de Leiden e Delft na Holanda, Universidade de Syracuse nos Estados Unidos, e de Bocconi, em Itália. Muito do foco deste trabalho académico tem sido no desenvolvimento de indicadores métricos, através do tratamento de informação de bases de dados, de comportamentos e interações dos utilizadores, para avaliar como se instalam e desenvolvem atitudes de intolerância e incivilidade (Crook, 2018). É provável que Twitter mantenha a atitude de “duplicidade” no que se trata de permitir discurso político e liberdade de expressão na plataforma: por um lado não autorizar grupos extremistas, com discurso de ódio ou de dano, na plataforma, face a casos pontuais, de maior mediatismo, de suspensão de utilizadores por opiniões ou argumentos no limite do inaceitável em testes mais liberais, ao mesmo tempo que assume uma atitude paternalista de proteção dos utilizadores no uso da arena de discussão, mercado de ideias, receção e transmissão de discurso e/ou conteúdos.

4.3 A tribalização do debate político *online*

Nathaniel Persily, em 2017, no rescaldo da eleição Presidencial nos Estados Unidos, publicou um artigo no Journal da Democracia onde faz a pergunta pertinente se a democracia pode sobreviver à era digital (Persily, 2017). Ainda mais quando o futuro parece trazer mais polarização e extremismo em posições políticas, e na forma como estas são defendidas no mercado de ideias. Cass Sunstein, no livro Republic.com, argumenta que a Internet pode limitar o alcance do acesso à informação, e a discussão da mesma, não pela incapacidade de se aceder a informação, mas pela tendência, natural, de privilegiar-se a informação com que nos identificamos, enquanto se rejeita, ou se tem menos tolerância, com aquelas que não concordamos (Sunstein, 2001). Raciocínio similar foi feito por Robert Putnam, que desenvolve o conceito de ciberbalcanização, que consiste em comunicar, propositadamente, apenas com aqueles com que se partilha interesses, fazendo uma triagem da informação que desafia pontos de vista estabelecidos (Putnam, 2000). Está demonstrado que pessoas com diferentes ideologias políticas tem interpretações radicalmente opostas aquando expostos a um pacote de informação igual (Jerit & Barabas, 2012). Isso acontece devido ao que se denomina de raciocínio motivado, que tem a função principal de acomodar, seja de uma forma voluntária ou involuntária, a informação acedida às crenças bases, não possibilitando mudanças no sistema de valores ou de identidade política. Esta é uma conduta conhecida por cientistas da área comportamental e do desenvolvimento do ser humano a que se denomina como viés cognitivo.

Vieses são conhecidos na área da psicologia e da economia comportamental, transferíveis também para o âmbito do discurso político: viés de confirmação com a tendência da procura de informação que confirma pré-concepções, independentemente de serem ou não verdadeiras; viés de atenção, com a tendência de prestar atenção a estímulos emocionais dominantes; viés de crença, onde a avaliação da consistência lógica de um argumento é influenciada pela credibilidade da conclusão; viés Bayesiano com uma atitude conservadora nas crenças quando confrontadas com nova evidência; e até um efeito de viés inverso, quando um individuo acirra ainda mais as suas crenças e valores quando confrontado com evidência e factos contrários. Estes processos são reforçados quando existe uma identificação completa com uma ideologia política (Flynn, Nyhan, & Reifler, 2017). A aceitação incondicional, ou quase incondicional, de uma ideia política não é, naturalmente, um bom ponto de partida para se mudar de opinião. Esses vieses são exacerbados na arena digital para criar o que vulgarmente se chamam de bolhas ideológicas ou camaras de eco. Existe uma maior facilidade no desenvolvimento de raciocínio motivados nesses ecossistemas pela fraternidade entre os que acreditam que a sua opinião é mais válida do que a dos outros, ao que se junta

o sentimento da aprovação por todos os outros que partilham a mesma bolha ideológica ou camara de eco.

Quando *online*, a participação no debate político está vulnerável a outro fator de risco: o *feedback negativo*. Uma diferença entre opiniões declaradas ou completas pode levar a posições irredutíveis, não havendo um “meio-termo” a que os dois, ou mais, participantes possam chegar. Devia aplicar-se, aquando da exaustão da troca de argumentos, a ideia “vamos concordar em não concordar”. Porém, o *feedback* negativo causa o distanciamento entre posições, e ainda pior, entre pessoas. Não conseguir usar a máxima Voltairiana de “eu não posso concordar com a sua opinião, mas...” é um problema, mas outro mais preocupante é tornar o oponente num indesejado ou inaceitável. À medida que aumenta o tempo de permanência, e o número de interações, em debates políticos *online* tende a criar-se uma atitude cada vez mais negativa, levando a respostas similares, criando-se assim um círculo vicioso (Zollo et al, 2015). Isto ajuda a explicar, por exemplo, o que chama a “Lei de Goodwin” que postula que: à medida que uma discussão se estende no tempo, existe uma probabilidade de 100% de ser realizado um argumento de natureza comparativa, com uma posição assumida, que contem uma menção ao Nazismo ou a Adolf Hitler (Goodwin, 1994). Corre-se assim o risco de se perder um dos princípios básicos da interação no discurso político: ouvir-se uma opinião diferente da nossa. Como defende Mill, “[s]e existirem aqueles que contestem uma opinião recebida, ou que o façam se a lei ou a opinião o permitir, devemos agradecer-lhes por isso, abrir a nossa mente para os ouvir, e rejubilar que haja alguém eu faça por nós aquilo que devemos fazer” (Mill, 1859: 43). Um dos problemas evidentes da tribalização do debate político *online* é que a contestação de uma opinião recebida é cada vez menos construtiva. Existe uma progressiva facilidade em se assumir a opinião da tribo, de ter raciocínios motivados e vieses cognitivos, e entrar numa espiral de defesa argumentativa que levam a um reforço das crenças à partida e ao afastamento de negociar posições com quem defenda o contrário.

Depois existe a questão do que é popular, identitário e afetivo. Como já referido, chega-se, por vezes, a opiniões e crenças não por métodos racionais, mas por respostas emocionais. Sabe-se que posições políticas (entre outras) resultam de regulações e dinâmicas emocionais³¹ que não sendo um ponto central neste trabalho, ajudam a entender a tribalização que se observa atualmente. O sentimento de identificação com um grupo, principalmente em questões políticas, resulta primariamente na necessidade de um reforço da própria posição. Se antes da era digital, certas posições teriam muita dificuldade em ganhar

³¹ Para saber mais sobre este tema ver Lee & Fowler (2013) e Zollo *et al* (2015).

tração, atendendo à sua falsidade, inadequação ou simplesmente por serem absurdas ou descabidas, na sociedade moderna, essas opiniões não só ganham aderentes, como por vezes se tornam, nos respetivos grupos, em ideologias políticas e dogmas doutrinários. Questões de justiça social, sexualidade, iniquidades, classe, género, tornaram-se causas primordiais para uns, o que causa uma resposta negativa para outros. Essa diferença é exacerbada pela facilidade a ser-se exposto, ou pior, quando se procura ativamente, essas opiniões e o extremismo que causam. Posições sensatas podem ser anuladas por posições extremas e posições conciliadoras são suprimidas se não passam por testes de “pureza ideológica”. A tribalização via total concordância, ou a rotulação de quem não demonstra tal pureza como “outro”, remete-nos novamente para John Stuart Mill, e o seu aviso sobre a caracterização de algumas posições como “opiniões heréticas” (Mill, 1859: 44). O filósofo não poderia imaginar o quanto na era digital não ser partidário de “verdades absolutas” acarreta o peso da desaprovação social. Se o modelo Madisoniano de funcionamento da democracia passa pelo debate robusto de ideias e opiniões, e pela avaliação do mérito de argumentos na procura de soluções negociadas em comunidade, observa-se no mundo digital o desenvolvimento de um estado natural Hobbesiano. Tal desenvolvimento não augura um processo salutar, recompensante e proveitoso no envolvimento no debate político para o progresso civilizacional.

Capítulo 5. Discurso político *online* e soberania

Existe uma preocupação válida na criação de legislação e regras para a regulação de plataformas digitais transnacionais: como articular leis diferentes em países diferentes? Num exemplo paradigmático, em 2000, um tribunal francês decidiu que a plataforma Yahoo teria de retirar o acesso, em qualquer ponto do mundo, a um sítio na Internet que vendia relíquias Nazis. Aquando da decisão do tribunal, a Vice-presidente da empresa, Heather Killen, perguntou, corretamente, “[o]K, qual a lei que devemos seguir? Há muitos países e leis, mas só uma Internet” (Goldsmith & Wu, 2008: 2). Em instituições supranacionais, como é o caso da União Europeia, existe a necessidade de criar um equilíbrio entre legislação intragovernamental e intergovernamental, mediante as condições próprias de cada um dos Estados-membro. Por exemplo, tem havido a relutância em codificar alguns tipos de discurso e/ou conteúdos que são considerados, indiscutivelmente, como discurso de ódio e a serem incluídos no princípio de dano. A posição base da União Europeia, pelo menos até a tomada de posse da nova Comissão liderada por Ursula von der Leyen, tem sido que são os Estados-membro a criar legislação adicional em áreas que achem ser necessárias (Lomas, 2019). No caso apresentado anteriormente da Sociedade Civil para Iniciativa em Política de Drogas Leves, a decisão do Tribunal Distrital de Varsóvia, relativamente ao Facebook censurar conteúdos na página da organização, foi uma decisão com jurisdição na Polónia, e à luz da lei polaca.

Estes precedentes legais podem ser depois invocados em outros países, incluindo os que pertencem a organizações supranacionais como a União Europeia. Isto é particularmente importante quando diferentes sociedades têm diferentes testes para avaliar o efeito da liberdade de expressão e como esta interfere com a igualdade e dignidade do individuo, ou de grupos de indivíduos. Por exemplo, enquanto nos Estados Unidos é preciso identificar que existe uma intenção de causar dano para o controlo e censura de discurso e/ou conteúdos, em certos países da Europa, como é o caso da França, Alemanha e Áustria, banem-se material tornado público não só pela probabilidade de criar dano, mas também pelo tipo de opiniões e argumentos que contêm (UNESCO, 2015: 11). Noutro exemplo, no Reino Unido, existe o *Ato da Ordem Pública* de 1986, onde a Secção 5 proíbe linguagem ameaçadora, abusiva ou insultuosa, e onde o acusado precisa de provar, para sua defesa que, “não tinha qualquer razão para acreditar que existiam pessoas à distância de ouvir ou ler o discurso/conteúdo que causam assédio, alarme ou angústia” (UK Government, 1986). Já a Secção 18 do mesmo ato determina que um orador é acusado de ofensa se usar linguagem, comportamentos, ou que exiba material escrito, que sejam ameaçadores, abusivos ou insultuosos, e se tem como intenção, a) causar ódio racial, ou b) tendo tido em conta todas as

circunstâncias que permitem que ódio racial seja despoletado proximamente (UK Government, 1986a).

Uma uniformização das regras de como apresentar queixas, como argumentar a favor de queixosos, assim como uma uniformização nos acórdãos de decisões dos tribunais, pode ajudar a resolver problemas de perda de liberdade de expressão em diferentes países, com diferentes idiomas, custos e sistemas legais (Panoptikon Foundation, 2019). Uma das decisões do Tribunal de Justiça da União Europeia criou o caminho para isso acontecer, ao decretar que o Facebook pode ser ordenado a procurar, e identificar, todos os comentários idênticos a um primeiro comentário difamatório, se considerado como ilegal. No caso *Eva Glawischnig-Piesczek v. Facebook Ireland Limited*, ficou provado que tinha sido requerido à plataforma, pela utilizadora, a remoção de um comentário difamatório que a envolvia, na altura sendo Presidente do Grupo Parlamentar dos Verdes na Áustria, algo que tinha sido adjudicado pelo Supremo Tribunal Austríaco. O Tribunal Europeu considerou que, ao abrigo da *Diretiva e-Comércio de 2000/31/EC*, apesar de não ser da responsabilidade da plataforma quais os conteúdos nela colocados, uma vez sendo solicitado a remoção desses conteúdos a empresa tem a obrigação de o fazer de uma forma célere. O facto mais importante no caso de Eva Glawischnig-Piesczek, é que foi determinado que não deve existir condicionalismos territoriais para a remoção do material considerado como ilegal (CJEU, 2019). Como Alexander Tsesis já propunha em 2001, “se as transmissões da Internet são enviadas e recebidas em localizações particulares, então instâncias específicas retêm jurisdição para indiciar atividades ilegais na Internet” (Tsesis, 2001: 822).

O artigo 19.º do *Pacto Internacional de Direitos Cíveis e Políticos* (UNHRC, 1966) define que podem ser aplicadas restrições legais, desde que seja necessário por respeito aos direitos e reputações de outros, e para a proteção da segurança nacional, ordem pública, e saúde e morais públicas. No entanto, tais restrições precisam de ter o reforço de uma legislação formal, ou normas legais equivalentes que expressem, de uma forma clara e adequada, qual o direito, e nível de interferência no exercício da livre expressão por parte de governos e de reguladores supranacionais. As Nações Unidas têm proposto normas e procedimentos que se sobrepõem a decisões do sector privado, tanto na Internet como na imprensa tradicional. No documento *Princípios Orientadores sobre Empresas e Direitos Humanos*³², e em particular no ponto 11, pode-se ler que “[e]mpresas devem respeitar os direitos humanos. Isto significa que devem evitar infringir os direitos humanos de outros, e devem contemplar os impactos adversos que as suas atividades possam ter nos direitos

³² *Guiding Principles on Business and Human Rights*, no original.

humanos” (UNHRC, 2011a: 13). Para que isso possa ser operacionalizado, as empresas de plataformas digitais devem avaliar atuais, e potenciais, efeitos que possam originar violações de direitos, com a capacidade de intervir nesses casos após a sua descoberta, monitorizar respostas e comunicar a entidades interessadas como quais os resultados dessas intervenções (UNHRC, 2011a: 17), e no caso de haver violações efetivas as empresas devem ser capazes de “providenciar e cooperar na sua remediação através de processos legítimos” (UNHRC, 2011a: 24).

Esforços de homogeneização de leis e regras comportam, no entanto, outros problemas, como é o caso de uma tendência, por parte de instituições supranacionais, de medidas paternalistas e de testes muito restritivos. Isto é feito, naturalmente, com o objetivo de defender os utilizadores de serem expostos a material indesejável. Porém, haverá sempre o risco de quem tem mais poder determinar aquilo que lhes mais convém. Na Europa, alguns dos países mais influentes tendem a empurrar outros com menos poder para testes mais restritos à livre expressão, como é o caso de leis de blasfémia ou de criminalização da negação do Holocausto. Num mundo progressivamente interligado, e com a facilidade em produzir e disseminar discurso e/ou conteúdos, qualquer cidadão deve beneficiar, na perspetiva de Thomas Paine e John Stuart Mill, de ser exposto a outras perspetivas sobre problemas comuns, independentemente de qual a cultura onde se encontra inserido. A liberdade de expressão serve para apontar para injustiças, tensões e assimetrias que merecem uma atenção de global a local. O Relator Especial do Conselho para os Direitos Humanos das Nações Unidas em Questões de Minorias apresenta uma proposta para um consenso mais alargado. Começando pelo nível mais elevado, deve ser encontrado o limite para que discurso e/ou conteúdos constituam crime à luz da lei internacional e passível de processo criminal. Segundo, definir as expressões que não sendo criminalmente puníveis, possam justificar restrição e/ou despoletar casos civis. Terceiro, as expressões que não atingem o limite de sanções civis ou criminais, mas que levantem preocupações a nível de tolerância, civilidade e respeito pela dignidade dos outros (USHRC, 2015).

5.1 A questão da União Europeia

A União Europeia, particularmente através do seu órgão executivo, a Comissão Europeia, tem estado atenta a questões de legalidade nacionais e supranacionais, ciente que existem diferenças entre Estados-membro na forma como regulam, ou desejam regular, plataformas digitais, incluindo as de redes sociais. Em 2014, a Comissão fez uma comunicação ao Parlamento Europeu, ao Conselho Europeu, e ao Comité Económico e Social Europeu e ao Comité das Regiões, com o título *A política e a governação da Internet. O Papel*

da Europa na configuração da governação da Internet no futuro³³ (European Commission, 2014). No ponto 8 da comunicação, na secção *Conflitos de Competências Jurisdicionais e Legislações*, pode-se ler que, “[a]plicação extraterritorial de legislação nacional, por vezes com base na geografia do sistema de nomes de domínio, levou a um certo número de decisões judiciais contraditórias” (European Commission, 2014: 10), e que por causa disso “a Comissão Europeia irá lançar uma análise aprofundada dos riscos, a nível internacional, de conflitos de leis e jurisdições suscitadas pela Internet, e avaliar todos os mecanismos, processos e instrumentos disponíveis e necessários para resolver tais conflitos” (European Commission, 2014: 12).

A nova Comissão Europeia, resultante das eleições para o Parlamento Europeu de 2019, e sob a Presidência de Ursula von der Leyen e das duas Vice-Presidências de Frans Timmermans e Margrethe Vestager, promete a revisão de algumas das medidas e diretivas europeias para a regulação do setor. Em assuntos vitais, como o combate à apologia *online* de atos e ideologias terroristas, proteção de dados pessoais e utilizações indevidas de material ao abrigo dos direitos de autor, tem-se observado progressos, alguns deles apesar da resistência das empresas digitais. Por exemplo, e como apresentado anteriormente, em 2015 o Facebook mostrou preocupações relativamente à possibilidade de ser sujeito a regulações por parte da Comissão Europeia, encetando um conjunto de reuniões para apresentar apreensões, condicionalismos e propostas de regulamentação (CEO, 2018). Durante o período de 2014 a 2019 o então Presidente da Comissão Europeia, Jean-Claude Juncker, preferiu, ou foi incapaz, de promover alterações estruturais à forma como a União Europeia lidava com empresas de plataformas digitais. A diretiva na União Europeia relativamente à regulação deste setor, na altura da elaboração deste trabalho, continua a ser a *Diretiva e-Comércio de 2000/31/EC* (European Union, 2000), apesar de já proposta um novo *Ato de Serviços Digitais*, abordado seguidamente. Este documento determina um conjunto de regras no espaço da União Europeia, sendo que a mais importante, em futuras negociações ou atos executivos, é qual a responsabilidade, e implicações legais, da remoção de discurso e/ou conteúdos colocados nas plataformas que incorram no princípio de dano, seja antes ou pós publicação, assim como qual a velocidade com que isso tem de acontecer e punições no caso de incumprimento.

A questão de como incluir, numa única diretiva, diferentes exigências provenientes de diferentes Estados-membro será um dos problemas, para a União Europeia, a resolver num futuro próximo. Por exemplo, como articular com outros países uma lei draconiana como

³³ *Internet Policy and Governance. Europe's role in shaping the future of Internet Governance*, no original.

é o *Ato Alemão da Aplicação da Lei na Internet*³⁴, vulgarmente conhecida por Lei NetzDG, onde legisladores querem que as plataformas digitais tenham uma ação muito rápida no que se trata da remoção de discurso que incorre em dano. Multas de 50 milhões de Euros são as coimas aplicadas no caso de haver uma demora “não justificada” na remoção de material de índole criminal, uma vez sinalizados como tal. Alguns indicadores por parte do governo alemão apontam para um reforço ainda maior da lei, com a possibilidade de serem proactivamente removidos discurso e/ou conteúdos ilegais como é caso de ameaças de morte e discurso de ódio (Die Bundesregierung, 2019). Este é um possível modelo está também a ser explorado para aplicação em França (France Press, 2019) e no ex-membro da União, o Reino Unido (Dickson, 2019). Existem também sinais, tanto no governo alemão como do Reino Unido para a criação de uma unidade policial específica, em articulação com empresas digitais, para um acesso legal e rápido à informação pessoal e IP³⁵ do utilizador, o que permite saber a localização física de quem está a realizar uma atividade criminosa.

Noutro exemplo, a República da Áustria passou a lei *Rigor e Responsabilidade na Internet*, que, basicamente, força as plataformas de redes sociais a guardar dados dos utilizadores, a os entregar às autoridades policiais, mas também a outros utilizadores, no caso de haver uma queixa de possível ofensa. Obviamente, um dos riscos de tal lei é o potencial abuso por parte das autoridades, ou de indivíduos privados, de conseguir aceder a um conjunto de informação pessoal que pode ser usada para outros propósitos (Epicenter Works, 2019). O crescimento de tendências iliberais em certos membros da União, como é o caso de países do centro-este Europeu é também um facto extra a juntar à heterogeneidade de códigos legais. O dismantelar do Estado de Direito nessas nações e os comportamentos autoritários dos seus líderes, que se traduzem em controlo da imprensa tradicional e digital, promete tornar ainda mais complicado um consenso na revisão, a partir de Bruxelas, das regras da União Europeia relativamente a aspetos legais nos serviços da sociedade de informação no mercado interno.

5.2 Leis, normas e reguladores

A União Europeia propõe um novo ato que substituirá a *Diretiva e-Comércio de 2000/31/EC*. Esta diretiva foi criada para servir de estrutura legal na regulamentação de serviços digitais e com uma das preocupações iniciais o comércio eletrónico no mercado interno. A diretiva estabelece regras sobre informação e transparência, requerimentos para o serviço de plataformas digitais, comunicações comerciais, contratos eletrónicos e limitações

³⁴ *Netzwerkdurchsetzungsgesetz*, no original.

³⁵ *Internet provider*, no original.

de quais as responsabilidades legais das plataformas digitais como intermediárias digitais (EPC, 2000). No que se relaciona com a liberdade de expressão, a regra mais importante define que os intermediários de serviços digitais não serão responsabilizados pelo discurso e/ou conteúdos que os utilizadores colocarem nas plataformas digitais, espelhando a secção 230 do Ato da Decência nas Comunicações americano. As seguintes condições estão codificadas na diretiva: a) uma vez que a plataforma detete que existe material de natureza ilegal, é necessário proceder à sua remoção de uma forma expediente; b) as plataformas digitais devem ter um papel neutral, passivo e meramente técnico em relação aos conteúdos colocados pelos utilizadores (EPC, 2000: L178/6).

A preocupação na proteção dos intermediários existe também na Organização das Nações Unidas para a Educação, a Ciência e a Cultura (UNESCO) em 2011. Num documento da sua direção executiva, fica feito o alerta que empresas que providenciam acesso à Internet poderão um dia serem legalmente responsabilizadas por conteúdos ilegais. Na opinião da organização esse desenvolvimento pode resultar na diminuição de fluir de informação sem impedimentos (UNESCO, 2011). Também em 2012, num relatório produzido pela Comissão Europeia, menciona-se que existem questões sobre responsabilidade legais que necessitam clarificação (European Commission, 2012). Com o aproximar do final do mandato da Comissão liderada por Jean-Claude Juncker, e principalmente depois dos sucessivos escândalos sobre a falta de proteção dos utilizadores, a que se juntou o desinteresse das plataformas digitais na proteção da democracia, ou pior, a conviência para o fomento de polarização a troco de compras de publicidades, começaram a surgir os primeiros indícios que a União Europeia estaria a trabalhar num novo documento para substituir a *Diretiva e-Comércio de 2000/31/EC*, com o nome *Ato de Serviços Digitais*³⁶ (European Commission, 2020). A Presidente da Comissão Ursula von der Leyen, já expressara a intenção de criar uma agenda ambiciosa no que se trata da União Europeia lidar em questões digitais (Naranjo, 2019).

A génese do *Ato de Serviços Digitais* foi conhecida em julho de 2019 quando um memorando interno da Comissão Europeia foi publicado em sítios da Internet especializados em questões de informação e tecnologia, com posterior eco na imprensa generalizada. Neste documento já se podia encontrar algumas das preocupações de legisladores e reguladores para o futuro da operação de plataformas digitais, intermediários digitais e empresas que proporcionam serviços de acesso à Internet (DGCNCT, 2019). De uma forma mais precisa, quando o ato foi tornado público (European Commission, 2020), alguns pontos, principalmente

³⁶ *Digital Services Act*, no original.

de interesse para este trabalho, assumem especial relevo. Entre eles vale a pena destacar a modernização da proteção das plataformas para discurso e/ou conteúdos colocados pelos utilizadores, onde agora é acrescentada a necessidade de serem colocados avisos e descrição de quais os mecanismos para a identificação de conteúdo ilegal. As plataformas precisam de estabelecer sistemas internos para gestão de queixas, tentar resolver questões sem recorrer a tribunais, e dar prioridade a notificações de organizações que se qualificaram como “reguladores” de conteúdos. É requerido também que as empresas digitais sejam mais transparentes nos processos de combate a discurso e/ou conteúdos ilegais, e uma vez estes a serem removidos, uma explicação das razões para o fazer deve ser enviada para quem colocou na plataforma esse material. Um relatório deve estar disponível com discurso e/ou conteúdos ilegais, ou contrários aos termos e condições, que foram removidos. Relativamente a publicidades em plataformas digitais, estas devem conter informação clara que são publicidades, quem é responsável pelas mesmas e quais foram os parâmetros utilizados para chegar ao utilizador

Para plataformas com mais de 45 milhões de utilizadores ativos por mês (onde facilmente se inclui o Facebook e o Twitter), é necessário também uma análise sistémica de riscos pela utilização da plataforma para a aplicação de uma moderação efetiva na gestão desses riscos, seja material ilegal, ou violações de privacidade. Os parâmetros para o funcionamento de algoritmos (incluindo preferências de discurso e/ou conteúdos) devem ser transparentes. Igualmente, o utilizador deve ter a possibilidade de modificar esses parâmetros. Relativamente ao funcionamento destas plataformas digitais, a União Europeia exigirá que estabeleçam repositórios com a informação de quem comprou publicidades e dar acesso a dados internos de modo a poder monitorizar-se o cumprimento das regras do ato. Igualmente, essa informação será tornada disponível para investigadores e académicos para trabalhos sobre riscos sistémicos. Finalmente, e igualmente importante, é a intenção da Comissão Europeia de ter poderes de supervisão e de gestão em plataformas de grande dimensão a nível de utilizadores, o que será uma negociação que promete ser difícil com gigantes como Facebook, Twitter e Alphabet.

Discurso e/ou conteúdos que se enquadrem no princípio de dano, que não sendo necessariamente de âmbito criminal, deve ser revisto à luz das alterações que possam existir na definição de discurso de dano. Isto faz com que possam ser alterados os limites dos testes para queixas por parte de utilizadores. Nesse particular, já existe um precedente proveniente do Tribunal de Justiça Europeu, particularmente no que se trata do âmbito do artigo 14.º da *Diretiva e-Comércio de 2000/31/EC*, com o nome *Armazenagem em Servidor* (EPC, 2000). Este artigo reforça a necessidade de haver mecanismos transparentes, e exatos, para

responder a reclamações e pedidos de remoção de discurso e/ou conteúdos de plataformas digitais. Igualmente, está definido na diretiva que deve haver uma resposta em tempo útil, entendida também como o mais rápida possível, de maneira a informar os queixosos de quais as diligências tomadas, particularmente às razões para a remoção ou o cancelamento do acesso a de discurso e/ou conteúdos (European Commission, 2018: 5). Estas medidas transitam para o novo ato.

Finalmente, de referir a intenção de existir uma maior proximidade entre as empresas digitais, legisladores e reguladores. No *Ato de Serviços Digitais* (European Commission, 2020) fica expresso que as empresas que proporcionam serviços na União Europeia devem designar representantes legais para cooperar com autoridades reguladoras: a Comissão Europeia e o *European Board for Digital Services*, sendo este último um grupo pan-europeu que estará encarregue da harmonização do ato. Igualmente, para plataformas digitais com mais de 45 milhões de utilizadores mês, terá de ser designado um técnico exclusivamente responsável pela aplicação do ato, e que fará parte do processo de permitir uma auditoria anual por uma organização independente.

Quando o ato foi tornado público, a imprensa especializada, organizações não-governamentais da área digital e grupos cívicos começaram a expressar a intenção de contribuir para o processo da sua implementação. A comunidade digital, incluindo os que a constroem, gerem e utilizam, deve estar atenta ao processo, as decisões, e influências. Para a democracia florescer, principalmente uma democracia forte, como definida por Benjamim Barber, as instituições, têm de “lidar concretamente com os obstáculos que a modernidade parece colocar no que se relaciona à participação: nomeadamente, escala, tecnologia, complexidade e o paradoxo do paroquialismo” (Barber, 2003: 262). É um “admirável mundo novo”, mas também pode ser um de cada vez menos liberdades e direitos.

5.3 Qual o melhor equilíbrio?

Diferentes quadros legais e normativos em diferentes Estados-membro da União Europeia causam o risco que definições adotadas pela Comissão Europeia, possam manter-se ambíguas. Enquanto está definido, com mais ou menos clareza o que é ilegal quando do exercício da livre expressão, o que pode criar dano é subjetivo, dependendo do contexto, modo de disseminação, população a quem se destina, e a capacidade de se evitar esse tipo de discurso e/ou conteúdo. Estes receios resultam de leis existentes na União Europeia, e agora em ex-Estados-membro. Por exemplo, a tentativa da Comissão Europeia de regular

plataformas digitais segue exemplos como o do Reino Unido (UKDDCMS, 2019), ou da Lei NetzDG na Alemanha que ilegaliza discurso de ódio (Die Bundesregierung, 2019).

Porém, a alteração ao *status quo* que terá mais impacto, é a crescente exigência de se retirar discurso e/ou conteúdos considerados como ilegais, criando um maior ónus para as plataformas digitais agirem de uma forma proactiva na remoção de material. Por causa da crescente pressão de serem as empresas digitais a evitar problemas relacionados com aquilo que os utilizadores carregam para as suas plataformas, estas irão ter a tendência de “sobre-bloquear”, para evitar responsabilidades legais *à posteriori*, discurso e conteúdos, diminuindo assim a capacidade dos utilizadores de divulgarem ou acederem a informação. Estas alterações podem ser a nível de criação de termos e condições mais restritos, implementação de sistemas de filtros digitais e algoritmos para remoção de discurso e/ou conteúdos de uma forma assimétrica e desproporcional.

Vital para o futuro da relação entre plataformas digitais, legisladores e utilizadores, é a manutenção da medida conhecida como *Safe Harbour Law* que isenta as companhias digitais de responsabilidades legais por discurso e/ou conteúdos carregados para as respetivas plataformas. Esse não é o caso, por exemplo, da NetzDG, o *Ato Alemão da Aplicação da Lei na Internet*, mencionada anteriormente, que é uma das leis mais vezes invocada como modelo a seguir por países da Europa e da União Europeia. Neste ato encontram-se categorias como “incitamento ao ódio”, “disseminação de imagens violentas”, “fomento de organizações terroristas” e a utilização de “símbolos de organizações não constitucionais”. Outras categorias são mais universalmente utilizadas como “distribuição de pornografia infantil”, difamação, difamação de religiões, e de associações ideológicas e religiosas de maneira a causar distúrbios na paz pública, “violação da privacidade”, “ameaças para a concretização de um crime” e “falsificação de dados para intenção de prova” (TWG, 2019).

Porém, uma das categorias incluídas é o “insulto”, o que levanta um problema similar ao apresentado no princípio da ofensa. A principal preocupação para os defensores da liberdade de expressão com a lei NetzDG é a ausência de conceitos definidos de uma forma clara sobre o que é insulto, ofensa e simplesmente mau-gosto. Isso pode levar ao sobre-bloquear e sobre-remover conteúdos, devido às multas pesadas pela demora na remoção de material determinado como ilegal. A lei NetzDG não requer que haja uma decisão arbitral para remoção de discurso e/ou conteúdos, assim como está mal definido quais os mecanismos de apelo para aqueles que querem a decisão revista por instituições independentes (TWG, 2019: 3). O risco é de diminuir a importância dos tribunais ou outras organizações democráticas

existentes para o efeito. Também na República da Áustria, a já mencionada lei sobre a lei do *Rigor e Responsabilidade na Internet* causa a obrigação das plataformas digitais de partilhar informação de utilizadores tanto com autoridades, mas também com terceiros que apresentem queixas sobre o exercício de discurso ou a colocação de conteúdos (Epicenter Works, 2019). Isto é particularmente preocupante por causa do acesso, sem ordem judicial, a dados sensíveis, como determinados pelo *Regulamento Geral de Proteção de Dados*, como é o caso de nome, morada, números de identificação sociais ou IP's.

A outra consequência indesejada é autocensura por parte de utilizadores para evitar complicações, o que diminui o debate político na praça pública, uma das condições essenciais, na perspetiva de James Madison, para uma democracia participativa. No *White Paper on Online Harms*, do Reino Unido, um dos conceitos repetidos durante o documento é a necessidade de ser repensado o controlo legislativo sobre o serviço oferecido pelas plataformas de redes sociais. Um dos aspetos positivos é a apresentação da defesa da liberdade de expressão, como valor central e alvo de uma proteção extra por parte de reguladores (UK Government, 2019). Aquando da publicação do *Código de Conduta da União Europeia para Discurso de Ódio*³⁷ (European Commission, 2016), foi definido como medidas de sucesso, para as empresas digitais, o número de contas *online* eliminadas, o volume de conteúdos removidos e a velocidade com que isso acontece. Isso pode causar que as plataformas digitais removam “por excesso” conteúdos e discurso, o que pode levar à retração dos utilizadores de expressarem as suas opiniões e argumentos, exatamente pelo receio de serem banidos das plataformas, ou terem os seus contributos censurados ou alterados.

Aquando da produção deste trabalho, continua a servir como referência o artigo 14.º da *Diretiva e-Comércio de 2000/31/EC*, uma vez que o atual ato é uma proposta legislativa que terá de ser aprovada pelo Conselho Europeu e Parlamento Europeu, o que se espera demorar cerca de 18 meses (verão de 2022). Assim as empresas digitais continuam a não ser obrigadas a pré avaliar o conteúdo que é colocado pelos utilizadores. Esta decisão foi complementada por decisões do Tribunal Europeu de Justiça que proíbe os Estados-membro de imporem a obrigação de uma monitorização generalizada das plataformas digitais na procura de conteúdos ilegais (Fanta & Rudl, 2019).

³⁷ *EU's Code of Conduct on Hate Speech*, no original.

Capítulo 6. Criar condições para o debate político *online*

A livre expressão, enquadrada numa ideia de governação democrática através de uma discussão pública, é base essencial para deliberações democráticas, como defendido por James Madison. Os melhores resultados são conseguidos através de uma discussão alargada, onde as condições de igualdade são asseguradas à partida, mesmo entre pessoas de diferentes posições, estatuto ou autoridade. Neste sistema inclui-se a desconfiança ao viés institucional, a defesa do carácter experimental do debate, e a igualdade política entre participantes na arena de discussão (Sunstein, 1995: 248). O mesmo deve ser produzido quando o debate público acontece no mundo digital. Para isso é necessário criar as condições para que o debate possa acontecer, mas mais importante, que seja bem-sucedido. Essas condições são de natureza prática e conceptual: prática como é o caso de alterações de processos informáticos e competências dos utilizadores, mas também qual a melhor regulação e caracterização de conceitos chave como é a arena de discussão e o mercado de ideias.

6.1 Competências digitais do cidadão

No sentido democrático, a participação do cidadão na comunidade tem o intuito de produzir um sistema de autogovernação. Por sua vez, uma comunidade que se autogoverna necessita da participação do cidadão (Barber, 2003). Na perspetiva de uma democracia forte, estes dois conceitos: participação e comunidade, “são os aspetos de um modo único de ser-se social: cidadania” (Barber, 2003: 155). Quando, numa comunidade os cidadãos não procuram ativamente a participação, corre-se o risco de se gerarem consensos sem uma reflexão associada. Conformidade pode ser requerida para quem tem ideias diferentes que eventualmente termina num coletivismo unitário, suprimindo cidadania e autonomia da qual depende a atividade política. Por outro lado, participação sem comunidade cria iniciativas privadas sem direção e coerência, levando a conflitos competitivos mediante interesses individuais. Assim, a não participação racionaliza o coletivismo, dando-lhe uma aura de legitimidade, enquanto a participação sem comunidade racionaliza individualismo, dando-lhe uma aura de democracia (Barber, 2003: 155).

Cidadania é a utilização dos direitos individuais numa lógica construtiva, e com objetivo de ajudar no desenvolvimento coletivo. Esta lógica aplica-se também à “vida” *online*. A “cidadania virtual”, sendo um termo que ainda não se encontra suficientemente codificado, é, assim, um motivo de interesse e de investigação. Uma das principais perguntas é: quais as relações que o mundo digital constrói internamente, e quais os efeitos que transporta para o

mundo real? Para responder a tais questões pode-se, primeiro, começar por definir o que é um “cidadão digital”. Em Portugal, estatísticas apontam para um tempo, em média, de permanência *online* de 2 horas por dia, com 27 visitas, também em média, à Internet (Garcia, 2019). Com este tempo passado a olhar para um ecrã, a utilização, e usufruto, do mundo digital requer competências básicas, suficiente para navegar no ciberespaço, ou utilizado outra expressão clássica, as “autoestradas de informação”.

Uma definição proposta para o cidadão digital é aquele que “usa tecnologia frequentemente, que usa tecnologia para informação política e assim cumprir um dever cívico, e aquele que usa a tecnologia em casa para benefício económico” (Mossberger, Tolbert, & McNeal, 2008). Ao cidadão digital, tal como no mundo real, é esperado que vá além do interesse pelas redes sociais, ou pela aquisição de serviços ou bens *online* e da participação em fóruns públicos de discussão, políticos ou não. Espera-se também que o utilizador tenha em conta o bem comum. Alexis de Tocqueville defendeu que a participação de cidadãos informados é crucial para o exercício de democracia (Tocqueville, 1835), na perspetiva Madisoniana da construção de uma República. As decisões, ações e interações do cidadão devem ser baseadas em conhecimento factual, verídico e informado, de maneira a que a sua participação no funcionamento da sociedade seja produtivo. Este enquadramento é ao mesmo tempo republicano e liberal: republicano no sentido que existe uma procura de desenvolvimento da sociedade de uma forma virtuosa, mas igualmente liberal por causa da procura de existirem condições para a igualdade entre membros da sociedade. Não é descabido acreditar que a Internet proporciona essa igualdade, ou, pelo menos, tem as condições para tal. O maior acesso a informação, principalmente de índole política, permite obter um nível mais elevado de conhecimento, criando assim mais interesse sobre a ação política e uma participação mais assídua em processos de tomadas de decisões administrativas (Mossberger, Tolbert, & McNeal, 2008).

Importante é também o que se denomina como “o significado social do uso da tecnologia”, e quais as implicações a nível de normas, comportamentos éticos, proteção de direitos humanos e responsabilidade societal (UNESCO, 2015: 47). O “cidadão virtual”, ativo na arena política, não é só um consumidor de discurso e/ou conteúdos, mas também pode ser um cronista, um jornalista, um formador de opiniões e um curador de informação. Existe a necessidade de um trabalho permanente na área da literacia digital para o consumo de informação e de como utilizar a Internet. Este trabalho deve refletir a evolução tecnológica e das dinâmicas que se desenvolvem *online* entre utilizadores (UNESCO, 2014). Para além de soluções tecnológicas, é necessária uma mudança cultural para uma cidadania positiva, com objetivo de assegurar direitos fundamentais também no mundo virtual. Os utilizadores devem

ser encorajados a conhecer, respeitar e defender os direitos humanos. Complacência pode também ser problema no mundo virtual. Existe a tendência de pensar-se na Internet como uma arena de direitos, o que é refletido pela forma supérflua como é utilizada pela grande maioria dos utilizadores. É importante criar uma cultura de responsabilidade, através da educação contínua, com material produzido pelo trabalho de grupos que se interessam por direitos humanos, democracia, governação e indústria (Puddephatt, 2011: 17). Programas de desenvolvimento de literacia digital devem incluir o desenvolvimento de *skills* técnicos para uma navegação segura, onde se inclui a capacidade de avaliar a qualidade de discurso e/ou conteúdos, interpretar informação complexa, produção de opiniões e argumentos com valor adicionado de forma a enriquecer o mercado de ideias.

Uma das populações que necessita de atenção especial, no que se trata de desenvolver competências digitais, são os jovens e crianças. Esta população mais jovem, e impressionável, é facilmente manipulável para reproduzir, e mesmo criar, discurso de dano e de ódio. É preciso envolver a comunidade de agentes que interagem com os jovens, como pais, educadores e professores. Todos estes agentes devem ser treinados em como proteger e educar os mais jovens para saber lidar com a vida *online*. Deste trabalho destaca-se a preparação para lidar com a diversidade, fazendo isso de uma forma franca e esclarecida. Assuntos fraturantes como raça, religião, tendências sexuais, migrações, ou ideologia política não são assuntos novos, mas são exacerbados por causa do efeito amplificador da Internet, tanto a nível de alcance como de volume de material *online*. As gerações que agora passam grande parte do dia na Internet são as que mais precisam de entender os perigos da falta de controlo do discurso, ou da ausência de regras para um debate robusto de ideias.

Existem exemplos de medidas governamentais, com sucesso, para preparar os utilizadores do mundo virtual, particularmente desde uma idade mais jovem. A Finlândia, que ocupa o topo da lista de países Europeus no Índice de Literacia Mediática, tem instituído cursos escolares onde os alunos são treinados em como identificar imagens e vídeos *online* alterados, detetar informação maliciosa, combater intimidação, e lidar com perfis falsos, *trolls* e *bots* (Mackintosh, 2016). Na Estónia, desde o jardim-de-infância até ao 12º ano, são ensinados *skills* digitais de maneira a aumentar a segurança e usufruto da vida *online*. Juntamente com parceiros do setor privado e académico, o governo estoniano criou um currículo para desenvolver pensamento crítico aquando da exposição a discurso e/ou conteúdos na Internet, assim como proteger informação pessoal e compreender as regras relativamente à utilização de material licenciado (Rohaidi, 2019).

A nível da União Europeia, a Comissão Europeia desenvolveu o *DigComp 2.0: Estrutura de Competência Digital para Cidadãos*³⁸. Este documento visa oferecer ferramentas para aumentar a competência digital dentro de um projeto mais alargado chamado de *EU Science Hub*, um serviço da Comissão Europeia para ciência e conhecimento³⁹, e nele é apresentado um primeiro quadro conceptual para o desenvolvimento de cinco grandes áreas: informação e literacia digital; comunicação e colaboração; criação de conteúdos digitais; segurança e resolução de problemas (Vuorikari, Punie, Carretero & Van den Brande, 2016). Na sociedade civil, e a nível de trabalho de organizações não-governamentais, existem grupos com programas para o desenvolvimento de competências digitais e feitos de forma a chegar a para um leque alargado de agentes societais: jovens; educadores; pais; políticos; *bloggers* e imprensa. Alguns exemplos dignos de apresentar são a Anti Defamation League⁴⁰, a No Hate Speech Movement⁴¹ e a MediaSmarts⁴².

Em Portugal, a grande maioria dos utilizadores da Internet não conhece a legislação que regula a publicação de discurso e/ou conteúdos ilegais nas plataformas digitais, ou os mecanismos para apresentar uma reclamação ou uma queixa (ECRI, 2018: 22). É sugerido que o Estado português desenvolva campanhas de informação sobre legislação existente, e quais as autoridades a que se pode recorrer no caso de haver infrações. Existem sistemas para a prevenção, deteção e remoção de conteúdos ilegais, tanto *online* como na imprensa tradicional e são aplicados por organizações como o Conselho Deontológico do Sindicato de Jornalistas, a Comissão de Carteira Profissional de Jornalista (CCPJ), que está encarregue da distribuição de carteiras profissionais à imprensa, e a Entidade Reguladora para a Comunicação Social (ERC). Por ordem, o Conselho Deontológico delibera sobre queixas que se relacionem com possíveis violações do respetivo código. A CCPJ pode retirar a carteira a um jornalista a quem tenha sido comprovado matéria suficiente para a punição. Já a ERC pode intervir aquando de queixas sobre conteúdos inaceitáveis. De referir que desde 2009, a ERC recebeu 21 queixas formais (ECRI, 2018: 22).

Igualmente, é recomendado às autoridades portuguesas que a imprensa tradicional, e as plataformas digitais, proíbam o discurso de ódio e que mantenham esses conteúdos inacessíveis. Igualmente, é encorajado a formação dos agentes envolvidos no processo sobre questões de prevenção e combate ao discurso de ódio, tanto para jornalistas, como para quem é responsável por monitorizar sítios da Internet e redes sociais. Finalmente, a ERC deve ter

³⁸ *The Digital Competence Framework for Citizens*, no original.

³⁹ <https://ec.europa.eu/jrc/en>

⁴⁰ <https://www.adl.org/>

⁴¹ <https://www.coe.int/en/web/no-hate-campaign>

⁴² <http://mediasmarts.ca/>

uma ação determinante em casos comprovados de publicação de conteúdos inaceitáveis e com aplicação de sanções apropriadas (ECRI, 2018: 22). Nas organizações de utilizadores de plataformas digitais com uma atenção especial aos processos de regulação da Internet, destaca-se em Portugal a Associação 3D⁴³ com o seu trabalho realizado na informação e proteção dos cidadãos digitais. Esta organização é responsável por ações como, por exemplo, a entrega de uma queixa à Comissão Europeia sobre retenção indiscriminada de metadados⁴⁴, o direito à navegação anónima na Internet⁴⁵, ou com pedidos de esclarecimento sobre partilha de informação à margem da lei⁴⁶.

6.2 Melhorar a “arena de discussão” *online*

O diálogo está no centro de uma democracia forte, não apenas como a expressão de uma ideia e de um argumento, mas como meio para a interação humana. O processo político “ossificar-se-ia sem a criatividade, a variedade, a abertura e flexibilidade, a sua capacidade para a invenção, a sua capacidade para a descoberta, a sua subtilidade e complexidade, a sua eloquência, o seu potencial para empatia e expressão afetiva, e o seu carácter para dóxico (alguns podem dizer mesmo dialético) que ilustra a natureza completa do humano como um agente ativo, com propósito e interdependente” (Barber, 2003: 174). O diálogo numa democracia forte implica haver uma igual medida de atenção ao que é dito, e à forma de expressão. Essa dialética resulta da passagem do domínio da reflexão pura para o mundo da ação (Barber, 2003: 174). Porém, para haver essa ação é necessário que estejam asseguradas as condições para a liberdade de expressão, e que propósito, e independência, não sejam diminuídos. A não ser que incorram no princípio de dano.

Com o aparecimento de cada vez mais países com tendências iliberais (Wike, Silver & Castillo, 2019) o princípio da escala deslizante determina que é de esperar mais sistemas de controlo sobre discurso público e produção de conteúdos. O equilíbrio entre controlo e liberdade tem de ser afinado a todo o momento. A relação entre instituições governamentais, supragovernamentais, grupos tecnológicos e organizações não governativas, é um ponto de partida para se avançar na resolução desta necessidade. Grupos como accessnow.org⁴⁷, European Digital Rights⁴⁸, digitalcourage⁴⁹, ~~privacy international~~⁵⁰ (*sic*) são exemplos de

⁴³ <https://www.direitosdigitais.pt/>

⁴⁴ https://stopdataretention.eu/stop_data_retention_open_letter.pdf

⁴⁵ http://direitosdigitais.pt/media/ficheiros/russia%20telegram_statement.pdf

⁴⁶ <https://privacyinternational.org/sites/default/files/2018-04/Privacy%20International%20Briefing%20-%20Intelligence%20Sharing%20FINAL.pdf>

⁴⁷ <https://www.accessnow.org/>

⁴⁸ <https://edri.org/>

⁴⁹ <https://digitalcourage.de/en>

⁵⁰ <https://privacyinternational.org/>

organizações que expõem o crescente controlo governamental e de empresas privadas, em prol da proteção dos utilizadores. A arena de discussão política, diferente de outras arenas de discussão populares no *zeitgeist*, tem de ser um espaço onde impere a civilidade, num exercício consentâneo com cidadania. No caso de uma troca robusta de argumentos e ideias, civilidade é mais do que apenas a aplicação de boas maneiras entre pessoas pertencentes à mesma comunidade. No dicionário português da Priberam lê-se sobre civilidade como “modo de se corresponderem pessoas bem-educadas, cortesia, etiqueta, respeito, dignidade, tolerância”⁵¹. No entanto, existe a questão de saber até onde é preciso ser-se tolerante na proteção da civilidade. Um dos resultados da emergência do politicamente correto, é a atitude, natural ou forçada, de haver tolerância contínua, e muitas vezes excessiva. Karl Popper apresenta o que é o “paradoxo da tolerância”, com o aviso que a tolerância sem limites levará ao desaparecimento da tolerância (Popper, 1966). A procura de tolerância até para os intolerantes, não deve ser reclamada como um direito, nem pedida como uma demonstração de virtude. A tolerância é algo que deve reinar quando se trata de permitir argumentos e opiniões que ajudam a comunidade na perspetiva Madisoniana, de boa governação, de tomadas de decisão em conjunto, de robustez de debate. Ser-se tolerante “só porque sim”, é confundir a arena de discussão e permitir excessos na argumentação e na defesa da opinião, que não são produtivos nem contribuem para o progresso societal.

É uma questão em aberto se fortalecer a arena de discussão pode ser um mecanismo de criação de novos limites para a liberdade de expressão. Discurso racista, xenófobo ou misógino é um entrave para uma coexistência salutar em sociedade. Como tal, deve ser proibido? Numa perspetiva maximalista de liberdade de expressão, a resposta é “não”. É, no entanto, legítimo pedir que esse tipo de expressão seja regulado, no lugar de ser censurado. Ataques verbais a características como raça, género, doença, deficiência devem ser considerados como argumentos e/ou conteúdos que não merecem proteção. Argumentos e opiniões com discurso que cause ofensa, ou até de ódio sem causar dano, devem ser protegidos por lei. Argumentos em defesa da eutanásia, da interrupção voluntária da gravidez ou da igualdade de minorias sexuais podem ser considerados, particularmente por pessoas com fortes convicções religiosas, como mais do que simplesmente ofensivos, passando a ser odiosos. No entanto, isso não deve ser uma razão para se suprimirem tais opiniões ou argumentos.

Noutro exemplo, relativamente à controvérsia mencionada anteriormente da publicação dos *cartoons* com o Profeta Maomé pelo jornal Jyllands-Posten, Timothy Ash

⁵¹ <https://dicionario.priberam.org/civilidade>

sugere que arte e humor sejam vistos como transgressões aceitáveis à norma. Porém, e como sugere o autor, isso depende do conceito de “abertura”, ou seja, a capacidade de lidar com argumentos e opiniões sem preconceitos à partida, confrontando discurso e/ou conteúdos com que não concordamos, ou até que nos possam ofender, mas com uma atitude civilizada. Esta abertura, é na opinião de Ash, muito importante pois, “[a]bertura sem civilidade pode levar à anarquia, mas civilidade sem abertura é uma receita para perda de liberdade (Ash, 2016: 211).

Outro potente regulador da arena de discussão, tanto no mundo real como digital, é a desaprovação social. Apesar de não ter o poder de impedir discurso de ódio, dano ou ofensivo, pode criar um ambiente onde se paga um preço elevado pelo exercício da liberdade de expressão. Mill partilha dessa hipótese, preocupando-se inclusive com a possibilidade de a desaprovação causar consensos onde “todos vivem como sujeitos a olho de hostilidade e da temida censura” (Mill, 1859: 58). Porém, o polemista cede algum terreno ao contemporizar que o princípio da ofensa pode ser aplicado em situações de desaprovação social: “[n]ovamente, existem muitos atos que, sendo diretamente injuriosos apenas para os agentes, não devem ser legalmente proibidos, mas que, se feitos publicamente, são uma violação de boas maneiras e, sendo incluído na categoria de ofensa contra outros, pode ser, justamente, proibido” (Mill, 1859: 97). Isso aplicar-se-á àqueles que demonstrem atitudes de crueldade, malícia, ressentimento, egoísmo, inveja ou insinceridade, e que assim colocam-se suscetíveis a uma desaprovação social, com isolamento e ostracização. Mesmo assim, continua o debate em aberto, uma vez que algumas destas atitudes menos civilizadas, não implica, necessariamente, que resulte em violação de outros direitos, escapando assim ao princípio de dano, mas mantendo o princípio da ofensa.

Finalmente, é importante analisar as dinâmicas que se estabelecem entre os diferentes intervenientes na arena de discussão. Exemplos são as dinâmicas utilizador-utilizador, utilizador-intermediário digital, relação utilizador-plataforma digital (não se restringindo só às redes sociais) é finalmente a relação plataforma-utilizador-autoridades. Pode-se imaginar estas dinâmicas num contínuo. Como podem utilizadores lesados expressarem as suas queixas e estas serem corrigidas diretamente com quem está a cometer a infração? No caso de isso não for suficiente, o utilizador poder recorrer ao intermediário digital para estes contactarem o infrator e pedir para que o comportamento seja alterado. De dizer que no Facebook e no Twitter isso é possível, tanto pelo utilizador como pela plataforma. Num último nível, serão envolvidos agentes de autoridade e/ou jurídicos para a resolução de problemas, preferencialmente sem ser por meios punitivos.

6.3 Alimentar o mercado de ideias

O conceito de existir um mercado de ideias foi expresso pela primeira vez numa opinião legalista sobre limites da liberdade de expressão pelo Juiz Oliver Wendell Holmes, na opinião dissidente à decisão por maioria do Supremo Tribunal Americano no caso, referido anteriormente, de *Abrams v. United States* (USSC, 1919). Em particular, o Juiz Holmes referiu à necessidade de existir uma “troca livre de ideias” e uma “competição” entre as mesmas, como de um mercado se tratasse. Seria mais tarde, em 1953, que o Juiz Justice William O. Douglas usaria a expressão mercado de ideias⁵², no enquadramento do acórdão da maioria do Supremo Tribunal na decisão de *United States v. Rumely*: “[t]al como os autores de jornais, revistas ou livros, este autor licita para as mentes do homem no mercado de ideias” (USSC, 1953). Este “mercado” pode ser visto no sentido figurado de não haver interferência do Estado ou de reguladores no controlo de discursos e/ou conteúdos, pois tal irá “distorcer” o mercado, tirando a capacidade do indivíduo, neste caso do utilizador de plataformas digitais, de fazer uma apreciação independente de ideias e argumentos em debate (Hudson, 2018). Porém, para autores como Cass Sunstein esta expressão não capta perfeitamente o que deve ser a troca de ideias e argumentos. Para o autor, a metáfora “produziu um bem considerável, mas também é enganadora. Noções de mercado, ou de agregação, ignoram a extensão de quanto os resultados políticos são supostos serem resultados de discussão e debate, igualdade política e razões apresentadas a favor ou contra as alternativas” (Sunstein, 1995: 249). Muito antes de a expressão existir, já John Milton, na obra *Aeropagítica*, proponha algo similar sugerindo que: “[d]eixemos a sua [verdade] e a falsidade chocarem; quem alguma vez conheceu a verdade entregue ao pior num encontro livre e aberto? A sua refutação é a melhor e mais segura supressão” (Milton, 1644: 37).

O conceito de que a opinião dos outros deve ser ouvida, e inclusive ser alvo de uma proteção especial, foi expressa, pela primeira vez com grande efeito, por Voltaire. Já ficou definido neste trabalho que a autoria da frase não é do filósofo humanista, mas o conceito de que “eu desaprovo aquilo que disse, mas defenderei até à morte o seu direito de o dizer”⁵³ passou a fazer escola entre defensores e absolutistas da liberdade de expressão. Também John Stuart Mill ecoou pensamento igual, ao escrever que “[s]e toda a humanidade, menos um, fosse da mesma opinião, e se uma pessoa fosse de uma opinião contrária, a humanidade não estaria mais justificada em silenciar essa pessoa, do que essa pessoa, se tivesse o poder, em silenciar a humanidade” (Mill, 1859: 18), e que “o peculiar mal de silenciar a expressão de

⁵² *Marketplace of ideas*, no original.

⁵³ *Je ne suis pas d'accord avec ce que vous dites, mais je me battrai jusqu'à la mort pour que vous ayez le droit de le dire*, no original.

uma opinião, isso é roubar à raça humana: posterioridade e gerações existentes, daqueles que divergem da opinião, ainda mais do que aqueles que a detêm” (Mill, 1859: 19). De facto, suprimir-se uma opinião, porque se assume que esta é errada, que vai contra um consenso estabelecido, ou que causa ofensa, é assumir que a nossa versão da verdade, do nosso consenso, da nossa sensibilidade é, à partida, mais correta do que a dos outros. Isto conduz ao que Mill descreve como a assunção de infalibilidade, que condiciona, ou até mesmo pode cancelar uma troca salutar de ideias e argumentos. Há uma diferença perniciosa entre ter a certeza da validade de uma ideia ou argumento, por causa da ausência de uma melhor ideia ou argumento, ou assumir que não existe nenhuma outra posição que possa obrigar a repensar a nossa posição. O processo de refinamento de ideias e argumentos, que depois se traduz na produção de discurso e/ou conteúdos, deve ser contínuo, por muito que esse processo seja incomodativo. Mill sugere, inclusive, que a validade de uma ideia ou de um argumento, numa discussão livre sobre ideias e argumentos, seja “levado ao extremo” (Mill, 1859: 23), uma vez que, se as razões apresentadas para a diferença de opinião não são boas o suficiente para serem usadas em situações extremas, não são boas para serem usadas em qualquer caso. Este é um argumento que, seguramente, é bem aceite pelos absolutistas da liberdade de expressão.

A liberdade de expressar ideias e/ou argumentos, ou de produzir discurso e/ou conteúdos deve ser pensada, particularmente no debate político, como um exercício de onde saem argumentos com mais ou menos valor, na procura de consensos ou de desafio de ideias “estabelecidas” se estas não têm mérito. Desse processo resulta o desenvolvimento de conhecimento político, produzido com enquadramento histórico e pelo acumular de experiências. Todos esses ganhos resultam na aplicação de ações conjuntas para benefícios futuros. Existem questões que se colocam quando se avalia estes benefícios: “o que devemos fazer?”; “como devemos reconciliar as nossas diferenças?”; “como devemos nos comportar como membros de uma comunidade justa?”. Para estas questões não existem respostas como “verdadeiro” ou “falso”, ou posições corretas ou incorretas. O que existem são visões alternativas que competem para uma aceitação comunitária (Barber, 2003: 169). Uma vez isso aceite, devem ser feitas três perguntas “[o] que se vai fazer? Qual a intenção? Há mais a ganhar ou a perder ao diminuir-se a liberdade de expressão?” (Fish, 1994: 127).

Cass Sunstein sugere que produzir discurso e/ou conteúdos tem semelhanças com o mercado económico por causa das relações que se constroem entre custo/retorno, valor/não valor, vantagens/desvantagens. Dessa forma é possível uma melhor regulação da liberdade de expressão (Sunstein, 1995: 17). Por exemplo, quem exerce a sua liberdade de expressão pode ser visto como um produtor de discurso e/ou conteúdos que detêm um conjunto de

características: qual o trabalho necessário, e quanto tempo e atenção é dado a essa produção? Qual a expectativa de chegar a outras pessoas? Qual a extensão de alcance desejado? Qual o risco de a mensagem ser recebida de uma forma errada, e com isso haver um preço elevado a pagar pela produção do discurso e/ou conteúdo? Qual a eventualidade desse contributo sofrer desaprovação social ou de marginalização da mensagem ou até mesmo de quem a emite? Porém, condições à partida precisam de ser asseguradas: igualdade no acesso e na difusão, na avaliação da validade, na definição da importância e contributo do discurso e/ou conteúdo, o ceticismo para com argumentos de autoridade ou a rejeição de tentativas de desvirtuar, ou manipular, o fluxo de informação. O princípio Madisoniano deve resultar no criar de uma atenção alargada por parte de um número substancial de cidadãos, com uma abordagem detalhada de questões políticas que são de interesse para a comunidade, e com uma exposição ao público de uma diversidade de pontos de vista, desde que apropriados à discussão (Sunstein, 1995: 20-21). Um mercado de ideias pode ser visto como um sistema desregulado, usando um conceito do liberalismo clássico, de existir uma “mão invisível”, na lógica de John Smith, de uma política de *laissez-faire* que moderará o discurso na procura de uma deliberação democrática (Sunstein, 1995: 18). Essa é uma posição que será mais próxima dos absolutistas da liberdade de expressão, mas cada vez menos provável devido a cada vez mais testes conservadores para discurso de ofensa e dano.

6.4 Qual a melhor regulação?

Em 2019, a Comissão Europeia produziu um relatório sobre códigos de conduta de discurso de natureza ilegal em plataformas digitais, com a participação voluntária do Facebook, Twitter e YouTube. Para além de um conjunto de sugestões, que iremos referir a seguir, o relatório revela alguns desenvolvimentos positivos. A velocidade com que as plataformas retiram discurso e/ou conteúdos de natureza ilegal aumentou consideravelmente. Cerca de 89% dos *posts* que tinham sido alvo de pedidos de remoção, foram retirados num período de 24 horas. Desses, 72% tinham sido caracterizados, corretamente, como material ilegal (European Commission, 2019). O facto de 18% terem sido retirados indevidamente pode causar uma de duas reações: ou é um preço aceitável a pagar, ou então é uma margem de erro intolerável. Qual a percentagem aceitável é uma contenda entre absolutistas e os que defendem testes mais restritivos. A outra preocupação importante, apresentada pela Comissão Europeia, é a falta de transparência em processos decisórios para avaliar o que é, ou não, passível de ser retirado pelas plataformas digitais mediante os pedidos de remoção recebidos (European Commission, 2018), algo que tenta ser corrigido com o novo *Ato de Serviços Digitais*.

Um outro problema, com efeito na qualidade da discussão política *online* e com consequência o aumento de comportamentos aberrantes, ofensivos e causadores de dano, era a possibilidade de estar no mundo virtual de uma forma anónima. Inicialmente, as redes sociais, não antecipando o seu eventual uso para fóruns de discussão de ideias e de opiniões, permitiam a criação de contas *online* e de perfil de utilizador sem necessidade de ser recolhida informação relativa ao mundo real. Em 2011 era feito o aviso que a facilidade na criação de discurso anónimo, ou sob a utilização de pseudónimos, podia acelerar o comportamento destrutivo do discurso público (Citron & Norton, 2011). Como descrito, aquando da apresentação da proposta de lei GSVimNetz da Áustria, que defende a necessidade de criar *Rigor e Responsabilidade na Internet*, uma das medidas é deixar de haver anonimato quando se deixa um comentário *online*. Os utilizadores têm de providenciar à plataforma nome completo e morada (Al-Youssef, 2019). As empresas digitais têm avançado para a definição de condições na criação de contas, como é o caso de endereços de e-mails reais, preenchimento de campos obrigatórios de informação fácil de confirmar (moradas, números de telefones), sendo até proposto um sistema de *number plates* que tornam obrigatória um conjunto de informação fidedigna na identificação de utilizador para este aceder a fóruns de discussão e redes sociais.

O Twitter deu um passo importante com o que se denomina *blue check marks*, uma forma de autenticar o utilizador e afirmar que a conta é de interesse público (Twitter, 2019a). A organização não governativa alemã Sociedade para os Direitos Civis⁵⁴ contrapõe que o final do anonimato não impede utilizadores de cometerem atos ilegais. Uma das razões apresentadas para este paradoxo é existir, por parte das autoridades, ou um interesse reduzido ou uma incapacidade de ação. Isso causa, no utilizador, uma sensação de impunidade que o leva a repetir os comportamentos (Gesellschaft für Freiheitsrechte, 2019). Para um processo de punição, ou de ressarcimento, é preciso criar um sistema célere, independente e justo, para a análise de queixas, persecução de criminosos e aplicação de penas. Uma sugestão é que esse processo seja da exclusiva responsabilidade de tribunais independentes, no lugar de ser deixado a instituições governamentais ou às plataformas digitais (Gesellschaft für Freiheitsrechte, 2019).

A possibilidade de utilizadores privados, instituições governamentais ou empresariais poderem aceder a dados pessoais, como resultado de uma reclamação ou queixa, causa também o risco de deixar de haver denúncias anónimas, se, quem as faz não consegue proteger a sua identidade. Fica assim em risco a existência de *whistleblowers*. Expor más

⁵⁴ Gesellschaft für Freiheitsrechte, no original.

práticas, mas sem tornar pública a identidade e a posição, fica comprometido se não existir proteções atribuídas por leis. Curiosamente, estas propostas de termino de anonimato coincidem com os primeiros passos concretos por parte da União Europeia, através do Parlamento Europeu e do Conselho da União Europeia, de desenvolver legislação para proteger *whistleblowers* na Europa (TIEU, 2019).

Outro sistema de controlo é a autorregulação. No entanto, isso não significa chegar ao ponto de inibir a participação no debate *online*. É necessário haver a “coragem de convicções”, se o argumento for válido e dentro das regras do debate, de não abandonar uma ideia ou uma opinião por não ser popular. Haverá sempre quem aprecie um ponto de vista diferente ou uma ideia provocadora. No entanto, é necessário mediar o debate, no caso de a autorregulação ser inexistente para prevenir que limites sejam ultrapassados. Essa é a função de moderadores. Algumas plataformas já instituíram o modelo clássico de moderação, como é o caso do Redditt e Kialo⁵⁵ empossando moderadores, por experiência ou legitimidade, para exercerem a fiscalização na troca de argumentos. O mesmo ainda não acontece no Twitter e no Facebook, a primeira por questões de arquitetura da plataforma, e a segunda por decisão organizacional.

Outro mecanismo à disposição de instituições nacionais ou supranacionais são ações de defesa da concorrência (*antitrust*)⁵⁶. No momento da elaboração deste documento, os Procuradores-gerais dos cinquenta estados dos Estados Unidos da América estão em conversações com o Departamento de Justiça para a eventualidade de serem iniciadas investigações *antitrust* contra o gigante digital do Google, com a justificação que essa plataforma “domina todos os aspetos de publicidade e de motor de busca na Internet” (Kelly, 2019). Igualmente, a Procuradora-geral do Estado de Nova York anunciou uma investigação similar com o Facebook, com a justificação que a empresa “coloca em risco os dados dos consumidores, reduz a qualidade de escolha dos consumidores, e aumenta o preço de anúncios publicitários” (Lecher, 2019). Juntando-se à Procuradora-geral de Nova Iorque os Procuradores-gerais dos Estados de Colorado, Flórida, Iowa, Nebraska, Carolina do Norte, Ohio, Tennessee e o Distrito de Columbia. Também na Europa, legisladores, principalmente no Parlamento Europeu, têm referido a possibilidade de leis *antitrust* serem aplicadas na área digital. No entanto, no momento da elaboração deste documento, a Comissária Europeia para a Concorrência, que acumula a posição de Vice-presidente Executiva da Comissão Europeia, Margrethe Vestager, com o *portfolio* da *Europa Pronta para a Era Digital* fez saber que se

⁵⁵ <https://www.kialo.com/>

⁵⁶ *Antitrust laws*, no original.

deve contemplar tal tipo de iniciativa apenas como uma solução extrema, privilegiando mecanismos de entendimento que envolvam todos os *stakeholders* (Lomas, 2019a).

6.4.1 Regulação de algoritmos

A União Europeia terá como uma das áreas de intervenção, uma vez aprovado o *Ato de Serviços Digitais*, a supervisão da mecânica das plataformas digitais, principalmente em como operam os algoritmos. A pergunta mais importante para responder nos próximos anos, à medida que se caminha para a substituição da *Diretiva e-Comércio de 2000/31/EC*, é que filtros podem ser implementados em algoritmos para um controlo de qualidade de discurso e/ou conteúdos, principalmente antes de estes aparecem *online*. Tem sido a posição, tanto pela Comissão Europeia em funções antes das eleições Europeias de 2019, como a Comissão constituída no verão do mesmo ano, que as plataformas digitais de redes sociais não serão forçadas a monitorizar, de uma forma ativa, o tipo de material colocado pelos utilizadores. Essa parece ser uma base fundamental da regulação da Internet: a existência de um espaço para todos terem direito à liberdade de expressão, e de comunicação, sem a necessidade de passar por “guardiões de informação” (EDRi, 2019a). Porém, também existem movimentações, algumas apresentadas anteriormente, em Estados-membro, e também em Bruxelas, para ser exigido às empresas digitais que automatizem um controlo prévio.

Em 2018, num documento da Comissão Europeia com recomendações sobre como efetivamente combater material ilegal *online*, pode-se ler que “em certos casos, tal solução [filtros automáticos de controlo pré-publicação] pode ser um elemento importante para combater esse tipo situação” (European Commission, 2018a: 6), sem prejuízo do artigo 15.º(1) da *Diretiva e-Comércio de 2000/31/EC*, que se refere à “ausência de obrigação geral de vigilância” (EPC, 2000). Porém, esta posição abre a porta para empresas digitais argumento que, pelo número de utilizadores nas plataformas, e pela quantidade de publicações que acolhem, é impossível cumprirem com essa exigência. Isso abre a possibilidade de uma automatização completa do processo de publicação de *posts*. Uma solução sugerida pela Comissão Europeia em funções antes de 2019, e transferida para o *Ato de Serviços Digitais*, é a possibilidade da automatização ser efetuada quando o “caracter ilegal do conteúdo tenha sido já estabelecido, ou onde o tipo de conteúdo é tal que a sua contextualização não é essencial” (European Commission, 2018a: 6). É de esperar que as empresas exijam continuar a beneficiar de um *safe harbour*, para as ilibar de responsabilidades legais no caso de chegarmos a um tempo onde algoritmos sejam os “guardiões” de conteúdos, à priori ou a posteriori. Uma comunicação da Comissão Europeia, em 28 de setembro de 2017, sobre *Como Combater Conteúdos Ilegais Online. A Caminho de uma Responsabilidade Aumentada*

das *Plataformas Online*, já apontava para isso (European Commission, 2017: 10), onde está expresso que medidas proativas de controlo de discurso e/ou conteúdos não levarão as plataformas a perderem os benefícios das exceções proporcionadas no artigo 14.º da *Diretiva e-Comércio de 2000/31/EC* (EPC, 2000). No *Ato de Serviços Digitais*, essa proteção mantém-se, acrescentando as alterações apresentadas anteriormente.

O *Ato de Serviços Digitais* menciona a necessidade da avaliação da transparência no processo de como os algoritmos produzem recomendações para as páginas pessoais dos utilizadores. Tal avaliação não é desejada pelas empresas com plataformas digitais de redes sociais. Será necessário esperar para ver o quanto as intenções políticas se traduzem em regulação, e que critérios serão usados nos códigos digitais, e como estes decidem o que deve, ou não, ser censurado (HoL, 2015: 21). Organizações que fazem *lobbying* na União Europeia em nome de empresas digitais, como é o caso da CCIA Europe⁵⁷ e da EDiMA⁵⁸, têm expresso preocupação com eventuais obrigações das plataformas de redes sociais terem de criar filtros de *upload*⁵⁹ para bloquear material ilegal (Kayali, 2019). Na opinião destas organizações, este é um passo na direção de ser implementada censura prévia, aplicando-se novamente o princípio da escala deslizante. EDiMa, por exemplo, numa declaração pela sua diretor-geral, fez saber que “filtros de *upload* e direitos de vizinhança⁶⁰ irão resultar em censura da livre expressão *online* e, à desintegração da Internet como a conhecemos” (EdiMA, 2018). Se o caminho anunciado é a tentativa de um reforço de sistemas de controlo de discurso e/ou conteúdos à base de filtros de *upload*, levanta-se outro problema, que é o volume de material que existe e que é carregado para as plataformas digitais. Por exemplo, no YouTube, são carregados, a cada minuto, mais de 500 minutos de vídeos para os servidores do Google. Até mesmo os melhores *softwares* a nível de controlo de conteúdos, como é o caso do YouTube ContentID, gera muitos erros, e são necessários milhares de moderadores humanos acrescentados ao sistema para um controlo suficiente do que não deve ser publicado (Fanta & Rudl, 2019).

6.4.2 Regulação da arquitetura digital

A arquitetura digital desenvolvida para funcionamento das plataformas digitais está diretamente associada a qual o modelo económico da respetiva empresa. No caso do Facebook, o que vulgarmente se chama do *feed* de notícias, parte principal do *front page* da plataforma, as publicações são determinadas por algoritmos. Porém, o objetivo é amplificar

⁵⁷ <https://www.ccianet.org/tag/europe/>

⁵⁸ <http://edima-eu.org/>

⁵⁹ Carregamento de informação digital, traduzido do original

⁶⁰ Direitos de trabalho criativo que não estão diretamente relacionados com o trabalho do atual autor.

discurso e conteúdos que aumentam a taxa de envolvimento do utilizador, causando que fique mais tempo na plataforma e aumentando assim o tempo de exposição a publicidades (Evans, 2019). Esta é a forma como o Facebook assegura o seu poderio económico (Macrotrends, 2019). Porém, esta arquitetura digital não favorece a troca de conteúdos com outros utilizadores que tenham pontos de vista e argumentos diferentes, reunindo-se as condições para o desenvolvimento das bolhas ideológicas, camaras de eco e polarização societal. Argumentos contrários a convicções à partida, não são assunto que despertem interesse ao utilizador, o que faz com que não interaja com o mesmo, o que diminui a probabilidade de tornar a aparecer na sua página pessoal. Este problema é exacerbado pela política do Facebook de todos os conteúdos têm o mesmo valor informativo, fazendo com que os algoritmos sejam “conteúdo-agnósticos” (Evans, 2019). Isso passa para o utilizador a responsabilidade de avaliar qual a qualidade da informação que aparece no seu *feed*, assim como qual a veracidade nas publicidades políticas, o que conduz a mais desinformação e polarização.

Um outro exemplo de interação entre arquitetura digital e a propagação de material prejudicial é o YouTube. A forma como a plataforma está construída, principalmente no que se trata de sugestões associadas ao tema do último vídeo observado, sejam sugestões no final do vídeo ou na barra lateral, têm como objetivo a manutenção do utilizador na plataforma o máximo de tempo possível. Junta-se o facto adicional de quanto mais o vídeo for apelativo, mais vezes este aparecerá como recomendação, novamente sem uma avaliação da veracidade ou natureza do discurso que nele é apresentado. Por causa disso, a plataforma tem estado sob um apertado cerco de críticas devido à propagação de desinformação e criação de polarização e extremismo (Friedersdorf, 2018). YouTube teve de anunciar, no início de 2019, que começaria a “reduzir recomendações de conteúdos marginais e conteúdos que possam ter informação que crie dano aos utilizadores” (YouTube, 2019). Exemplos são “curas milagrosas”, “opiniões que a terra é plana” ou “falsidades claras sobre eventos históricos tal como o 9/11” (YouTube, 2019). No entanto, apesar do compromisso pelo gigante digital, no início de 2020, o grupo AVAAZ⁶¹ detetou, numa amostra de 5,537 vídeos sobre “alterações climáticas”, “aquecimento global” e “manipulação climática” que 8% dos vídeos no primeiro tema, 16% no segundo, e 21% no terceiro, continham informação que vão contra o consenso científico sobre alterações climáticas e relação com atividade humana (AVAAZ, 2020: 11).

Outro ponto de interesse em arquiteturas de plataformas é qual o debate público, propriamente dito, que acontece. Conversações no Twitter são normalmente espontâneas,

⁶¹ <https://www.avaaz.org/page/en/>

mas dependentes do assunto e da renovação de temas. A troca de opiniões e argumentos nesta plataforma é muito volátil devido à forma como os tópicos em tendência mudam rapidamente, inclusive dependentes da região e das áreas de interesse no Twitter. O Facebook não privilegia tanto o imediatismo da interação. Os *posts* podem ter trocas de mensagens nas caixas de comentários que se alongam por muito mais tempo, tornando mais difícil o controlo de discurso ofensivo, de ódio ou de dano. Finalmente, outra diferença interessante entre a arquitetura das duas maiores plataformas de redes sociais, é a recusa pelo Twitter, até à data da elaboração deste documento, em criar a opção de edição de *tweets*, seja para corrigir um erro ortográfico ou de uma opinião política. Para os responsáveis da plataforma, uma vez havendo respostas ao *tweet*, seja em *likes*, *retweets* ou comentários, o *tweet* original não pode ser alterado (apesar de poder ser apagado). A nível de arquitetura tornar-se-ia muito complicado para o Twitter rever os algoritmos para atualização dos conteúdos: a correção de um *tweet* pode desvirtuar, ou inverter, a mensagem inicial e como tal comprometer todas as interações precedentes, como *likes* ou *retweets*. Já o Facebook não parece particularmente preocupado com este tipo de precisão.

Capítulo 7. Sistematização dos mecanismos para proteção de discurso e/ou conteúdos políticos *online*

Direitos adquiridos como liberdade de expressão, privacidade, possibilidade de produzir, transmitir e receber informação, na ubiquidade do mundo digital, estão expostos a novos condicionalismos. Tal pode acontecer por ação do Estado na ânsia de proteger os cidadãos numa atitude paternalista, ou pelos gigantes digitais, por política interna ou como resposta a imposições aplicadas por legisladores. Depois, existem ameaças que resultam da evolução do ecossistema da Internet: velocidade da inovação, facilidade da utilização e capacidade do utilizador de navegar em fóruns públicos cada vez mais intolerantes e polarizados. A possibilidade, bem real, das empresas intermediárias digitais terem de lidar com censura generalizada, sem base legal ou enquadramento legislativo, deve preocupar qualquer utilizador, mesmo os defensores de testes restritivos à liberdade de expressão. Adiciona-se a este risco a prática de vigilância maciça, de tratamento de metadados, de aplicação de inteligência artificial e mecanização algorítmica. Estão reunidas as condições onde, autocensura, impossibilidade de publicar opiniões e argumentos, controlo do discurso e dos conteúdos, dificuldade em acionar mecanismos de revisão de queixas ou penas, não só serão coisas habituais, mas aceites com normalidade.

As leis, normas e regras tradicionais, no sentido de terem sido idealizadas na era pré-digital e para resolverem questões de comportamento comunal, não se aplicam adequadamente ao mundo digital. Ao mesmo tempo, há a probabilidade elevada que, à medida que a tecnologia progride, mais desatualizadas fiquem. Soluções técnicas como transferência instantânea de informação entre países, *upload* de dados em sistemas de *cloud*, uso generalizado de telefones inteligentes com acesso a redes cada vez mais poderosas, são agora facilmente acessíveis. Isso tem proporcionado que governos e empresas privadas desenvolvam protocolos para rastreio, censura de discurso e/ou conteúdos, bloqueio ou redirecção de sistemas de nomes de domínio⁶² e de endereços de Internet, negação de serviço⁶³, remoção de informação em motores de busca, proibição de visionamento de sítios da Internet, entre outros. Cabe aos utilizadores e organizações que são *stakeholders* do mundo digital exercer contrarrespostas a esta erosão do exercício de liberdades básicas *online*. Assim, existem áreas que são de charneira para a construção de recomendações para a manutenção da Internet como um espaço de liberdade, troca de ideias e argumentos e deliberação democrática.

⁶² *Domain Name Systems*, no original.

⁶³ *Denial of Service Attacks*, ou DDoS, no original.

7.1 Leis e o sistema judicial

Legislação nacional e internacional deve ser independente de influências para aproveitamento político e/ou comercial, e não devem causar atos discriminatórios ou arbitrários. Quem seja acusado de transgressões deve ter um mecanismo para reclamação de uma queixa, determinando-se à *posteriori* a validade da reclamação. Principalmente, devem existir formas para que processos legais sejam acionados. Esta necessidade é particularmente importante quando a construção de quadros normativo é atribuída às empresas digitais. O procedimento legal que serviu, satisfatoriamente, até ao início da era digital consistia em ser iniciado um processo de tribunal quando uma potencial vítima de excesso de liberdade de expressão se sentia prejudicada. Seriam então apresentados os argumentos da acusação e da defesa, o que permitia determinar a justiça dos factos e ilibar-se os inocentes ou punir-se os culpados. Se ficasse provado que tinha sido efetuado dano, a pessoa alvo desse discurso e/ou conteúdo nefasto era ressarcida através de uma compensação financeira, um pedido de desculpas, uma retração pública. Porém, as pessoas não eram “banidas” de poderem exercer a sua liberdade de expressão. Outro aspeto positivo desse processo era a (desejada) melhoria do debate, não só por correção do discurso, mas também por mudança na atitude dos participantes. Corre-se o risco de se perder esse mecanismo por causa da opacidade com que as empresas digitais exercem censura na publicação de discurso e/ou conteúdos, e pela pressão para que leis sejam mais restritivas no valor e da qualidade dos materiais colocados *online*.

Quando, neste momento, a esmagadora maioria dos utilizadores *online* produz a quase totalidade do seu discurso de natureza política em plataformas geridas por apenas quatro empresas, há a preocupação da aglomeração de poder. Poder esse que se traduz em influência nos processos, nos utilizadores, nos legisladores, e em última análise, na liberdade de informação e de expressão. Isto é particularmente importante na Europa onde as grandes empresas digitais usam testes adaptados à cultura anglo-saxónica, muitas vezes desajustada para diferenças entre países. Empresas como a Alphabet, que detém os gigantes Google e YouTube, e as plataformas de redes sociais como o Facebook, o Twitter e o Reddit, preocupam-se principalmente com o mercado interno, i.e., os Estados Unidos da América. Apesar destas empresas terem delegações para influenciar os centros de poder na Europa, esse mercado acaba por ser secundário em relação ao Americano. Para além disso, a Ásia, Euro Ásia e África são mercados em crescimento e onde não existe uma ação tão premente por parte de legisladores para controlo das plataformas digitais. Quando se pensa numa organização supranacional como a União Europeia, uma tendência para a uniformização de

leis e normas continuará a sofrer de existirem diferenças locais que complicarão esse tipo de legislação comunitária.

Outro passo para uma liberalização da arena de discussão é alterar leis de difamação. A livre expressão pode ser restrita quando necessário, desde que de uma forma proporcional ao interesse da proteção da reputação e da privacidade (CD&T, 2011: 2). O direito à reputação implica a possibilidade de serem efetuadas queixas por difamação. Isso, por sua vez, obriga a existirem mecanismos de proteção legal que possam ser acionados. No entanto, estes podem diminuir o direito à livre expressão. Estas leis podem ser utilizadas por instituições com o intuito de proteger interesses políticos e/ou económicos, silenciando vozes dissidentes (Article XIX, 2009: 1). No relatório do International Press Institute de 2014, é apresentado que as leis contra a difamação na União Europeia ameaçam a liberdade de expressão, principalmente quando envolve figuras públicas e, especialmente, com membros de aparelhos governativos (IPI, 2015). Uma solução para este condicionalismo é proposta no *Relatório do Relator Especial Frank La Rue* de 2011, para o Conselho dos Direitos Humanos das Nações Unidas, sobre a promoção e proteção dos direitos de liberdade de opinião e de expressão. Nesse documento é inequivocamente sugerido que se deve descriminalização o ato de difamação (La Rue, 2011: 11). No mínimo deve haver uma atenuação do peso punitivo de leis anti-difamação e que sejam aumentados os limites dos testes que causam a necessidade de respostas legais.

Novas leis, regras e termos e condições também terão de ser ajustados à crescente na necessidade de remoções indevidas de discurso e/ou conteúdos. A utilização do sistema legal causa, inevitavelmente, a necessidade de tempo e recursos. Se a intenção é não atribuir às empresas digitais a possibilidade de serem juízes em causa própria, e com isso incorrerem em excessos de controlo e de censura, então o sistema judicial terá de se tornar mais ágil na avaliação de queixas sobre controlo da liberdade de expressão. Esta necessidade é ainda mais premente quando as plataformas digitais, se não forem obrigadas a tal, irão resistir a gastar dinheiro a montar sistemas híbridos (humanos/máquinas) para controlar milhões de utilizadores e mil milhões de formas diferentes de discursos e conteúdos. Uma das soluções para resolver a questão do controlo partilhado entre entidades privadas e públicas são os modelos mistos de poder judicial e supervisão legislativa. Isto implica a recalibração do sistema judicial e legislativo para providenciar o julgamento legal correto e otimização do exercício de uns direitos sem comprometer outros (UNESCO, 2016). Definições legais terão de ser atualizadas para questões como o direito à privacidade, a ser-se “esquecido”, ao ónus da prova de difamação, à responsabilidade legal de intermediários, aplicação de filtros ou regulação de algoritmos.

7.2 Plataformas digitais de redes sociais

De maneira a garantir transparência e equidade na prevenção de remoção indevida de discurso e/ou conteúdos, as plataformas digitais devem codificar de uma forma precisa quais os critérios para a remoção, e, se feito indevidamente, como podem ser acionados mecanismos para repor material censurado. Isso permitirá responder a reclamações, com as razões para ter sido acionado o processo de controlo da liberdade de expressão, e como reverter a decisão inicial. Essa codificação é uma condição essencial, não só para avaliação *a posteriori* do tipo de discurso e/ou conteúdos publicado na plataforma, mas também, no caso de existirem filtros de controlo *à priori*. É também necessário acautelar que a robustez do sistema não impeça soluções parcimoniosas quando existem situações de óbvia ilegalidade. Se o exercício da liberdade de expressão se enquadrar na produção de atos ilegais, deve ser acionado o trabalho conjunto com autoridades para a punição dos prevaricadores. Por outro lado, quando for ostensivamente claro que esse não é o caso, também não deve ser pedido, ou esperado, que as plataformas digitais tenham de responder a reclamações que não têm o mérito de ser avaliadas. Para evitar a criação de desconfiança por parte dos utilizadores e dos *stakeholders*, as plataformas digitais devem cultivar uma total transparência no que se relaciona com a partilha de dados com agentes de autoridade, das negociações com o Estado e instituições supranacionais, e com legisladores e reguladores. Estas necessidades são particularmente importantes uma vez que se antecipa que, cada vez mais, o utilizador não só desfrute das plataformas digitais, mas seja um produtor, e um produto, que beneficia a empresa que proporciona o serviço.

Uniformização de termos e condições é igualmente uma solução óbvia. O Yahoo, por exemplo, tem nos termos de serviço que não se pode colocar na plataforma discurso/conteúdos que sejam de dano, ameaçadores, abusivos, de assédio, de tortura, difamatórios, vulgares, obscenos, invasivos de privacidade, de ódio, de teor racial ou étnico⁶⁴. Já o Twitter alerta para utilizadores não poderem expressar conteúdos que possam ser ofensivos, de dano, imprecisos, ou de qualquer forma inapropriados ou em algum caso *posts* que possam ser falsos ou enganadores⁶⁵. O YouTube encoraja liberdade de expressão mesmo de pontos de vista não populares. Porém, não permite discurso de ódio, discurso que ataque um grupo baseado em raça ou origem étnica, religião, deficiência, género, idade, estatuto de ex-combatente, orientação sexual e identidade de género⁶⁶. Quanto ao Facebook, proíbe conteúdos que causem dano, com ameaças, ou que tenham potencial para provocar

⁶⁴ <https://policies.yahoo.com/us/en/yahoo/terms/index.htm>

⁶⁵ <https://twitter.com/pt/tos>

⁶⁶ <https://www.youtube.com/static?template=terms&gl=PT>

ódio ou incitem violência, removendo discurso de ódio baseado em raça, etnicidade, origem nacional, afiliação religiosa, orientação sexual, sexo, género ou identidade de género, com serias deficiências ou doenças⁶⁷. Esta heterogeneidade pode ser um dos pontos de partida, por parte de legisladores, na procura de um conjunto de normas comuns, envolvendo os *stakeholders* e as empresas digitais.

Um modelo de negócio onde o utilizador, e a sua interação com a plataforma digital, é um “bem” no sentido económico do termo, levanta questões cruciais sobre os incentivos financeiros das empresas digitais, e na forma como oferecem e mantêm o serviço. A monetização do tempo de atenção do utilizador, seja pela análise dos gostos, tendências e interesses, ou da exposição a publicidades pagas, mesmo de foro político, faz com que as plataformas digitais, principalmente as redes sociais, pensem na experiência do utilizador não como um fim, mas como um meio. Isto é particularmente observável no caso do YouTube e o Facebook, com estas companhias a afinarem os algoritmos para privilegiar discurso e conteúdos que criam viés negativo para manter o utilizador na plataforma. Estudos científicos (Trussler & Soroka, 2014; Lengauer, Esser & Berganza, 2011) mostram que os seres humanos têm um interesse aumentado por notícias de cariz catastrófico, controverso, polémico e que causam ultraje, irritação e afronta. Por causa disso, algoritmos privilegiam conteúdos polarizantes, o que gera radicalização dos utilizadores (Moskalenko, 2018). Se juntarmos a isto a possibilidade de *microtargeting* político, reforçado pela existência de bolhas ideológicas e camaras de eco, estão reunidas as condições para que a utilização das redes sociais, mesmo que seja só por lazer ou entretenimento, tenda a colocar cada vez mais umas pessoas contra outras. Uma das soluções possíveis para impedir que tal aconteça é a supervisão editorial dos *posts online*, de maneira a corrigir falsidades, moderar exageros e propor correções. Seria o equivalente a fazer a transferência da prática centenária de cartas ao editor, mas para uma escala global. Por outras palavras, seria transformar as plataformas digitais de redes sociais em órgãos de comunicação social e haver um processo de editorialização do discurso e conteúdos, algo que já foi pedido ao Facebook para acontecer (Sullivan, 2016) e que a empresa sugeriu ser possível acontecer num futuro próximo (Waterson, 2019). Uma alteração tão profunda na forma como estas empresas se colocam no mercado, precisaria de um substancial investimento financeiro e de recursos humanos. Nem uma coisa, nem outra, parecem estar iminentes.

Outra expectativa é que os responsáveis por estas empresas sejam recetivos a uma aposta na investigação científica para estudar o efeito da evolução das tecnologias de

⁶⁷ <https://www.facebook.com/communitystandards/>

informação e comunicação no utilizador. Tais investigações focam-se, tanto a nível individual como coletivo, no tempo de permanência e de exposição a material *online*, ao volume de interações e dos seus efeitos. Estes estudos devem conter diferentes amostras com diferentes grupos etários. Em jovens as questões incidem sobre níveis de autoconfiança, de assédio, privacidade, de entrada prematura em grupos de discussão sobre tópicos sociais. Em adultos interessam questões de polarização, proteção de dados, radicalização, e em idosos, desinformação, burlas e teorias da conspiração.

No que se trata da qualidade do ecossistema onde o debate político digital acontece, quando este trabalho está a ser finalizado, o gigante tecnológico Alphabet finalizou quais as regras do alcance de publicidades políticas (Google, 2019). Nas duas plataformas da empresa, Google e YouTube, utilizadores estarão expostos a anúncios que dependem de três categorias: idade, género e localização, sendo que esta última pode ir até a zona de código postal. Desta forma, organizações políticas e/ou candidatos a cargos políticos não podem direcionar as suas publicidades políticas recorrendo a outras informações, tal como afiliação política. O Twitter assume uma atitude mais direta, banindo anúncios políticos da plataforma. Quanto ao Facebook, não mostra sinais de adotar sistemas de proteção relativamente a anúncios em *microtargeting* e sem controlo da veracidade. Isto é importante pelo alcance que o Facebook tem. Relativamente ao efeito da arquitetura digital, políticos, ou grupos políticos, com um grande número de seguidores têm a vantagem de os conteúdos serem amplificados organicamente. Políticos com milhões de seguidores, por exemplo Narendra Modi Twitter, ou Recep Erdoğan no Facebook, têm menos a perder com o final de publicidades políticas ou de *microtargeting* do que aqueles que têm menos alcance e que querem fazer a sua mensagem chegar a mais pessoas. Para estes, uma maior penetração na blogosfera depende das publicidades que possam apresentar ao eleitorado, para além de ser uma oportunidade para corrigir desinformação que possa existir sobre a sua pessoa, ideias e argumentos, e razões para a candidatura a um lugar político.

7.3 Utilizadores e *stakeholders*

Os direitos dos utilizadores devem ser centrais no desenvolvimento de legislação, regras e normas no mundo digital. Porém, os termos e condições mantêm-se densos e complicados, a salvaguarda da privacidade pela utilização do serviço continua insatisfatória, a proteção da informação pessoal para esta não ser usada para *microtargeting* económico e político é insuficiente, os mecanismos de reclamação no caso de serem suprimidos discurso ou conteúdos são opacos, a juntar a uma atitude sobranceira que os responsáveis pelas empresas mostram e o paternalismo da regulação para o “benefício” do utilizador. A erosão

do direito à livre expressão *online* obriga à criação de programas para desenvolvimento da literacia digital para ganho de conhecimento, *skills*, e para a exposição a exemplos práticos para entender processos, políticas e enquadramentos legais (UNESCO, 2016).

Organizações que resultam da sociedade civil precisam também de desempenhar um papel de supervisão e denúncia de abusos de poder, seja pelo Estado ou por empresas privadas. Como já descrito anteriormente, as grandes empresas digitais estão representadas, com embaixadores, antigos vice-primeiros ministros e Comissários Europeus. Isso faz com que as organizações não governamentais precisem de reclamar uma contribuição na definição de políticas e legislação. Este é mais um campo de ação onde algumas organizações podem ter uma função vital, como é o caso da Global Network Initiative⁶⁸, a Internet Corporation for Assigned Names and Numbers⁶⁹, a Internet Engineering Task Force⁷⁰, o Internet Architecture Board⁷¹, a International Telecommunications Union⁷², e instituições regulatórias como o World Summit on the Information Society⁷³ e o Internet Governance Forum⁷⁴ na égide das Nações Unidas. Estes grupos terão de continuar vigilantes às mudanças na cultura da utilização da Internet, principalmente no que se prende ao exercício da liberdade de expressão e do mercado de ideias. O cidadão pode também assumir a iniciativa, e fazer parte de organizações civis, associando-se e contribuindo para o trabalho que essas organizações efetuam. Alguns exemplos são os grupos accessnow.org⁷⁵, [European Digital Rights](https://edri.org/)⁷⁶, [digitalcourage](https://digitalcourage.de/en)⁷⁷, [privacy international \(sic\)](https://privacyinternational.org/)⁷⁸, e em Portugal a Associação 3D⁷⁹.

A área da psicologia e sociologia é de principal importância, na criação de quadros clínicos na primeira, e de integração (ou desintegração) social na segunda. Isto servirá para se conhecer melhor quais os efeitos negativos e positivos da permanência em plataformas digitais de redes sociais, assim como da exposição à arena de discussão e mercado de ideias. Igualmente, a nível escolar deve-se apostar em informação específica nos currículos. Essa informação deve apostar no conhecimento de como viver *online* e como adquirir, e desenvolver, comportamentos e encontrar ferramentas para evitar problemas da utilização das plataformas digitais. Por exemplo, jovens mulheres precisam de ser preparadas para

⁶⁸ <https://globalnetworkinitiative.org/>

⁶⁹ <https://www.icann.org/>

⁷⁰ <https://ietf.org/>

⁷¹ <https://www.iab.org/>

⁷² <https://www.itu.int/en/Pages/default.aspx>

⁷³ <https://www.itu.int/net/wsis/>

⁷⁴ <https://www.intgovforum.org/multilingual/>

⁷⁵ <https://www.accessnow.org/>

⁷⁶ <https://edri.org/>

⁷⁷ <https://digitalcourage.de/en>

⁷⁸ <https://privacyinternational.org/>

⁷⁹ <https://www.direitosdigitais.pt/>

saber como lidar com saberão como melhor lidar com abuso e perversão e as minorias em como lidar com discurso de ódio *online*.

7.4 Quadros regulamentares

A criação, ou atualização de quadros regulamentares, nacionais ou supranacionais, devem atender a preocupações como: quais os problemas que precisam de ser resolvidos; que efeitos causam no mundo digital e real; qual a evolução dos problemas e das soluções; qual a evidência científica que se acumula à medida que o tempo passa. A outra preocupação é evitar entrar numa escala deslizando, onde se começam a criar novos limites para testes de liberdade de expressão na avaliação do princípio de dano, mas principalmente de ofensa, para assim ser mais fácil censurar discurso e conteúdos, com novos precedentes legais e normativos. Na União Europeia para além *Ato de Serviços Digitais*, existem outras diretivas para enquadrar atos ilegais na Internet, como é o caso da *Diretiva 2011/93/EU* para a luta contra o abuso sexual, exploração sexual e pornografia infantil (EPC, 2011), e a *Diretiva 2017/541/EU* contra o terrorismo (EPC, 2017). Igualmente, quando existe um receio real que o discurso passe da intenção à ação, e com isso um atentado à segurança pública, está codificada como as plataformas digitais e as autoridades partilham informação de um utilizador com comportamentos alarmantes (European Commission, 2019).

Nas Nações Unidas, destaca-se o relatório do Representante Especial para o Secretário-Geral das Nações Unidas, John Ruggie, com o nome *Proteger, Respeitar e Remediar: Um Quadro Regulamentar*⁸⁰ (Ruggie, 2008), que contém um conjunto de recomendações para empresas transnacionais e legisladores. Três pilares são apresentados: o dever do Estado em prevenir e lutar contra abusos dos direitos humanos por terceiros incluindo empresas comerciais, as empresas devem ter em atenção esses direitos e atuar de forma a não os infringirem assim como resolver situações onde isso aconteça, e deve existir acesso de queixosos a sistemas de resolução de problemas, tanto judiciais como não-judiciais (Ruggie, 2008). Também dentro das Nações Unidas, em particular via a Organização das Nações Unidas para a Educação, Ciência e a Cultura (UNESCO), um outro documento, de 2016, merece destaque, em particular a publicação *Privacidade, Livre Expressão e Transparência. Redefinição dos Limites na Idade Digital*⁸¹ (UNESCO, 2016) que trata proteção e promoção do exercício da liberdade de expressão, e da produção e difusão de discurso e conteúdos, incluindo os que se relacionam com o processo político.

⁸⁰ *Protect, Respect and Remedy: a Framework*, no original.

⁸¹ *Privacy, free expression and transparency. Redefining their new boundaries in the digital age*, no original.

Conclusão

Em dezembro de 2019, o Tribunal Judicial da Comarca de Lisboa Oeste não deu razão à queixa da empresária Isabel dos Santos contra a diplomata, e ex-membro do Parlamento Europeu, Ana Martins Gomes. Isabel dos Santos reclamava que um conjunto de *tweets* da ex-eurodeputada fossem apagados por os considerar difamatórios. Num dos *tweets* pode ler-se que “[i]sabel dos Santos endivida-se mto porque, ao liquidar as dívidas, “lava” q se farta!”. O tribunal pronunciou que “o direito à liberdade de expressão e de informação prevalece sobre os direitos de personalidade da requerente indeferindo-se por isso a providência requerida” (TJCLO, 2019: 16). No acórdão fica expresso que “[c]omo é sabido, a liberdade de informação e de expressão está inscrita no quadro dos direitos, liberdades e garantias pessoais e tem por fim último garantir a plenitude da democracia, a pluralidade de opiniões e de pensamento” (TJCLO, 2019: 13). A Juiz de Direito refere a “tendência predominante” de privilegiar direitos fundamentais individuais que se relacionam com o bom nome e a reputação, sobre o exercício do direito de opinião e de expressão.

Porém, em sentido contrário, uma das razões para a decisão foi que “[a]ctualmente, por serem ambos direitos fundamentais individuais havendo por isso necessidade de compatibilização e de considerar não só a legislação interna mas as normas da Convenção Europeia dos Direitos Humanos (também conhecida por Convenção Europeia dos Direitos do Homem) e a jurisprudência do TEDH [Tribunal Europeu dos Direitos Humanos], órgão que, nos termos da CEDH [Convenção Europeia dos Direitos Humanos], está especificamente vocacionado para uma interpretação qualificada e controlo da aplicação dos preceitos de Direito Internacional convencional que a integram e que vinculam o Estado Português, não é possível conferir-se abstratamente precedência de um em relação ao outro, tornando-se necessário averiguar no caso concreto, qual deles deve ser sacrificado em relação” (TJCLO, 2019: 14). A Juíza apresenta que “[c]omo se escreve no citado acórdão do Supremo Tribunal de Justiça «O TEDH vem entendendo que – particularmente no âmbito dos artigos que visam essencialmente a expressão da opinião e a crítica a aspectos ligados à vida pública e a temas de manifesto interesse público - está coberta pela liberdade de expressão, não apenas a discordância respeitosa, *a crítica puramente objectiva e moldada pela elevação do debate* – mas também *a crítica contundente, sarcástica, mordaz, com uma carga exageradamente depreciativa ou caricatural da acção e capacidades do visado* – justificando a necessidade de uma particular tolerância deste às opiniões adversas que *criticam acerbamente, chocam, ofendem ou exageram*, envolvendo porventura o uso de expressões agressivas ou virulentas (...)»” (TJCLO, 2019: 15).

Este tipo de desenvolvimento dá alento a defensores da liberdade de expressão, até mesmo de maximalistas, ao referenciar alguns dos pontos abordados neste trabalho no que se trata da defesa de opiniões e argumentos políticos, em particular nas redes sociais: o direito à crítica exacerbada, a ofender sensibilidades políticas, a expressar livremente temas de interesse público político na arena de discussão e no mercado de ideias. Este é o caminho para a manutenção de testes liberais no que se trata da avaliação do valor e qualidade da mensagem. Ainda mais quando, e como é referido justamente na mesma sentença, que o tribunal “não pretende dizer que concorda com o teor dos *tweets*” (TJCLO, 2019: 15).

No filme pornográfico de 1982, a *Academia de Devassas*⁸², uma tradução inusitada do título original, após os créditos finais aparece um fotograma com uma declaração, e uma voz feminina que faz a narração em *voz-off*. Nestas frases vai-se para além de questões legais ou normativas, mas para reflexão importante sobre os limites da liberdade de expressão: “[a]gora que o filme terminou, e exercemos, conjuntamente, os nossos direitos Constitucionais, gostaríamos de o deixar com um pensamento muito importante. Num futuro próximo, você pode ter a oportunidade de participar num julgamento sobre um caso de alegada obscenidade. É prudente lembrar-se que as mesmas pessoas que o querem impedir de ver um filme para adultos podem estar de volta no ano seguinte, para queixarem-se sobre um livro ou até mesmo um programa de televisão. Se lhe pode ser dito o que ver ou ler, então segue-se que lhe possa ser dito o que dizer ou pensar. Defenda os seus direitos constitucionalmente protegidos. Ninguém mais o fará por si”. Thomas Paine, numa obra de teor radicalmente diferente, começa o seu livro seminal *The Age of Reason* com uma das proclamações mais sensatas, e produtivas, no que se trata da defesa da liberdade de expressão e de um ecossistema salutar na arena de discussão: “[e]u coloco este trabalho sobre a sua proteção. Contem como estas as minhas opiniões sobre Religião. Você [o leitor] far-me-á justiça ao lembrar-se, que eu sempre apoiei vigorosamente o Direito de qualquer Homem para ter a sua opinião, por muito diferente que esta seja da minha. Aquele que nega o direito a outro, faz de si um escravo à sua presente opinião, porque se impede a si mesmo de a mudar, A mais formidável arma contra erros de qualquer tipo é a Razão” (Payne, 1794).

Estes dois argumentos, por muito diferentes na origem, resumem perfeitamente a defesa da liberdade de expressão: uma luta de direitos, constitucionalmente garantidos, ou não, de impedir escalas deslizantes que conduzem ao totalitarismo, partir de uma posição de abertura a discurso e conteúdos diferentes, e atingir uma pluralidade de ideias e argumentos em prol de uma deliberação democrática. Esse será o processo racional para atingir

⁸² “*A brief affair*”, no original.

consensos e deliberações. Esta necessidade é particularmente premente quando atravessamos um momento onde, por exemplo, num país como os Estados Unidos onde a pluralidade de ideias e liberdade de expressão sempre foi um baluarte societal, uma sondagem realizada em julho de 2020 mostrou que 62% dos inquiridos acredita que o clima político atual impede de se dizer em que se acredita, pelo receio que “outros podem achar tais opiniões ofensivas” (Cato Institute, 2020). Como em muitos outros aspetos relacionado com o exercício de liberdades, desde que não se infrinja na liberdade dos outros, opiniões e argumentos não devem ser censurados por serem ofensivos.

Não existe tal coisa como ser-se livre de ofensas e existem muitas formas daqueles que se sentem ofendidos exercerem a sua liberdade e expressar desagrado, ou até mesmo responder com um tom igual. Haverá sempre uma tensão entre quem acredite que o “direito a ofender” não implica, necessariamente, o “dever de ofender”, e os que acham que nada é “sagrado” ou “imperturbável”, protegendo opiniões e argumentos de valor baixo ou de conteúdo incomodativo. Essa é um equilíbrio que necessitará de ajustes contínuos para uma convivência salutar. No entanto, evitar o princípio do dano é a principal condição à partida. Quando está em causa a integridade física dos outros, assim como a perda de direitos, consequência resultante da produção de discurso de ódio, existem razões legítimas para que seja efetuado um controlo da qualidade, e, em última análise, uma censura. Porém, tendo como guia o princípio do dano, todo o resto da liberdade de expressão deve ser protegida. Como diz John Stuart Mill, “[a] única liberdade que merece esse nome é aquela que permite procurar o nosso próprio bem da nossa maneira, desde que não haja uma tentativa de privar os outros da sua [liberdade] ou impedir os seus esforços para o conseguir” (Mill, 1859: 22).

As leis, normas e regras que têm servido para a regulação da liberdade de expressão, arena de discussão e mercado de ideias, precisam de ser atualizadas para o mundo digital, que, apesar de ainda se encontrar na sua infância, tornou-se um fator incontornável para o futuro da Humanidade. No livro *A Ordem Mundial*, Henry Kissinger faz a previsão que “proporcionada escolha ao povo, o cidadão comum optaria pela paz - uma tese partilhada pelo filósofo iluminista Kant (...) e pelos apologistas contemporâneos da Internet aberta” (Kissinger, 2014: 300). Se a paz mundial será um objetivo ousado para aqueles que acreditam nas potencialidades da Internet, a esperança de que o mundo virtual seja um local de respeito, troca livre de ideias, segurança e progresso, uma extensão de direitos humanos fundamentais e de expressão de cidadania, é totalmente aceitável e, esperemos, possível. Isto ainda é particularmente mais importante num mundo cada vez mais interligado, e onde soluções digitais são cada vez mais utilizadas para relacionar cidadãos com o Estado, os Estados entre si, e com organizações supranacionais. Na União Europeia, o futuro do tecido político e social

na arena digital irá depender de quais as decisões que os Estados-membro, juntamente com a Comissão Europeia e Parlamento Europeu, nas novas regras do *Ato de Serviços Digitais*, e em futura legislação para plataformas como Facebook e Twitter.

Uma das principais preocupações relativamente à liberdade de expressão *online*, principalmente no que se trata de discurso político na arena digital, não é tanto saber como as plataformas de redes sociais irão definir, ou serem obrigadas a definir, o que será considerado como discurso e/ou conteúdos passíveis de serem censurados. A pergunta deve ser: quais os efeitos no precário balanço entre a defesa de liberdades e a aceitação não criteriosa de soluções que sejam propostas. A opção de existir uma pré-avaliação, automática ou editorial, de opiniões e argumentos, e até mesmo factos, para que estes não choquem com os termos e condições das plataformas digitais, pode tornar-se numa realidade. Compete a todos os *stakeholders*, incluindo os diretores das empresas digitais, de não se incorrer em censura prévia para *à posteriori* ver-se se é aceitável, ou não, o discurso e/os conteúdos. Não deve ser um ponto de partida ter de se provar a “inocência”, ou a “não criminalidade”, mas sim quem apresenta a queixa deve ser capaz de provar que existiu dolo e intenção de causar dano ou de promover o ódio ou a violência iminente. Antecipa-se que leis *safe harbour* se mantenham, mantendo-se a proteção legal das plataformas digitais não serem punidas pelo aparecimento de discurso e/ou conteúdos, mas pela incapacidade da remoção dos mesmos, uma vez provados que estes causam dano.

Não é de esperar que as plataformas digitais de redes sociais sejam um dia geridas como o modelo dos tradicionais órgãos sociais noticiosos. O Facebook e o Twitter argumentam que o volume de discurso e/ou conteúdos produzido diariamente, a nível global, torna impossível uma intervenção humana cabal, ao contrário do que acontece na imprensa tradicional. Poderem ser os utilizadores a obrigar as empresas digitais a entrar em ação para a remoção de discurso/conteúdos com argumentos impopulares, ou que criem ofensa, poderá trazer uma “solução” desproporcionada para o problema. Isto, enquanto ao mesmo tempo, as empresas digitais farão o que for preciso para defenderem as suas margens de lucro.

Com todos os condicionalismos e desafios, e com a necessidade de continuar a desenvolver quadros legais e normativos, os defensores da liberdade de expressão na arena política digital têm o seu trabalho bem definido. O potencial da Internet é inegável, assim como são os perigos. É necessário que, todos os que acreditam que o debate político vai passar a acontecer, e a ser resolvido, em grande medida, e inevitavelmente, nas plataformas digitais, continuem a trabalhar para que a Internet sirva como forma de criar uma sociedade mais igualitária, mais livre e mais democrática.

Bibliografia

- Al-Youssef, M. (2019). *Government seeks to eliminate Internet anonymity – with severe penalties* (online). Recuperado de <https://www.derstandard.at/story/2000101677286/government-seeks-to-eliminate-Internet-anonymity-with-severe-penalties>
- Article XIX (2009). *Civil defamation: undermining free expression* (online). Recuperado de <https://www.article19.org/data/files/medialibrary/1427/civil-defamation.pdf>
- Ash, T. G. (2016). *Free Speech: Ten Principles for a Connected World*. London: Atlantic Book, London.
- AVAAZ (2020). *Why is YouTube broadcasting climate misinformation to millions?* (online). Recuperado de https://avaazimages.avaaz.org/youtube_climate_misinformation.pdf
- Barber, B. (2003). *Strong Democracy. Participatory Politics for a New Age*. Berkley: University of California Press.
- Barlow, J. (1996). *A Declaration of the Independence of Cyberspace* (online). Recuperado de <https://www.eff.org/cyberspace-independence>
- Barron, R. (2005). On the duty of not taking offense. *Journal of Moral Education*, 34, 3, 265-267.
- Benesch, S. (2012). Words as weapons. *World Policy Journal*, 29, 1, 7–12.
- Benesch, S. (2013). *The Dangerous Speech Project. Dangerous speech: a proposal to prevent group violence* (online). Recuperado de <https://perma.cc/DEN4-LKKG>
- Benesch, S., Buerger, C., Glavinic, T., & Manion, S. (2018). *Dangerous speech: a practical guide* (online). Recuperado de <https://dangerousspeech.org/guide/>
- Bengali, S. (2021). *Facebook banned Trump but has failed to react quickly to other leaders who incited violence* (online). Recuperado de <https://www.latimes.com/world-nation/story/2021-01-15/facebook-social-media-bans-trump-capitol-riot>
- Black, T. (2014). *Yes, freedom of speech should be absolute* (online). Recuperado de <https://www.spiked-online.com/author/tim-black/>
- Bollinger, L. (2010). *Uninhibited, Robust and Wide Open: A Free Press for a New Century*. Oxford: Oxford University Press.
- Bonin, D. (2011). *Should race matter? Unusual Answers to Usual Questions*. New York: Cambridge University Press.
- Burch, S. (2020). *Facebook Bans Blackface Posts, Anti-Semitic Conspiracies in Updated Hate Speech Policy* (online). Recuperado de

<https://www.yahoo.com/entertainment/facebook-bans-blackface-posts-anti-222415724.html>

- Burch, S. (2020a). *Facebook Removed 7 Million 'Harmful' COVID-19 Misinformation Posts* (online). Recuperado de <https://www.thewrap.com/facebook-removed-7-million-harmful-covid-19-misinformation-posts>
- Buyse, A. (2014). Words of Violence: “fear speech,” or how violent conflict escalation relates to the freedom of expression. *Human Rights Quarterly*, 36, 4, 779–797.
- Byers, D. (2020). *Facebook bans new political ads in week before election*. (online). Recuperado de <https://www.nbcnews.com/tech/tech-news/facebook-bans-new-political-ads-week-election-n1239159>
- Campese, S. (2015). *Voltaire n’a jamais dit : “Je ne suis pas d’accord avec vous, mais je me battraï...”* (online). Recuperado de <https://www.projet-voltaire.fr/culture-generale/voltaire-citation-apocryphe-je-ne-suis-pas-d-accord-avec-vous/>
- Cato Institute (2020). *Poll: 62% of Americans Say They Have Political Views They’re Afraid to Share* (online). Recuperado de <https://www.cato.org/publications/survey-reports/poll-62-americans-say-they-have-political-views-theyre-afraid-share>
- CD&T [Center for Democracy & Technology] (2011). *Regardless of frontiers. The international freedom of expression in the digital age* (online). Recuperado de https://cdt.org/files/pdfs/CDT-Regardless_of_Frontiers_v0.5.pdf
- CEO [Corporate Europe Observatory] (2018). comunicação pessoal.
- Chafee, Z. (1946). *Free Speech in the United States* (2ª ed). Cambridge: Harvard University Press.
- Citron, K. D., & Norton, H. (2011). Intermediaries and hate speech: fostering digital citizenship for our information age. *Boston University Law Review*, 91, 1435–1484.
- CJEU [Court of Justice of the European Union] (2019). *Advocate General Sharpston: the Court should rule that, by refusing to comply with the provisional and time-limited mechanism for the mandatory relocation of applicants for international protection, Poland, Hungary and the Czech Republic have failed to fulfill their obligations under EU law* (online). Recuperado de <https://curia.europa.eu/jcms/upload/docs/application/pdf/2019-10/cp190133en.pdf>
- CoE [Council of Europe] (2003). *Explanatory Report to the Additional Protocol to the Convention on Cybercrime, concerning the criminalisation of acts of a racist and xenophobic nature committed through computer systems* (online). Recuperado de <https://rm.coe.int/16800d37ae>

- Crook, J. (2018). *Twitter turns to academics to improve conversational health on the platform* (online). Recuperado de <https://techcrunch.com/2018/07/30/twitter-turns-to-academics-to-improve-conversational-health-on-the-platform/>
- DGCNCT [Directorate-General for Communications Networks, Content and Technology] (2019). *Update would encompass a REFIT of the E-Commerce Directive* (online). Recuperado de <https://cdn.netzpolitik.org/wp-upload/2019/07/Digital-Services-Act-note-DG-Connect-June-2019.pdf>
- Diário da República (1978). Pacto Internacional sobre os Direitos Civis e Políticos (online). Recuperado de <https://dre.pt/web/guest/pesquisa/-/search/426144/details/normal?q=Lei+29%2F78>
- Dickson, A. (2019). *UK races to finalize law policing 'wild west' Internet* (online). Recuperado de <https://www.politico.eu/article/uk-races-to-finalize-law-policing-wild-west-Internet-google-facebook>
- Die Bundesregierung (2019). *Bundesregierung geht gegen Rechtsextremismus und Hasskriminalität vor* (online). Recuperado de <https://www.bundesregierung.de/breg-de/aktuelles/gegen-extremismus-und-hass-1686442>
- Dorsey, J. (2016). *Twitter creator Jack Dorsey illuminates the site's founding document* (online). Recuperado de <https://latimesblogs.latimes.com/technology/2009/02/twitter-creator.html>
- Downs, D. A. (1985). Skokie revisited: hate group speech and the First Amendment. *Notre Dame Law Review*, 60, 4, 629-685.
- Dwoskin, E., & Timberg, C. (2021). *Misinformation dropped dramatically the week after Twitter banned Trump and some allies* (online). Recuperado de <https://www.washingtonpost.com/technology/2021/01/16/misinformation-trump-twitter/>
- ECHR [European Court of Human Rights] (1950). *Convention for the Protection of Human Rights and Fundamental Freedoms* (online). Recuperado de https://www.echr.coe.int/Documents/Convention_ENG.pdf
- ECHR (1962). *Case of DeBecker v. Belgium* (online). Recuperado de <https://www.legal-tools.org/doc/49edbf/pdf/>
- ECHR (1979). *Case of The Sunday Times v. The United Kingdom* (online). Recuperado de <http://hudoc.echr.coe.int/rus?i=001-57584>
- ECHR (1986). *Case of Lingens v. Austria* (online). Recuperado de <http://hudoc.echr.coe.int/fre?i=001-57523>
- ECHR (1976). *Case of Handyside v. The United Kingdom* (online). Recuperado de <http://hudoc.echr.coe.int/eng?i=001-57499>

- ECRI [European Commission against Racism and Intolerance] (2018). *Relatório da ECRI sobre Portugal* (online). Recuperado de <https://rm.coe.int/fifth-report-on-portugal-portuguese-translation-/16808de7db>
- Economy, E.C. (2018). *The great firewall of China: Xi Jinping's internet shutdown* (online). Recuperado de <https://www.theguardian.com/news/2018/jun/29/the-great-firewall-of-china-xi-jinpings-internet-shutdown>
- EdiMA (2018). *EDiMA reaction: EU Copyright Directive is not fit for digital era* (online). Recuperado de <https://edima-eu.org/news/edima-reaction-eu-copyright-directive-is-not-fit-for-digital-era>
- EDRi [European Digital Rights] (2019). *UK: Online Harms Strategy must “design in” fundamental rights* (online). Recuperado de <https://edri.org/uk-online-harms-strategy-must-design-in-fundamental-rights/>
- EDRi (2019a). *More responsibility to online platforms – but at what cost?* (online). Recuperado de <https://edri.org/more-responsibility-to-online-platforms-but-at-what-cost>
- EPC [European Parliament and the Council] (2000). *On certain legal aspects of information society services, in particular electronic commerce, in the Internal Market (Directive on electronic commerce)* (online). Recuperado de <https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32000L0031&from=EN>
- EPC (2011). *On combating the sexual abuse and sexual exploitation of children and child pornography, and replacing Council Framework Decision 2004/68/JHA* (online). Recuperado de <https://op.europa.eu/en/publication-detail/-/publication/d20901a4-66cd-439e-b15e-faeb92811424/language-en>
- EPC (2017). *On combating terrorism* (online). Recuperado de <https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32017L0541&from=EN>
- Epicenter Works (2019). *Austria: new “responsibility” law will lead to self-censorship* (online). Recuperado de <https://edri.org/austria-new-responsibility-law-will-lead-to-self-censorship/>
- Estado Português (2019). *Código Penal. Disposições relevantes em matéria de comunicação social* (online). Recuperado de <http://bocc.ubi.pt/pag/estado-portugues-codigo-penal.html>
- European Commission (2012). *Online services, including e-commerce, in the Single Market* (online). Recuperado de <https://op.europa.eu/en/publication-detail/-/publication/2d17d3dd-4f00-4e3c-b75c-98c03b135047/language-en>

- European Commission (2014). *Internet Policy and Governance Europe's role in shaping the future of Internet Governance* (online). Recuperado de <https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:2014:0072:FIN:EN:PDF>
- European Commission (2016). *Justice and fundamental rights. Combatting discrimination racism and xenophobia. Countering illegal hate speech online* (online). Recuperado de https://ec.europa.eu/newsroom/document.cfm?doc_id=42830
- European Commission (2017). *Communication on Tackling Illegal Content Online - Towards an enhanced responsibility of online platforms* (online). Recuperado de <https://ec.europa.eu/digital-single-market/en/news/communication-tackling-illegal-content-online-towards-enhanced-responsibility-online-platforms>
- European Commission (2018). *On measures to effectively tackle illegal content online* (online). Recuperado de https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=50095
- European Commission (2018a). *Commission Recommendation on measures to effectively tackle illegal content online* (online). Recuperado de <https://ec.europa.eu/digital-single-market/en/news/commission-recommendation-measures-effectively-tackle-illegal-content-online>
- European Commission (2019). *Factsheet - 4th monitoring round of the Code of Conduct* (online). Recuperado de https://ec.europa.eu/info/files/factsheet-4th-monitoring-round-code-conduct_en
- European Commission (2020). *Proposal for a regulation of the European Parliament and of the Council on a a Single Market For Digital Services (Digital Services Act) and amending Directive 2000/31/EC* (online). Recuperado de <https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52020PC0825&from=en>
- European Union (2000). *Charter of Fundamental Rights of the European Union* (online). Recuperado de https://www.europarl.europa.eu/charter/pdf/text_en.pdf
- European Union (2012). *Consolidated Version of the Treaty on European Union* (online). Recuperado de https://eur-lex.europa.eu/resource.html?uri=cellar:2bf140bf-a3f8-4ab2-b506-fd71826e6da6.0023.02/DOC_1&format=PDF
- Evans, J. (2019). *Facebook isn't free speech, it's algorithmic amplification optimized for outrage* (online). Recuperado de <https://techcrunch.com/2019/10/20/facebook-isnt-free-speech-its-algorithmic-amplification-optimized-for-outrage>
- Fanta, A., & Rudl, T. (2019). *Leaked document: EU Commission mulls new law to regulate online platforms* (online). Recuperado de <https://netzpolitik.org/2019/leaked-document-eu-commission-mulls-new-law-to-regulate-online-platforms>

- FTC [Federal Trade Commission] (2020). *Complaint for injunctive and other equitable relief Federal Trade Commission v. Facebook, Inc.* (online). Recuperado de <https://www.ftc.gov/system/files/documents/cases/1910134fbcomplaint.pdf>
- Feinberg, J. (1988). *The Moral Limits of the Criminal Law. Volume 2: Offense to Others*. New York: Oxford University Press.
- Fish, S. (1994). *There's No Such Thing as Free Speech...and it's a good thing too*. New York: Oxford University Press.
- Flynn, D. J., Nyhan, B., & Reifler, J. (2017). The nature and origins of misperceptions: understanding false and unsupported beliefs about politics. *Advances in Political Psychology*, 38, 127– 150.
- France Press (2019). *France online hate speech law to force social media sites to act quickly* (online). Recuperado de <https://www.theguardian.com/world/2019/jul/09/france-online-hate-speech-law-social-media>
- Freedom House (2014). *Freedom on the NET 2014* (online). Recuperado de https://freedomhouse.org/sites/default/files/FOTN_2014_Full_Report_compressed_v2_0.pdf
- Freedom House (2019). *Freedom on the NET. The Crisis of Social Media* (online). Recuperado de https://www.freedomonthenet.org/sites/default/files/2019-11/11042019_Report_FH_FOTN_2019_final_Public_Download.pdf
- Friedersdorf, C. (2018). *YouTube extremism and the long tail* (online). Recuperado de <https://www.theatlantic.com/politics/archive/2018/03/youtube-extremism-and-the-long-tail/555350/>
- Fung, B. (2020). Facebook must be broken up, the US government says in a groundbreaking lawsuit(online). Recuperado de <https://edition.cnn.com/2020/12/09/tech/facebook-antitrust-lawsuit-ftc-attorney-generals/index.html>
- Garcia, F. (2019). *Estudo revela que portugueses passam em média duas horas por dia na Internet* (online). Recuperado de <https://visao.sapo.pt/exameinformatica/noticias-ei/mercados/2019-10-25-estudo-revela-que-portugueses-passam-em-media-duas-horas-por-dia-na-internet/>
- Gesellschaft für Freiheitsrechte (2019). *Fighting online hate speech: an alternative to mandatory real names* (online). Recuperado de <https://edri.org/fighting-online-hatespeech-an-alternative-to-mandatory-real-names>
- Goldsmith, J., & Wu, T. (2008). *Who Controls the Internet? Illusions of a Borderless World*. New York: Oxford University Press.
- Goodwin, M. (1994). *Meme, counter-meme* (online). Recuperado de <https://www.wired.com/1994/10/godwin-if-2/>

- Google (2019). *An update on our political ads policy* (online). Recuperado de <https://blog.google/technology/ads/update-our-political-ads-policy>
- Greenspan, R.E. (2020). *How 'cancel culture' quickly became one of the buzziest and most controversial ideas on the internet* (online). Recuperado de <https://www.insider.com/cancel-culture-meaning-history-origin-phrase-used-negatively-2020-7>
- Grewal, D. (2018). *Do Trump tweets spur hate crimes?* (online). Recuperado de <https://www.scientificamerican.com/article/do-trump-tweets-spur-hate-crimes1>
- Griffiths, J. (2019). *Singapore just used its fake news law. Critics say it's just what they feared* (online). Recuperado de <https://www.cnn.com/2019/11/29/media/singapore-fake-news-facebook-intl-hnk/index.html>
- Gunia, A. (2019). *Facebook tightens live-stream rules in response to the Christchurch massacre* (online). Recuperado de <https://time.com/5589478/facebook-livestream-rules-new-zealand-christchurch-attack/>
- Hannon, E. (2019). *Twitter suspends 200,000 suspected China-sponsored accounts targeting Hong Kong protests* (online). Recuperado de <https://slate.com/news-and-politics/2019/08/twitter-suspends-suspected-china-sponsored-accounts-targeting-hong-kong-protests-facebook.html>
- Hatmaker, T. (2020). *Twitter broadly bans any COVID-19 tweets that could help the virus spread* (online). Recuperado de <https://techcrunch.com/2020/03/18/twitter-coronavirus-covid-19-misinformation-policy/>
- HoL [House of Lords] (2015). *Data protection law: a 'right to be forgotten'?* (online). Recuperado de <https://publications.parliament.uk/pa/ld201415/ldselect/ldcom/40/40.pdf>
- Holmes, A. (2019). *Twitter just suspended over 88,000 accounts tied to a Saudi disinformation campaign* (online). Recuperado de <https://www.businessinsider.com/twitter-removed-accounts-saudi-arabia-disinformation-campaign-2019-12>
- Hudson, D. L. (2018). *In the age of social media, expand the reach of the First Amendment* (online). Recuperado de <https://www.americanbar.org/content/dam/aba/administrative/crsj/human-rights-magazine/hr-v43-n4.pdf>
- Hughes, C. (2019). *It's time to break up Facebook* (online). Recuperado de <https://www.nytimes.com/2019/05/09/opinion/sunday/chris-hughes-facebook-zuckerberg.html>

- ILS [Internet Live Stats] (2019). Google search statistics (online). Recuperado de <https://www.Internetlivestats.com/google-search-statistics/>
- IPI [International Press Institute] (2015). *Out of balance. Defamation Law in the European Union: a comparative overview for journalists, civil society and policymakers* (online). Recuperado de <https://ipi.media/wp-content/uploads/2016/08/IPI-OutofBalance-Final-Jan2015.pdf>
- Isaac, M., & Conger, K (2021). *Facebook bars Trump through end of his term.* (online). Recuperado de <https://www.nytimes.com/2021/01/07/technology/facebook-trump-ban.html>
- Jackson, B. (2014). Censorship and Freedom of Expression in the Age of Facebook. *New Mexico Law Review*, 44, 1, 121-167.
- Jacobson, D. (2000). Freedom of speech acts: a response to Langton. *Philosophy and Public Affairs*, 24, 64-79
- Jerit, J., & Barabas, J. (2012). Partisan perceptual bias and the information environment. *Journal of Politics*, 74, 3, 672–684.
- Kastrenakes, J. (2018). *Twitter reports a million fewer users as a result of ongoing crackdown on bots* (online). Recuperado de <https://www.theverge.com/2018/7/27/17620440/twitter-q2-2018-earnings-1-million-mau-fall-in-spam-crackdown>
- Kayali, L. (2019). *Inside Facebook's fight against European regulation* (online). Recuperado de <https://www.politico.eu/article/inside-story-facebook-fight-against-european-regulation/>
- Kaye, D. (2020). *The Republic of Facebook* (online). Recuperado de <https://www.justsecurity.org/70035/the-republic-of-facebook>
- Kelly, M. (2019). *Google under antitrust investigation by 50 attorneys general* (online). Recuperado de <https://www.theverge.com/2019/9/9/20857440/google-antitrust-investigation-attorneys-general-advertising-search>
- Kissinger, H. (2014). *A Ordem Mundial*. (4ª ed). Alfragide: D. Quixote.
- La Rue, F. (2011). *United Nations Report of the Special Rapporteur of the promotion and protection of the right to freedom of opinion and expression. Humans Right Council. Seventeen session. A/HRC/17/27* (online). Recuperado de https://www2.ohchr.org/english/bodies/hrcouncil/docs/17session/a.hrc.17.27_en.pdf
- Lecher, C. (2018). *Twitter reportedly suspended 70 million accounts in past two months in crackdown* (online). Recuperado de

- <https://www.theverge.com/2018/7/6/17542480/twitter-account-suspensions-crackdown>
- Lecher, C. (2019). *Facebook is under investigation by state attorneys general for antitrust violations* (online). Recuperado de <https://www.theverge.com/2019/9/6/20852566/facebook-antitrust-investigation-states-attorneys-general-letitia-james-big-tech>
- Lee, L.L., Sohn, Y., & Fowler J.H. (2013). Emotion regulation as the foundation of political attitudes: does reappraisal decrease support for conservative policies? *PLoS ONE*, 812, e83143
- Lengauer, G., Esser F., & Berganza, R. (2011). Negativity in political news: a review of concepts, operationalizations and key findings. *Journalism*, 13, 2, 179-202.
- Locke, J. (1689). *A Letter concerning toleration* (online). Recuperado de <https://www.constitution.org/jl/tolerati.htm>
- Lomas, N. (2015). *Twitter makes tweaks to limit abuse and abusers* (online). Recuperado de <https://techcrunch.com/2015/04/21/twitter-makes-tweaks-to-limit-abuse-and-abusers/>
- Lomas, N. (2019). *Online platforms still not clear enough about hate speech takedowns: EC* (online). Recuperado de <https://techcrunch.com/2019/02/04/online-platforms-still-not-clear-enough-about-hate-speech-takedowns-ec/>
- Lomas, N. (2019a). *Europe's recharged antitrust chief makes her five-year pitch to be digital EVP* (online). Recuperado de <https://techcrunch.com/2019/10/08/europes-recharged-antitrust-chief-makes-her-five-year-pitch-to-be-digital-evp/>
- Macrotrends (2019). *Facebook net worth 2009-2020* (online). Recuperado de <https://www.macrotrends.net/stocks/charts/FB/facebook/net-worth>
- Madison, J. (1787). *Federalist No. 10* (online). Recuperado de <https://billofrightsinstitute.org/founding-documents/primary-source-documents/the-federalist-papers/federalist-papers-no-10/>
- Madison, J. (1798). *Virginia and Kentucky Resolution* (online). Recuperado de <https://billofrightsinstitute.org/founding-documents/primary-source-documents/virginia-and-kentucky-resolutions/>
- McCain, C. (2017). *Why I support free speech absolutism (even after Charlottesville)* (online). Recuperado de <https://www.theodysseyonline.com/support-free-speech-absolutism-charlottesville>
- McGill, M. H., & Lippman, D. (2019). *White House drafting executive order to tackle Silicon Valley's alleged anti-conservative bias* (online). Recuperado de <https://www.politico.com/story/2019/08/07/white-house-tech-censorship-1639051>

- McKinnon, R. (2012). *Consent of the Networked. The Worldwide Struggle for Internet Freedom*. New York: Basic Books.
- Meiklejohn, A. (1961). The First Amendment is an absolute. *The Supreme Court Review*, 1961, 245-266.
- Microsoft (2020). *Microsoft appoints senior government affairs leaders in Brussels and New York, establishes New York office to work with the United Nations* (online). Recuperado de <https://blogs.microsoft.com/eupolicy/2020/01/17/senior-gov-affairs-leaders-appointed-brussels-new-york/>
- Mill, J. S. (1859). *On Liberty*. Batoche Book Limited. 2001. Ontario: Batoche Book.
- Milton, J. (1644). *Areopagitica* (online). Recuperado de <http://pinkmonkey.com/dl/library1/milt16.pdf>
- Moskalenko, S. (2018). *Why social media makes us angrier—and more extreme* (online). Recuperado de <https://www.psychologytoday.com/gb/blog/friction/201807/why-social-media-makes-us-angrier-and-more-extreme>
- Mossberger, K., Tolbert, C.J., & McNeal, R.S. (2008). *Digital citizenship: The Internet, society, and participation*. Cambridge: The MIT Press.
- Naranjo, D. (2019). *Von der Leyen: an ambitious agenda for digital rights* (online). Recuperado de <https://edri.org/von-der-leyen-ambitious-agenda-for-digital-rights>
- Neto, I. (2019). *O dia a dia traumático dos moderadores do Facebook* (online). Recuperado de <https://www.noticiasmagazine.pt/2019/o-dia-a-dia-traumatico-dos-moderadores-do-facebook/historias/241365/>
- Newton, C. (2019). *The trauma floor* (online). Recuperado de <https://www.theverge.com/2019/2/25/18229714/cognizant-facebook-content-moderator-interviews-trauma-working-conditions-arizona>
- Newton, C. (2019a). *Facebook's embrace of Breitbart doesn't add up* (online). Recuperado de <https://www.theverge.com/interface/2019/10/29/20936441/facebook-news-breitbart-mosseri-trust-political-ads>
- Nguyen, M. (2017). *Twitter's role in politics* (online). Recuperado de <https://northwesternbusinessreview.org/twitters-role-in-politics-b3ed620465c9>
- O'Connor, M. E. (2019). *How Much Time Do People Spend on Social Media in 2019?* (online). Recuperado de <https://medium.com/@JBBC/how-much-time-do-people-spend-on-social-media-in-2019-infographic-cc02c63bede8>
- O'Neil, T. (2019). *Absolutist* (online). Recuperado de <https://www.mtsu.edu/first-amendment/article/887/absolutists>

- Panoptykon Foundation (2019). *"SIN vs Facebook": first victory against privatised censorship* (online). Recuperado de <https://edri.org/sin-vs-facebook-first-victory-against-privatised-censorship>
- Payne, T. (1794). *Age of Reason. Dedication* (online). Recuperado de <http://www.let.rug.nl/usa/documents/1786-1800/thomas-paine-the-age-of-reason/dedication.php>
- Persily, N. (2017). Can Democracy Survive the Internet? *Journal of Democracy*, 28, 63-76.
- Popper, K. (1945). *The Open Society and Its Enemies* (7ª ed). Kindle Edition. London: Routledge & Kegan Paul Ltd.
- Puddephatt, A. (2011). *Freedom of expression rights in the digital age* (online). Recuperado de <https://www.opensocietyfoundations.org/publications/mapping-digital-media-freedom-expression-rights-digital-age>
- Putnam, R. D. (2000). *Bowling Alone: The Collapse and Revival of American Community*. New York: Simon & Schuster
- Reuters Institute (2018). *Digital News Report 2018* (online). Recuperado de https://obercom.pt/wp-content/uploads/2018/09/DNR_PT_2018.pdf
- Rocheleau, M. (2017). *Here's the full list of Russia-linked Twitter handles* (online). Recuperado de <https://www3.bostonglobe.com/metro/2017/11/01/here-full-list-russia-linked-twitter-handles/UqBLnttWWTSpw16vsvOPO/story.html?arc404=true>
- Rodrigues, L.B., & Carneiro A.S. (2011). *Direitos humanos: textos fundamentais*. Lisboa: Quid Juris.
- Rodriguez, S. (2019). *Snopes quits fact-checking partnership with Facebook* (online). Recuperado de <https://www.cnn.com/2019/02/01/snopes-leaves-facebook-fact-check-program.html>
- Rohaidi, N. (2019). *Exclusive: how Estonia is training young people to spot fake news* (online). Recuperado de <https://govinsider.asia/connected-gov/mart-laidmets-kristel-rillo-estonian-ministry-of-education-and-research-fake-news-digital-competence>
- Room, T. (2020). *Trump threatens to veto major defense bill unless Congress repeals Section 230, a legal shield for tech giants* (online). <https://www.washingtonpost.com/technology/2020/12/01/trump-repeal-section-230-ndaa/>
- Roth, A. (2019). *Russia's great firewall: is it meant to keep information in – or out?* (online). Recuperado de <https://www.theguardian.com/technology/2019/apr/28/russia-great-firewall-sovereign-internet-bill-keeping-information-in-or-out>

- Ruggie, J. (2008). *United Nations Report of the Special Representative of the Secretary-General on the issue of human rights and transnational corporations and other business enterprises - Protect, Respect and Remedy: a Framework. Eight session. A/HRC/8/5* (online). Recuperado de <https://daccess-ods.un.org/TMP/7027755.97572327.html>
- Sandler, R. (2019). *Twitter usually leaves Trump's tweets alone—except when it comes to copyright* (online). Recuperado de <https://www.forbes.com/sites/rachelsandler/2019/10/03/twitter-usually-leaves-trumps-tweets-aloneexcept-when-it-comes-to-copyright/#6c6d05461b00>
- Scalon, T. (2015). *The Difficulty of Tolerance: Essays in Political Philosophy*. Cambridge: Cambridge University Press.
- Schauer, F. (1985). Slippery Slopes. *Harvard Law Review*, 99, 361-383.
- SCI [Supreme Court Illinois] (1978). *Village of Skokie v. Nat'l Socialist Party of America* (online). Recuperado de <https://law.justia.com/cases/illinois/supreme-court/1978/49769-6.html>
- Scola, N., & Overly, S. (2019). *Twitter drops all political ads in shot at Zuckerberg* (online). Recuperado de <https://www.politico.com/news/2019/10/30/twitter-dropping-all-political-ads-000308>
- SEP [Stanford Encyclopedia of Philosophy] (2019). *Freedom of Speech* (online). Recuperado de <https://plato.stanford.edu/entries/freedom-speech/>
- Shieber, J. (2018). *After selling his company to Facebook for \$19 billion, Brian Acton joins #deleteFacebook* (online). Recuperado de <https://techcrunch.com/2018/03/20/after-selling-his-company-to-facebook-for-19-billion-brian-acton-joins-deletfacebook>
- Starr, S. (2019). *Hate speech on the Internet* (online). Recuperado de <https://www.osce.org/fom/13846?download=true>
- Statista (2019). *Number of monthly active Facebook users worldwide* (online). Recuperado de <https://www.statista.com/statistics/264810/number-of-monthly-active-facebook-users-worldwide>
- Statista (2019a). *Number of monthly active Twitter users* (online). Recuperado de <https://www.statista.com/statistics/282087/number-of-monthly-active-twitter-users>
- Statista (2019b). *Number of daily active WhatsApp status users* (online). Recuperado de <https://www.statista.com/statistics/730306/whatsapp-status-dau/>
- Sullivan, M. (2016). *Call it a 'crazy idea,' Facebook, but you need an executive editor* (online). Recuperado de https://www.washingtonpost.com/lifestyle/style/call-it-what-you-want-facebook-but-you-need-an-executive-editor/2016/11/20/67aa5320-aaa6-11e6-a31b-4b6397e625d0_story.html

- Sunstein, C. (1995). *Democracy and the Problem of Free Speech*. New York: The Free Press.
- Singer, P.W., & Brooking, E. (2018). *Here's China's massive plan to retool the web* (online). Recuperado de <https://www.popsoci.com/chinas-massive-plan-to-retool-web>
- Sweney, M. (2018). *Facebook hires Nick Clegg as head of global affairs* (online). Recuperado de <https://www.theguardian.com/technology/2018/oct/19/facebook-hires-nick-clegg-as-head-of-global-affairs>
- Swisher, K. (2019). *Trump Is Too Dangerous for Twitter* (online). Recuperado de <https://www.nytimes.com/2019/10/01/opinion/trump-civil-war.html>
- Tadeu, P., 2017. *Passos Coelho não é racista, o seu preconceito é outro* (online). Recuperado de <https://www.dn.pt/opiniaio/opiniaio-dn/pedro-tadeu/interior/passos-coelho-nao-e-racista-o-seu-preconceito-e-outro-8706053.html>
- TIEU [Transparency International EU] (2019). *Press release: historic day for whistleblowers as EU agrees pathbreaking legislation* (online). Recuperado de <http://transparency.eu/press-release-historic-day-whistleblowers/>
- TJCLO [Tribunal Judicial da Comarca de Lisboa Oeste] (2019). *Acórdão Judicial* (online). Recuperado de <https://www.anagomes.eu/PublicDocs/d0d94ade-a8c0-4d23-a6fe-b034a5a22b2d.pdf>
- Tocqueville, A. (1835). *Democracy in America*. Edited by Francis Bowen, 1863 (3ª ed). Cambridge: Sever and Francis.
- Tracy, R. (2020). *Social Media's Liability Shield Is Under Assault*. (online). Recuperado de <https://www.wsj.com/articles/social-medias-liability-shield-is-under-assault-11606402800>
- Trussler, M., & Soroka, T. (2014). Consumer Demand for Cynical and Negative News Frames. *The International Journal of Press/Politics*, 19, 3, 360-379.
- Tsisis, A. (2001). Hate in cyberspace: Regulating hate speech on the Internet. *San Diego Law Review*, 38, 817-54.
- TWG [Transatlantic Working Group] *An analysis of Germany's NetzDG Law* (2019). (online). Recuperado de https://www.ivir.nl/publicaties/download/NetzDG_Tworek_Leerssen_April_2019.pdf
- Twitter (2019). *World leaders on Twitter: principles & approach* (online). Recuperado de https://blog.twitter.com/official/en_us/topics/company/2019/worldleaders2019.html
- Twitter (2019a). *About verified accounts* (online). Recuperado de <https://help.twitter.com/en/managing-your-account/about-twitter-verified-accounts>
- Twitter (2021). *Permanent suspension of @realDonaldTrump* (online). https://blog.twitter.com/en_us/topics/company/2020/suspension.html

- UK Government (1986). *Public Order Act 1986* (online). Recuperado de <http://www.legislation.gov.uk/ukpga/1986/64/section/5>
- UK Government (1986a). *Public Order Act 1986* (online). Recuperado de <http://www.legislation.gov.uk/ukpga/1986/64/section/18>
- UK Government (2019). *Online Harms White Paper* (online). Recuperado de https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/793360/Online_Harms_White_Paper.pdf
- UKDDCMS [United Kingdom Department for Digital, Culture, Media & Sport] *Online Harms White Paper* (2019). (online). Recuperado de <https://www.gov.uk/government/consultations/online-harms-white-paper>
- UNESCO [United Nations Educational, Scientific and Cultural Organization] (2011). *Reflection and analysis by UNESCO on the Internet* (online). Recuperado de <https://unesdoc.unesco.org/ark:/48223/pf0000211062>
- UNESCO (2014). *Paris Declaration on media and information literacy in the digital era* (online). Recuperado de http://www.unesco.org/new/fileadmin/MULTIMEDIA/HQ/CI/CI/pdf/In_Focus/paris_mil_declaration_final.pdf
- UNESCO (2015). *Countering online hate speech* (online). Recuperado de <https://unesdoc.unesco.org/ark:/48223/pf0000233231>
- UNESCO (2016). *Privacy, free expression and transparency: redefining their new boundaries in the digital age* (online). Recuperado de <https://unesdoc.unesco.org/ark:/48223/pf0000246610>
- UNHRC [United Nations Human Rights Council] (1966). *International Covenant on Civil and Political Rights* (online). Recuperado de <https://www.ohchr.org/en/professionalinterest/pages/ccpr.aspx>
- UNHRC (2006). *Implementation of General Assembly Resolution 60/251 of 15 March 2006 Entitled "Human Rights Council"* (online). Recuperado de <https://undocs.org/A/HRC/2/3>
- UNHRC (2011). *General comment No. 34 Article 19: freedoms of opinion and expression* (online). Recuperado de <https://www2.ohchr.org/english/bodies/hrc/docs/gc34.pdf>
- UNHRC (2011a). *Guiding principles on business and Human Rights* (online). Recuperado de https://www.ohchr.org/documents/publications/GuidingprinciplesBusinesshr_eN.pdf
- UNHRC (2014). *Joint Declaration on Freedom of Expression and responses to conflict situations* (online). Recuperado de

<https://www.ohchr.org/EN/NewsEvents/Pages/DisplayNews.aspx?NewsID=15921&LangID=E>

UNHRC (2015). *Report of the Special Rapporteur on minority issues* (online). Recuperado de [https://documents-dds-](https://documents-dds-ny.un.org/doc/UNDOC/GEN/G15/000/32/PDF/G1500032.pdf?OpenElement)

[ny.un.org/doc/UNDOC/GEN/G15/000/32/PDF/G1500032.pdf?OpenElement](https://documents-dds-ny.un.org/doc/UNDOC/GEN/G15/000/32/PDF/G1500032.pdf?OpenElement)

UNHRC (2016). *The promotion, protection and enjoyment of human rights on the Internet* (online). Recuperado de <https://documents-dds-ny.un.org/doc/UNDOC/LTD/G16/131/89/PDF/G1613189.pdf?OpenElement>

UNHROHC [United Nations Human Rights Office of the High Commissioner] (2018). *The right to privacy in the digital age* (online). Recuperado de <https://www.ohchr.org/en/issues/digitalage/pages/digitalageindex.aspx>

Unikaitė-Jakuntavičienė, I., & Rakutienė, S. *Writing a Bachelor's Thesis in the Field of Political Science. Didactical Guidelines* (online). Recuperado de http://www.esparama.lt/es_parama_pletra/failai/ESFproduktai/2013_metodine_prie_mone_Writing_a_Bachelors_Thesis.pdf

United Nations (1948). *Universal Declaration of Human Rights* (online). Recuperado de <https://www.un.org/en/universal-declaration-human-rights/>

US Congress (1798). *The Alien and Sedition Acts of 1798* (online). Recuperado de <https://www.thoughtco.com/the-alien-and-sedition-acts-of-1798-4176452>

US Congress (1917). *An Act to punish acts of interference with the foreign relations, the neutrality, and the foreign commerce of the United States, to punish espionage, and better to enforce the criminal laws of the United States, and for other purposes* (online). Recuperado de <https://uslaw.link/citation/us-law/public/65/24>

USC [United States Code] (2018). *47 U.S. Code § 230. Protection for private blocking and screening of offensive material* (online). Recuperado de <https://www.law.cornell.edu/uscode/text/47/230>

USSC [United States Supreme Court] (1919). *Schenck v. United States* (online). Recuperado de <https://supreme.justia.com/cases/federal/us/249/47/>

USSC (1919a). *Abrams v. United States* (online). Recuperado de <https://supreme.justia.com/cases/federal/us/250/616/>

USSC (1927). *Whitney v. California* (online). Recuperado de <https://supreme.justia.com/cases/federal/us/274/357/>

USSC (1927a). *Justice Brandeis, Concurring Opinion in Whitney v California* (online). Recuperado de http://www.columbia.edu/itc/journalism/j6075/edit/readings/brandeis_concurring1.html

- USSC (1937). *Palko v. Connecticut* (online). Recuperado de <https://supreme.justia.com/cases/federal/us/302/319/>
- USSC (1953). *United States v. Rumely* (online). Recuperado de <https://supreme.justia.com/cases/federal/us/345/41/>
- USSC (1969). *Brandenburg v. Ohio* (online). Recuperado de <https://supreme.justia.com/cases/federal/us/395/444/>
- USSC (2016). *Packingham v. North Carolina* (online). Recuperado de <https://supreme.justia.com/cases/federal/us/582/15-1194/>
- Vuorikari, R., Punie, Y., Carretero, S., & Van den Brande, L. (2016). *DigComp 2.0: The Digital Competence Framework for Citizens* (online). Recuperado de https://publications.jrc.ec.europa.eu/repository/bitstream/JRC101254/jrc101254_digcomp%202.0%20the%20digital%20competence%20framework%20for%20citizens.%20update%20phase%201.pdf
- de Vries, I., Ruckert, N.W., Taelman, E., Ruers, C., & van der Leew-Veiksha, A. (2020). *The Digital Services Act package is here* (online). Recuperado de <https://www.allenoverly.com/en-gb/global/news-and-insights/publications/the-digital-services-act-package-is-here>
- Waimberg, J. (2015). *Schenck v. United States: defining the limits of free speech* (online). Recuperado de <https://constitutioncenter.org/blog/schenck-v-united-states-defining-the-limits-of-free-speech/>
- Waldon, J. (2012). *The Harm in Hate Speech*. Cambridge: Harvard University.
- Waterson, J. (2019). *Facebook considering hiring editors to pick quality news for users* (online). Recuperado de <https://www.theguardian.com/technology/2019/apr/01/facebook-mark-zuckerberg-considering-hiring-editors-journalism-pick-quality-news>
- Wermiel, S. J. (2018). *The ongoing challenge to define free speech* (online). Recuperado de <https://www.americanbar.org/content/dam/aba/administrative/crsj/human-rights-magazine/hr-v43-n4.pdf>
- Wike, R., Silver, L., & Castillo, A. (2019). *Many across the globe are dissatisfied with how democracy is working* (online). Recuperado de <https://www.pewresearch.org/global/2019/04/29/many-across-the-globe-are-dissatisfied-with-how-democracy-is-working/>
- Wong, J. C. (2018). *Congress grills Facebook CEO over data misuse – as it happened* (online). Recuperado de <https://www.theguardian.com/technology/live/2018/apr/10/mark-zuckerberg-testimony-live-congress-facebook-cambridge-analytica>

- Wu, T. (2011). *The Master Switch: The Rise and Fall of Information Empires*. New York: First Vintage Books.
- Yglesias, M. (2019). *Trump's Twitter tirades about treason and civil war reveal a dangerously unfit president* (online). Recuperado de <https://www.vox.com/policy-and-politics/2019/9/30/20891128/trump-tweets-schiff-treason-civil-war>
- YouTube (2019). *Continuing our work to improve recommendations on YouTube* (online). Recuperado de <https://youtube.googleblog.com/2019/01/continuing-our-work-to-improve.html>
- Zollo, F., et al. (2015). Emotional dynamics in the age of misinformation. *PLoS ONE*, 10, 9, e0138740.
- Zuckerberg, M (2021). *Facebook blog post*. (online). Recuperado de <https://www.facebook.com/zuck/posts/10112681480907401>