Escola de Sociologia e Políticas Públicas
Departamento de Sociologia

# *ClarEvidência* :
# O Ónus da Prova na Moderação de Conteúdos e a Sustentabilidade de Comunidades Online

André Teles Fortes

Dissertação submetida como requisito parcial para obtenção do grau de

Mestre em Comunicação, Cultura e Tecnologias de Informação

Orientador:
Professor Doutor Pedro Pereira Neto
Professor Auxiliar Convidado
ISCTE-Instituto Universitário de Lisboa

Outubro, 2019

**AGRADECIMENTOS**

Em primeiro lugar agradeço a "nossa senhora" minha mãe, Fátima (!), companheira de todas as horas e conselheira em todas as etapas que conduziram este processo, com tamanho carinho, respeito, fé, paciência, e apoio incondicional; sem os quais nada disto seria possível.

Um especial agradecimento ao Professor Doutor e meu Orientador, Pedro Pereira Neto, pelo valor inestimável das suas recomendações para este trabalho, e pelo invejável sentido crítico e reflexivo que lhe reconheço (e me inspirou) desde a primeira aula.

Um bem-haja à Professora Doutora e Coordenadora deste mestrado, Joana Azevedo, que desde logo me acompanhou e guiou em matérias sensíveis que transcendem estes trabalhos, sempre com um olhar meigo, uma palavra amiga, e um sorriso contagiável.

Um muito obrigado a todos os moderadores emancipados, que carregaram e partilharam comigo aqui, sem reservas, o fardo da prova para um debate mais esclarecido sobre a sustentabilidade de comunidades online, desta prática e de aqueles que nela irão/continuarão a exercer.

Por último, estendo a minha gratidão aos júris e demais leitores que porventura escolham suspender um entendimento imediato do conteúdo desta dissertação em favor de um compromisso reflexivo com o esclarecimento que pode ainda advir dela (i.e. na defesa).

**RESUMO**

Esta dissertação procura compreender de que forma o Ónus da Prova na moderação de conteúdos online está a ajudar a criar comunidades online mais sustentáveis, testando se é de facto uma questão de forma (i.e. evidência) ou de conteúdo (i.e. clareza). Primeiro, veremos que o Ónus da Prova online difere radicalmente dos típicos julgamentos offline, porque, [1] como o Autor do conteúdo está ausente na tomada de decisão e não tem o "benefício da dúvida" para poder defender-se e fornecer mais contexto, logo, [2] o Moderador tem o seu campo de análise limitado à clareza do contexto apresentado e não consegue construir um caso de defesa ou oposição sólidos para sustentar uma declaração de inocência ou culpa (i.e. Visão-Contentor).

Assim, colocamos o paradoxo da clareza: é o Autor que a articula (concreta) mas é o Moderador que a realiza (subjetiva) e, logo, decide o que se pode ter (in)tentado expressar. No estudo prático que se segue com 1 categoria, 6 conteúdos, e 12 moderadores, veremos como isso exacerba sérias preocupações (mas com potencial para um debate mais esclarecido) sobre o futuro da livre-expressão e do próprio livre-arbítrio, concluindo que, se não devemos concordar num modelo conceptual ou formal que nos permita um entendimento estável do fenómeno de clareza, devemos, porém, questionar os seus usos no trabalho ideológico que ela faz (ou evita) na maneira apropriada com que é articulada na criação e/ou publicação de conteúdos (pelo Autor) ou 'realizada' na análise e tomada de decisão (pelo Moderador).

**PALAVRAS-CHAVE**

Desenvolvimento Sustentável, Liberdade de Expressão, Livre-Arbítrio, Políticas Públicas, Redes Sociais Online, Moderação de Conteúdos, Ónus da Prova, ClarEvidência, Visão-Contentor, Dilemas Estético e Ético

**ABSTRACT**

This dissertation seeks to understand in what way the Burden of Proof in online content moderation is helping to create more sustainable communities online, but argues that it is not so much a matter of form (i.e. evidence) but of content (i.e. clarity). First, we will see that the Burden of Proof online differs radically from typical judgments offline because, [1] as the Author of the content is absent in the decision making process and doesn't have the "benefit of the doubt" to defend himself and provide more context, hence [2] the Moderator has his field of analysis limited to the clarity of the context presented and cannot always build a solid case of defense or opposition to support a declaration of innocence or guilt (i.e. Container-View).

Thus, we put forward a paradox of clarity: it is the Author who articulates it (objective) but it is the Moderator who realizes it (subjective) and then decides what one may have intended to express. In the following case study with 1 category, 6 contents, and 12 moderators, we will see how this exacerbates serious concerns (but with the potential for more enlightened debate) about the future of free speech and free will itself, concluding that, if we should not agree on a conceptual or formal model that allows us a stable understanding of the phenomenon of clarity, we must, nonetheless, question its uses in the ideological work it does (or avoids) in the way in which it is 'articulated' in the creation and/or publication of the content (by the Author) or 'realized' in the analysis and decision making (by the Moderator).

**KEYWORDS**

Sustainable Development, Free Speech, Free Will, Public Policy, Online Social Networks, Content Moderation, Burden of Proof, ClarEvidence, Container-View, Aesthetic and Ethical Dilemmas

# ÍNDICE

## INTRODUÇÃO

### A Grande Ideia e Mito do Progresso, da Idade das Trevas à Idade das Luzes, à atualidade

Quando atentamos estudar as raízes do debate atual sobre o desenvolvimento sustentável de comunidades online para entender de que forma está associado a uma questão mais profunda sobre o fardo/ carga/ ónus da prova na moderação de conteúdos, é incontornável conceber (ainda que brevemente) uma introdução à (r)evolução histórica e cultural da ideia de "progresso" social, cujo corolário mais saliente assevera que a humanidade tem vindo e continuará a mover-se numa direção desejável em função dos avanços científicos, tecnológicos, materiais e morais (Bury, 1920: 2; Nisbet, 1980).

A época tumultuosa pejorativamente designada de "Idade das Trevas" que tende a ser associada a um período – algures entre escuridão intelectual, flagelo moral, ruína cultural e atraso social no progresso – a partir do século IV ou V, só foi superada com necessidade de harmonizar as exigências da fé (na religião) e da razão (na filosofia) entre os séculos XI e XIV. No entanto, a retirada lenta mas inexorável das instituições monásticas e escolásticas no esquema das coisas e um conjunto de transformações – como o Renascimento italiano, os Descobrimentos portugueses, a Reforma Protestante alemã, bem como a invenção da impressora, da transição do feudalismo para o mercantilismo, e do escolasticismo para o humanismo – entre os séculos XIV e XVI, assinalou o fim do período medieval e uma maior preferência pelo conhecimento baseado no sentido crítico e orientado na evidência empírica para conhecer a realidade, outrora sustentada no dogma e na superstição (Cassirer et. al, 1987).

Em todo o caso, até ao século XVII havia ainda pouca distinção entre padres e "pais" da ciência, mas, devido a um entendimento atualizado dos conceitos de "evolução" (Aelianus Tacticus, 1616) e da influência do humanismo na tensão epistemológica entre o Empirismo britânico (baseado na experiência; indução) e o Racionalismo continental (baseado na razão; dedução), e da própria Revolução Científica – com as leis de moção planetária kepleriana (1609-19), as observações telescópicas galileanas (1609/10), o cultivo da dúvida cartesiana (1637/41), o nascimento da mecânica clássica newtoniana (1687), entre outros –, a religião e a ciência seguiram caminhos opostos (sobretudo após o Índex e a Inquisição); embora a crença (religiosa) e teoria (científica) no progresso fossem mais indistinguíveis e inseparáveis que nunca nesta nova "Idade da Razão", e (curiosamente ou não) foi então articulado pela primeira vez pelo cientista francês Fontenelle: A Grande Ideia de Progresso (1683).

A partir do século XVIII essa "ideia" foi sendo progressivamente naturalizada e secularizada na "Idade das Luzes" especialmente com as orientações filosóficas e culturais do Idealismo alemão, do Iluminismo francês (1750-1789) e da própria Revolução Industrial (1760 e 1820/40) que, até ao século seguinte, conduziu a um crescimento drástico e inédito na população, produção, consumo e riqueza, e impulsionou os princípios e teorias da governança, da liberdade civil (Hume, 1742), da separação tripartida dos poderes (Montesquieu, 1748), do contrato social (Hobbes; Locke; e Rosseau, 1762), do mercado livre (Smith, 1776), entre outras matérias na base do liberalismo, conservadorismo, individualismo, etc. que tiveram uma influência seminal na Declaração de Direitos do Homem (1789), nas revoluções americana e francesa, nas reformas abolicionista e sufragista, e nos demais eventos com ideais representativos e democráticos modernos (Ryan, 2012: 25).

No século XIX, a síntese entre o racionalismo e o empirismo de Kant no final do século anterior influenciara inúmeras escolas e autores, e surgira também o pragmatismo americano, o utilitarismo britânico, entre outros, como o positivismo francês do qual se destaca Comte com a sua "lei do progresso" e na aplicação estrita do método científico. É mais ou menos nesta conjuntura que, no século XX, surge a filosofia analítica com um ênfase na "clareza" conceptual e na lógica formal na análise da linguagem (comum), e uma tendência simultaneamente empirista, racionalista e behaviorista. Na virada do século XXI, porém, a experiência já ensinara (a alguns) que, independentemente do crescimento exponencial da ciência e da tecnologia, tanto a condição humana material quanto moral permanecem abertas tanto a um progresso quanto a um "regresso".

A "grande" (e ambiciosa) ideia de progresso, tal como a promessa iluminista de crescimento e desenvolvimento linear e contínuo da condição humana, expôs a sua fraqueza e ficção, perdeu o seu fascínio e encanto, e, até certo ponto, provou exatamente o inverso: o Grande Mito do Progresso (Rachel Carson's ,The silent Spring: 1962; Paul Ehrlich's, The population bomb: 1968). Afinal de contas, essa "ideia" não inspirou apenas a invenção da impressora, a reivindicação de direitos cívicos e políticos, o desenvolvimento das telecomunicações e a adoção de duas agendas mundiais para um desenvolvimento sustentável, mas justificou também a barbárie simbólica das técnicas de manipulação mediática e a censura, o epílogo sangrento da exploração colonial e a ressaca de duas guerras mundiais, a crescente devastação da biosfera e a ameaça constante da fissão nuclear, e tantas outras coisas. Mais felizes? Mais seguros? Mais esclarecidos? Em breve ficaria claro que alguma ponderação (e moderação) poderia desempenhar um papel preponderante para o progresso social.

Se, à data aproximada desta publicação, comprimirmos (sensivelmente) a história remota da humanidade desde o Homo Sapiens num único dia de calendário, a parte atualmente documentada só ocupa as duas últimas horas do dia, quando começam a aparecer as primeiras formas de arte rupestre, a escrita cuneiforme e hieroglífica, os pombos correio, bem como alguns estudos hermenêuticos e da exegese. Todos os estudos, debates, teses e teorias multiparadigmáticas da análise de conteúdos moderna, da comunicação científica sobre todos estes eventos, e do ramo epistémico das ciências da comunicação (da matemática, da cibernética, dos dois-passos, do framing, do gate-keeping, do agenda-setting e muitas outras) ou mesmo das tecnologias da informação, aconteceram somente a um minuto e meio antes do dia terminar, e grande parte foi sobre os efeitos (in)diretos e (i)limitados que os conteúdos têm no público e vice-versa.

Os ponteiros marcam as 23:59:55. Todos os conteúdos que alguma vez escutámos, observámos e criámos na Internet (na sua forma mais popular que acompanha a virada do milénio XXI), serão publicados nos próximos 5 segundos da história da humanidade, e, antes que toquem as badaladas da meia noite a anunciar um novo dia, um moderador está a chegar ao trabalho, puxa a cadeira e prepara-se para moderar conteúdos – desde comentários como "Feliz Aniversário atrasado!", a fotos de nudez com o hashtag #FreeTheNipple, piadas em forma de MEME sobre o 9/11, vídeos de corpos esventrados num acerto de contas por um cartel, etc. Porém, antes que o moderador encontre respostas para o real impacto das suas decisões na sustentabilidade do discurso em comunidades online e pondere o futuro da liberdade (como a experienciamos), esta dissertação não deslizou ainda no canhão do tempo.

Tudo está (prestes) a acontecer, a radical novidade de tudo ficará instantaneamente (mais) clara, mas para entender o que vem a seguir teremos de mudar de escala. Assim, no Capítulo I, começaremos com um breve contexto histórico da análise à moderação de conteúdos, que nos ajudará a situar de onde vimos e para onde vamos com o debate que se segue no enquadramento atual no Capítulo II à luz das obrigações do Estado na regulação e dos deveres das empresas de redes sociais online na moderação. Em seguida, no Capítulo III, começamos com uma leitura aproximada à teoria e prática da moderação de conteúdos, começando com uma questão semântica-quântica do que chamaremos de "ClarEvidência" para problematizar *quanta* evidência e clareza constitui carga analisável de prova, e, logo, com a delimitação do que designaremos de "Visão-Contentor" para percebermos os limites da quantidade e qualidade de evidência e clareza num contexto, e encerramos com alguns dilemas na estética, ética e dialética do esclarecimento para refletir sobre "verdade" e "justiça".

Ainda antes do desfecho desse (sub)capítulo, levantamos algumas questões do debate entre "forma" e "conteúdo" na leitura preferida das políticas, tanto na forma ou expressão literal das palavras dos autores (i.e. "Carta") quanto do conteúdo ou da intenção com estes as proferiram (i.e. "Espírito), e que será útil para conceber o conceito de ideologia que se irá aplicar em seguida na moderação de conteúdos. Neste sentido, no Capítulo IV, explicamos (brevemente) o desenho da pesquisa que precedeu este trabalho, antes de remeter para o estudo prático que a acompanha no Capítulo V e que procurará responder "de que forma o Ónus da Prova na moderação de conteúdos online está a ajudar a criar comunidades online mais sustentáveis?", na premissa de que alguma moderação é efetivamente necessária para um desenvolvimento deste tipo em sociedades modernas. Assim, formuladas as perguntas e hipóteses de resposta no capítulo anterior, procuraremos testar como cada um de 12 analistas moderaria 6 conteúdos de 1 única categoria (nomeadamente, Discurso de Ódio, Cruel e Insensível) e como justificam as suas decisões baseadas na "evidência" e/ou "clareza" do contexto apresentado, terminando com algumas reflexões e considerações sobre o estado atual e o futuro de comunidades online com moderação de conteúdos.

# I. CONTEXTO HISTÓRICO

## 1.1 (Sem) Sombra de Dúvida: da Análise à Moderação de Conteúdos

A análise de conteúdos é talvez melhor entendida como uma ampla família/conjunto de técnicas, uma metodologia não reativa/intrusiva, ou um instrumento polimorfo e polifuncional de pesquisa laboriosa e minuciosa sobre qualquer artefacto de comunicação – imagético, sonoro e/ou textual – com a finalidade de descrever objetiva e sistematicamente (recorrendo a inferências válidas e deduções lógicas) o conteúdo manifesto ou latente na comunicação para identificar significado do discurso popular (Neuman, 1997: 272/3; Bardin, 1977: 9; Weber, 1990: 9; Holsti, 1968: 5; Macnamara, 2003: 6; Cole, 1988).

A análise de conteúdos é uma prática muito antiga com antecedentes históricos nas áreas da hermenêutica e filologia da antiguidade clássica, cujo estudo incidiu, desde o início, na interpretação e análise crítica do discurso simbólico e polissémico na exegese de certos textos bíblicos, literários e "legais" de modo a compreender e descodificar significados, muitos dos quais obscuros e ocultos em parábolas, metáforas, símbolos, sinais e mensagens sagradas, misteriosas e profanas (Bardin, 1977: 14; Campos, 2004: 611).

A partir de 1920, podemos identificar uma fase embrionária da análise de conteúdos nos Estados Unidos da América com o estudo quantitativo de material essencialmente jornalístico pela Escola de Columbia, cujo louvor pela técnica, fascínio pela medida e obsessão pela contagem, se acentuava na justa medida que analisavam o tamanho dos artigos, o grau de sensacionalismo e o enquadramento das páginas (Bardin, 1977: 15). Paralelamente, e com o advento da I Guerra Mundial e a invenção da imprensa e da televisão, o sociólogo Harold Lasswell formulou as questões de partida e exigências centrais de ordem técnica da análise de conteúdo com o estudo da propaganda, que culminou na sua obra de 1927 (Janeira, 1972: 373).

De 1940 até inícios de 1950 – especialmente aquando da II Guerra Mundial –, o estudo da análise de conteúdos acentuou-se e proliferou com um objetivo mais pragmático e de intervenção dos departamentos de ciências políticas dos EUA que, sobretudo motivados pela psicologia experimental objetiva, adotavam uma atitude behaviorista preocupada em subordinar os fenómenos e a introspeção intuitiva ao mínimo quantificável, culminando na primeira publicação com a definição mais célebre da análise de conteúdos e as suas regras preconizadas na obra de Bernard Berelson e Paul Lazarsfeld em 1952 (Bardin, 1977: 15, 16-18; Campos, 2004: 612; Vala, 1986: 101).

As décadas de 1950 e de 1960 marcaram invenções como a cassete de áudio, o disco de vídeo e o videojogo, e assinalaram o apogeu da técnica de análise de conteúdos e dos estudos maioritariamente quantitativos que resumiam e sistematizavam as preocupações e orientações metodológicas, gnosiológicas e epistemológicas desta época; embora esta abordagem fosse alvo de crítica de autores como Siegfried Kracauer que contestou a maneira normativa, restritiva, simplista e reducionista da literatura neste período (Janeira, 1972: 370; Bardin, 1977: 18-20). A querela da dicotomia entre abordagens quantitativas ou qualitativas intensificou-se, e foi também nesta altura que foi organizado o Alberton House Conference (1955) que marcou a mutação e expansão das aplicações de análise de conteúdo a disciplinas muito diversificadas, suscitando novas questões e propondo inúmeras respostas no plano teórico e metodológico (Pool, 1959: 2; Bardin, 1977: 19-21, 114/5; Vala, 1986: 101).

Em 1959 realizou-se o primeiro congresso de analistas no qual A L. George tentou precisar as características de ambos os métodos e, contrariamente ao que se admitia até então, considerou-se que o procedimento na análise de conteúdo não é obrigatoriamente quantitativo como defendia Berelson, Lasswell e Lazarsfel (vulgarmente considerados os pioneiros desta técnica). Contudo, a esmagadora maioria de pesquisadores quantitativos mais proeminentes, como Kimberley Neuendorf, acreditam que a análise qualitativa de "conteúdo" é mais apropriadamente descrita e categorizada como análise retórica, interpretativa, estruturalista, semiótica, do discurso, narrativa, textual (ex: nos estudos hermenêuticos), ou crítica (ex: nos estudos literários) (Neuendorf, 2002: 10, 41; Macnamara, 2003: 15).

Independentemente da definição, distinção e separação entre métodos, esta era menos rígida do que o que se veio a tornar mais tarde (Jensen, 2002: 43) e, enquanto uns defendem que nos aproximamos de uma análise mais qualitativa de conteúdo (Bardin, 1977: 116), outros denunciam a efemeridade das propostas de A.L. George e antecipam o regresso aos métodos unicamente quantitativos na atualidade (Vala, 1986: 102). Se parece não haver consenso nesse aspeto também, facto é que a investigação e a prática na/da análise de conteúdo mudou significativamente desde os anos 60, sobretudo ao atingir uma popularidade renovada em anos recentes graças aos avanços tecnológicos e à aplicação "frutífera" da pesquisa quantitativa na análise de informação sem precedentes nos meios de comunicação de massa.

Entre 1980 e 2000, noticiámos a invenção do telemóvel e da fibra ótica, dos computadores pessoais e portáteis, da WorldWideWeb e da Internet, dos motores de busca e da maior aplicação de protocolos de partilha de ficheiros/ conteúdos das primeiras comunidades online baseadas em linguagem de texto primitivo como nos grupos BBS, MUDs e Usenet que surgiram no final da década de 70. Estes grupos permitiram métodos de deliberação que

envolvessem ativamente utilizadores relevantes na reflexão, discussão e decisão, sobre o tipo de futuro que estes queriam (ajudar ou tentar) criar e sustentar em dada comunidade virtual (Robinson, 2004: 380). Os mecanismos e ferramentas de moderação de conteúdos foi um de tais métodos que ocorreu, desde logo, por intermédio de mão de obra voluntária dos seus utilizadores que, juntos, conectados virtual e "glocalmente", criavam e analisavam os conteúdos entre si e acordavam quais (e se) regras deviam ser definidas e como deviam ser implementadas.

A partir de 2000, a Internet foi amplamente refinada, adotada e valorizada pelos públicos cada vez mais fragmentados, e a intensidade, velocidade e quantidade massiva, inédita e avassaladora de conteúdos gerados por eles só aumentou com o advento dos blogs, wikis, fóruns e, sobretudo, da Web 2.0 e das redes sociais online – como o LinkedIn (2002), Skype (2003), Facebook (2004), Youtube (2005), Twitter (2006), Tumblr (2007/8), Instagram (2010), SnapChat (2011), entre muitas outras. A esmagadora maioria dessas (grandes) empresas baseadas na Internet aperceberam-se que dificilmente poderiam ou conseguiriam arcar com o risco – legal, financeiro e de reputação – que alguns conteúdos UGC "sem filtro" (i.e. não moderados) poderiam causar.

Por isso as empresas apressaram-se em mediar (já que não podiam remediar) o que consideravam ser uma apropriação indevida dos espaços livres para a discussão, opinião e expressão dos membros de certas comunidades que não se regulavam bem a si mesmos (ex.: pornografia infantil, discurso de ódio, ameaças de terrorismo, etc.). Uma de tais estratégias para mitigar eventuais ou potenciais danos às redes sociais online (como Facebook, Google, etc.), foi recorrer a empresas especializadas e terceirizadas (como Accenture, Arvato, Majorel, Cognizant, Competence Call Center, etc.) sediadas a grandes distâncias da sede (desde Silicon Valey e Portugal até às Filipinas ou Índia) com "exércitos" ou "legiões" de "gatekeepers online" (i.e. moderadores) geralmente subcontratados em regime outsourcing.

De facto, e muito embora a Internet tenha surgido sem um órgão de governança centralizado, o processo de globalização que ocorreu desde final do século anterior estendeu a política doméstica a um subconjunto de programas públicos focados em matérias de governança global (como a liberdade versus censura online) e isto animou um punhado de entidades estatais a implementar leis ou políticas "públicas" que, direta ou indiretamente, procuram regular o "Estado online" sobretudo na análise e moderação dos conteúdos; como na negação do Holocausto (ex: Alemanha), pornografia geral (ex: Marrocos), protesto político (ex: Vietnam), discurso de ódio (ex: Estados Unidos), comunidade LGBT (ex: Qatar), terminação de gravidez (ex: Kuwait), curso eleitoral (ex: Zimbabwe), entre outros (Freedom on the Net, 2018; Citizen Lab, 2018).

No entanto, exceto as extensas contribuições e debates profícuos no âmbito do gatekeeping online (e talvez até do framing), a temática da moderação de conteúdos é ainda excessivamente subvalorizada e só atingiu algum mediatismo recentemente devido aos efeitos psicológicos que surte em alguns moderadores, ou dos efeitos das decisões dos moderadores nas liberdades do público em geral (Roberts, 2014: 1-169); mas absolutamente nada está a ser dito sobre o ónus da prova na prática e o que significa para a maneira como experienciamos a nossa liberdade de expressão e livre-arbítrio (como veremos).

A estes analistas foi (e ainda é) delegada a tamanha tarefa de gerir as interações entre os utilizadores e aplicar políticas "públicas" (dos clientes) que filtrem certos conteúdos (i.e. mantendo ou removendo-os) para salvaguardar a sustentabilidade do "Estado online" (Madrigal, 2018; Wray, 2018). Atualmente, num único minuto, estamos a falar de 1 milhão de utilizadores no Facebook, 4,5 milhões de visualizações no Youtube, 3,8 milhões de pesquisas no Google, 46,200 de novos conteúdos no Instagram, 87,500 no Twitter, 2,1 milhões no Snapchat, 188 milhões de emails; etc (Statista, 2019; Desjardins, 2019; etc.). São 4,437 milhões os utilizadores da Internet dos quais 3,499 milhões utilizam redes sociais online, ou seja, 58% da população mundial usa a Internet e 58% usa essas redes de algum modo (Kemp, 2019: 6, 9).

Assim, pela primeira vez na história, temos uma excelente oportunidade para escutar, observar, experimentar e aprender uns com os outros, mas também nunca foi tão difícil gerir tais interações com eficácia e aplicar políticas que muitas vezes adotam uma visão simplificada da linguagem que ignora a complexidade da semiose e limita o escopo da análise de conteúdo à medição e à contagem, sem refletir critica e reflexivamente sobre a sua adequação à ciência e discurso sociais. Em todo o caso, qualquer conhecimento que este trabalho almeja construir, torna-se, lamentavelmente, mais falível e desafiante na medida que procura explorar a sustentabilidade do discurso online cuja generalidade de fontes teóricas sobre o assunto está ainda na sua infância, para não dizer que rigorosamente nenhuma aborda o Ónus da Prova e um tópico que suspeito estar no cerne da questão: a ClarEvidência (que abordaremos mais adiante).

## II. ENQUADRAMENTO ATUAL

### 2.1 Obrigações do Estado: preocupações com legalidade e regulação

A Lei Internacional dos Direitos Humanos impõe obrigações ao Estado para assegurar, simultaneamente, ambientes favoráveis à proteção do exercício do direito da liberdade e diversidade de opiniões, mas também na promoção do acesso à informação pelos cidadãos. Não obstante as restrições estatais à liberdade de expressão devam atender unicamente a condições bem estabelecidas no quadro legal fornecido por lei para serem consideradas legítimas para a interferência do Estado, este pode (e deve), porém, interferir para proteger os interesses enumerados à luz do artigo 19 aquando da colisão com os direitos/ reputações/ moral de outros, e em matérias de privacidade, segurança nacional, ordem, saúde (e pudor) públicos; que tornam o apelo à regulação mais compreensível e "apetecível" (UNGA, 2018: 4).

Desde o início da era digital, muitos Estados adotaram um regime legal de uma espécie de "imunidade" para proteger os provedores de serviços interativos (como as empresas de redes sociais online) da responsabilidade dos conteúdos publicados por terceiros nas suas plataformas, exceto, claro, quando estes intermediários excedem o seu papel como um mero canal ou meio de comunicação dirigido aos internautas (ex: Lei da Decência das Comunicações dos Estados Unidos em 1996 e a Diretiva do Comércio Eletrónico da União Europeia em 2000). Mas esta é uma área cinzenta: Afinal de contas, o que define um meio de comunicação, para que serve e que regime de responsabilidade tem de ter?

No entanto, atualmente, a esmagadora maioria dos Estados não exige somente que as empresas limitem conteúdo manifesta ou expressamente ilegal (ex.: representações de abuso sexual infantil, ameaças diretas e críveis de dano e incitamento à violência), mas muitas vezes vão além disso para censurar ou criminalizar a polissemia do discurso social legítimo com leis restritivas ampla e escrupulosamente redigidas sobre o que define extremismo, blasfémia, difamação, ou discurso ofensivo, cruel, insensível, como pretextos para moldar o ambiente regulatório online e, logo, o exercício geral da liberdade de opinião, expressão e associação (e mesmo quando o fazem inadvertidamente, não é inocente).

Por outro lado, existe ainda a particularidade das ferramentas de desinformação e propaganda estatais serem, por vezes, deliberadamente concebidas para limitar o escopo da acessibilidade e a confiabilidade nas empresas de media independentes, naquelas instâncias em que os Estados as obrigam a denunciar ou escalar algumas violações às autoridades e a restringir ou proibir conteúdo *a priori* sob critérios jurídicos vagos ou complexos sem revisão judicial prévia e com aplicação de duras e pesadas penalidades no caso de incumprimento (VER Lei da Cibersegurança chinesa em 2016) (PEN, 2018: 21).

Paralelamente, as empresas distinguem entre pedidos de remoção de conteúdo supostamente ilegal enviado por meio de solicitações do público geral com base nos Termos de Serviço e Utilização que se aplicam "globalmente", e pedidos de entidades estatais ou legais regulares que se aplicam "localmente" (ex: governos ou jurisdições solicitantes) – conquanto, cada vez mais, estas entidades formalizem pedidos alheios aos procedimentos legais e judiciais regulares para o efeito e, algumas, chegam mesmo a estabelecer unidades governamentais especializadas para encaminhar os pedidos de remoção extraterritorial de conteúdos diretamente às empresas que não são necessariamente baseados na lei nacional ou local em vigor (ex: a Unidade de Referência da Internet da União Europeia) (UNGA, 2018: 8). Estas questões exacerbam sérias preocupações sobre a censura de informação dentro/além fronteiras.

Apesar do aumento e da demanda global por revisões e remoções rápidas e automáticas em prole da eficiência governamental não ter transparência suficiente ou caráter vinculativo, também tende a resultar em penalidades substanciais para as empresas na inconformidade com os prazos estabelecidos, e, sobretudo, a arriscar formas de restrição prévias capazes de minar a liberdade de expressão dos públicos. Por exemplo, os prazos curtos para a remoção de conteúdos nas leis de imposição de redes e leis do processo eleitoral, como a entrada da rede alemã NetzDG em 2017 e as eleições quenianas em 2016 (Freedom on The Net, 2018: 13).

Mas, acima de tudo, parece que alguns Estados tendem a evitar assumir responsabilidade quando delegam funções reguladoras a atores privados que muitas vezes carecem de ferramentas e conhecimentos específicos e básicos para a gestão de comunidades e políticas de moderação de conteúdos online em matérias complexas de facto e direito, especialmente quando colocam uma pressão significativa sobre as empresas cujos processos podem ser adversos aos quadros legais, códigos deontológicos e padrões éticos vigentes que de outra forma seriam considerados com a devida diligência na supervisão legislativa e outros mecanismos de responsabilização.

## 2.2 Deveres das Empresas: preocupações com responsabilidade e moderação

Na medida em que algumas das maiores empresas de redes sociais online reivindicam, direta ou indiretamente, um papel cívico preponderante na vida, ordem e segurança públicas, há naturalmente um "apelo" para a adoção e implementação de princípios dos direitos humanos nas suas operações. No entanto, só uma pequena minoria de redes sociais online aplicam estes princípios e, dessas, a maioria é cética em como esses direitos são vinculados pelos mesmos governos que as pressionam em tomar medidas menos ortodoxas de moderação. Porém, os Princípios Orientadores sobre Empresas e Direitos Humanos estabelecem padrões globais de expectativas de conduta empresarial que são/ ou devem ser aplicados em todas as atividades das empresas, e sempre quando e onde estas operem.

Uma das iniciativas para minimizar o impacto e mitigar o risco que apoia as empresas a enfrentar estes desafios é a Global Network Initiative que, entre outras recomendações, defende medidas de maior transparência (UNGA, 2018: 9). Com os anos, muitas empresas têm desenvolvido e publicado relatórios de transparência anuais com o número de solicitações e justificações governamentais para manter ou remover certos conteúdos online; conquanto não relatem ou divulguem de modo "transparente" informações suficientes sobre como processam estes pedidos de acordo com os mecanismos, regras e termos de serviço que são formulados e executados internamente (i.e vedados ao público). Se isso está bem ou mal não é o tema deste trabalho, até porque qualquer análise restrita a uma única política num período limitado torna-se rapidamente obsoleta uma vez que as políticas internas estão em constante mutação.

De qualquer forma, atualmente os padrões da comunidade publicados (i.e. visíveis ao público) nos mais variados canais (ex: rede social e blog da empresa, etc.) são redigidos em termos tão gerais que têm sido alvo de uma crescente crítica e escrutínio públicos, sobretudo expressos pela imprensa e por organizações da sociedade civil, num apelo maior para esclarecer e clarificar algumas políticas e exemplos hipotéticos da prática de moderação, embora a lista não pareça suficientemente exaustiva para aplacar qualquer suspeição ou dúvida razoável.

As políticas internas mais detalhadas de moderação são um subconjunto desses termos levados a cabo por equipas internas ou externas de "confiança e segurança" em matérias específicas como spam e fraude ou terrorismo e crime organizado. Estas políticas tendem a ser reativas porque são constantemente clarificadas, mas normalmente variam consoante uma série de outros fatores, desde o mercado- e público-alvo até à missão, valores, objetivos e tolerância ao risco da empresa (UNGA, 2018: 10, 13). O compromisso ou conformidade legal e com os direitos humanos tende a ser particularmente complicado quando a legislação do Estado é vaga

e sujeita a uma miríade de interpretações por vezes insensíveis às vicissitudes do discurso do dia a dia.

Entre os processos e ferramentas de filtragem e deteção/denúncia de abusos nos conteúdos destacam-se a 'sinalização automática' por algoritmos e a 'sinalização manual' por membros do público ou das próprias empresas. O aumento exponencial de volume de conteúdos gerados pelos usuários impulsionou a criação e implementação de ferramentas de moderação automatizadas por inteligência artificial para detetar, sinalizar ou mesmo filtrar conteúdos mais clarividentes (ex: através de algoritmos de correspondência e de processamento de linguagem natural) para uma decisão automática ou para revisão humana *a posteriori* (UNGA, 2018: 12). Em muitos casos esta automação é efetivamente útil na identificação de conteúdos como pornografia e violência gráfica, mas normalmente é necessária uma ponderação humana mais profunda da clareza do contexto e do significado que requer o espírito crítico e reflexivo do analista.

Por outro lado, a sinalização manual permite que membros do público, das próprias empresas ou de entidades públicas, solicitem ou reclamem a remoção de conteúdos de uma plataforma. Várias empresas cooperam entre si, com órgãos reguladores (ex: O Código de Conduta da União Europeia em 2016) e com um conjunto mais restrito de usuários especializados (i.e. de com grande impacto/ influência) como fontes confiáveis em diversas matérias específicas para sinalizar conteúdos (embora não pareça haver evidência alguma sobre os requisitos ou critérios na seleção destes "sinalizadores de confiança", nem o repertório das suas interpretações de documentos jurídicos, nem o escopo da sua análise dos padrões comunitários, nem o alcance real das suas recomendações).

A falta de contexto é, contudo, talvez a maior preocupação e, até certo ponto, parece perfeitamente razoável que as empresas e o público enfatizem a sua importância ao ponderar restrições específicas naqueles casos em que a política é apenas parcialmente informada por evidência e vice versa. Uns defendem que a atenção ao contexto não evitou já a sinalização e remoção de algumas representações (ex: de nudez ou de conflitos armados), com valor histórico, artístico, cultural ou educacional. Outros defendem que isso se deve somente à falta de recursos de tempo e dinheiro (i.e. os analistas são pressionados a tomar decisões rápidas) e recursos humanos (i.e. os analistas nem sempre são dotados em literacia digital ou nuances linguísticas e culturais; muito menos os algoritmos automáticos de correspondência de inteligência artificial). Até certo ponto todos estão certos, mas é algures aí que os autores chegam a um impasse (conforme veremos a seguir).

## III. FARDO DA PROVA: *QUANTA* CLAREVIDÊNCIA?

### 3.1 Uma teoria: Semântica-Quântica da Evidência

A tradução da palavra "clarividente" só tem um significado equivalente num conjunto limitado de idiomas e tem obviamente uma coloração mística que agride violentamente a sensibilidade moderna e a nossa atual conceção tão atrelada ao bom senso e à ciência mecânica clássica do que constitui "evidência" e "vidência". Para agitar as coisas mais ainda, traduções noutros idiomas variam desde "claro" (em grego) até "transparente" (em persa) e "obscuro" (em hebraico). Contudo, "clarividente" em português, pode significar tanto uma visão material do que se evidencia de forma clara (no sentido figurativo daquilo que é "evidente") como uma visão espiritual do que só é claro para alguns (no sentido próprio daquele que é "vidente").

Ambos sentidos diferem radicalmente entre si (ou não) em função do que que se manifesta com maior ou menor grau de clareza e, como tendemos a provar e concordar cientificamente somente naquilo que se manifesta, tendemos por isso a associar "clareza" com "evidência" e ambas com a "realidade". O termo "evidência" (do latin *evidentia* que significa "que se vê bem e distintamente") é pensado como sendo atributo de tudo aquilo que não suscita ou dá margem à dúvida por ser perfeitamente inteligível, inequívoco e incontestável na existência de algo.

Por outro lado, "clareza", apesar das frequentes invocações do termo pelas disciplinas da retórica e da composição, resiste a uma definição porque não tem um significado único mas sim toda uma série de palavras elusivas com múltiplas definições concorrentes e sentidos copiosos, que são frequentemente invocadas e incarnadas como sinónimos do que é "claro" em diferentes contextos; entre as mais comumente utilizadas destacam-se: (estilo) "simples", "realista", "verdadeiro", "nítido" e, sobretudo, "transparente". Todas elas tornam a definição de clareza ainda menos clara e menos arrumada que uma tabela de políticas ou padrões de comunidade pode (querer) fazê-las parecer, e tornam um modelo conceptual, universal e descontextualizado de clareza, num logro.

Ao invés de arriscar aquilo que pode ser interpretado como uma incursão nos campos da paraciência ou da ficção científica, quero desde já clarificar que aquilo que começou com uma analogia ou metáfora útil rapidamente apontou para uma investigação e especulação científica fundamentada do fenómeno que chamarei antes de "ClarEvidência" para problematizar (provando e refutando) *quanta* evidência e clareza se manifesta com carga suficiente e analisável de prova na moderação de conteúdos, e o paradoxo nesta afirmação.

### 3.2 Dois limites: Visão-Contentor do Contexto

Antes de começar, há dois ditames naturais e clássicos para qualquer Ónus da Prova na análise de conteúdo online [1] o Analista só tem visibilidade sobre um (conjunto limitado de) conteúdo(s) que foi denunciado, e [2] o Autor desse conteúdo não tem o 'benefício da dúvida' para poder defender-se e fornecer mais contexto porque ele, simplesmente, está ausente no ato de moderação. É precisamente aqui que os julgamentos online diferem dos típicos procedimentos legais offline, porque não existe um Ónus da Prova completo (i.e. não existe a implícita necessidade ou obrigação de um proponente [i.e. o Autor], de consubstanciar o significado da sua disposição/expressão com uma prova que a sustente).

Com efeito, como nem o Autor do conteúdo nem o Autor da política da empresa estão presentes no ato da tomada de decisão pelo Moderador, é impossível construir um caso de defesa ou oposição sólidos, e fornecer ou receber evidência com clareza suficiente do contexto para sustentar uma declaração de inocência ou culpa. Logo, o Moderador tem aquilo que chamaremos de Visão-Contentor (ver exemplos no Anexo A), ou seja, o seu campo de observação é limitado à quantidade e qualidade do contexto que está a analisar, e tudo o resto está fora do escopo de análise e seria mera suposição.

No entanto, em alguns conteúdos, o contexto é tão restrito (e a evidência e clareza insuficientes) que não permite ser significante para tomar uma decisão mais "certa" e, noutros conteúdos, o contexto é tão abrangente na abundância de evidência e clareza que comunga do mesmo resultado e pode degenerar para uma falência semelhante de sentido. Logo, como nem sempre a maior ou menor quantidade de contexto da evidência ou clareza é significante para tomar uma decisão "certa", isto só arranha a superfície de um problema maior e mais profundo sobre a qualidade do contexto apresentado, e é algures aqui que suspeito que a maioria dos autores falha no ponto:

Independentemente da quantidade de contexto, a evidência tem de ser clara, e essa clareza tem de ser evidente. A contradição está exatamente aí: a evidência da expressão do Autor nem sempre clara, e a clareza da sua intenção nem sempre é evidente para o Moderador. Isto gera um paradoxo e um aforisma do tipo "Clareza da Evidência não é Evidência de Clareza" porque uma pode existir sem a outra, e isto contradiz a lógica e intuição comuns do que define a realidade do que se está a analisar. Existe de facto aqueles conteúdos em que 'uma violação é evidente mas não é clara', mas a maior questão parece ser sobre aqueles casos em que 'a violação é clara mas não é evidente'.

### 3.3 Três dilemas: Estética, Ética e Dialética do Esclarecimento

É importante reter que no dia a dia há uma omissão ou exclusão natural e necessária na economia linguística (e talvez da própria perceção) como ponto de partida para um discurso claro (especialmente para audiências específicas). Isto porque o Autor escolhe uma entre infinitas possibilidades de expressar informações relevantes e significantes sobre uma mesma coisa para ser entendido, e o Recetor procura/deve procurar a interpretação mais fiel com que o Autor teve a liberdade de expressar-se.

Neste sentido é importante clarificar que "clareza" é simultaneamente uma propriedade da transição do significado inalterável entre o Orador-Signo-Público, e uma qualidade condicional, contextual e retoricamente contingente que depende da relação única entre eles e das possibilidades de interpretação que um discurso oferece (i.e. o discurso do Autor é "claro" se o significado na comunicação do Signo não mudou ou perdeu no processo de transmissão ao Público para o efeito).

Como tal, não devemos (e provavelmente não podemos nem conseguimos) apontar qualidades formais particulares que nos permitam um entendimento estável e concreto do que constitui o fenómeno da clareza. Em contrapartida, podemos dizer que "clareza" não é: (1) uma propriedade apenas da articulação de pensamentos ou palavras (pois se fosse, o Recetor não teria o privilégio de julgar como foi formulado), nem é (2) uma propriedade apenas da compreensão dos pensamentos ou palavras (pois se fosse, o Autor não teria o privilégio de reformular como foi/gostava de ser entendido).

Em todo o caso, ainda que o Autor tente ser tão claro quanto possível, se a evidência não for clara para o Recetor então tudo o que o Autor pode fazer é tentar reformular de forma em que seja entendido o significado do que está a tentar comunicar ao Recetor (ler Metafísica da Clareza no trabalho de Mccumber, 2003: 58). Isso quer dizer também que a articulação de clareza cabe ao Autor do conteúdo, mas a realização em si recai no Recetor (i.e. público), e isto contradiz o senso comum sobre a exata localidade do que é "claro" (e sem saber onde reside, dificilmente sabemos justificar como lá chegámos).

Ora se assumirmos que o Autor de um conteúdo tem tanta responsabilidade na articulação de clareza como o Recetor na realização dela, não podemos esquecer, porém, que a imputação de clareza pelo Autor tende a ser concreta mas a interpretação em si pelo Recetor é subjetiva. Isto significa que, desde logo, há uma profunda associação entre clareza e uma tendência estética tanto ao imputar como ao realizar valores de verdade numa forma/objeto do discurso para tentar corresponder à realidade compartilhada entre Autor-Recetor (McCumber, 2003: 60).

De facto, muitas vezes preferimos que o nosso discurso seja aplicado/ interpretado de forma criativa em vez de ser meramente repetido ou reproduzido com exatidão, mesmo que tal criatividade na sua receção seja sujeita a algum grau de mal-entendido também. Há algo de feio, ríspido, pobre e brutal na prática discursiva meramente literal e funcional, porque as coisas na nossa alma não são a semelhança ou correspondência exata do esquema das coisas, mas aproximações mais ou menos distorcidas.

Por um lado, se a clarificação diária das políticas internas fossem tornadas realmente "públicas", podíamos estar (ainda que irrefletidamente) a estipular o tipo de estética que é tolerado e, gradualmente, a dissuadir o público de expressar livremente o seu estilo a um ponto de tal forma neutro, estanque e mecânico, que podia inclusive tornar-se verdadeiramente impossível distinguir a beleza do conteúdo da brutalidade da forma com que teve a liberdade de expressar-se; e esta estética dita dura, não é muito diferente de uma ditadura estilística.

Em suma, ao exigir do Autor um discurso transparente e visível (i.e. mais claro) para ser entendido pelos demais utilizadores, o Autor naturalmente anula ou omite partes do discurso e, com efeito, pode paralisar as possibilidades e oportunidades do público de interpretar, interrogar e contestar classes ou categorias inteiras de informações potencialmente úteis e críticas que podem estar a ser omitidas ou excluídas. O irónico desta abordagem é que a clareza de uma coisa pode obscurecer outras ao mesmo tempo e que, o discurso "claro" que aparenta, à partida, ser sobre o público, pode tornar-se sobre removê-lo de interrogar e carregar a prova.

Por outro lado, como o grau de clareza varia e os mal-entendidos sobre o "estilo" são mais regra que exceção, cada Moderador define o que é claro para si e de que forma poderá constituir uma violação. É aqui que as coisas se tornam particularmente intrigantes, porque enquanto remover um conteúdo baseado na evidência de uma violação por política é mais justificável, remover um conteúdo baseado na clareza do Moderador (ou qualquer Recetor) tende a ser dificilmente aceitável. Afinal de contas, a evidência é perfeitamente inteligível e não suscita sombra de dúvida, e a clareza é uma qualidade contingente do significado apreendido por um Recetor.

Embora o princípio da correspondência e da contingência do estilo "claro" pareça fundamental no dia a dia, há um certo moralismo de maior clareza na linguagem que tende a assombrar todo o discurso moderno que equivale maior clareza com "verdade" e "correção", e ambos com "justiça"; embora não haja necessariamente conexão inerente alguma. Apropriando as palavras de Richards Heuer, um moderador pode orientar um argumento mais claro e persuasivo mesmo apoiando um julgamento e análise erróneos (Heuer, 1999: 178).

Assim, é manter ou remover um discurso somente com base na realização de clareza do Moderador sobre um estilo, que pode ser eticamente suspeito. Aliás, de um modo geral, esta tendência estética sobre o que é "verdadeiro ou falso" para o Autor que se apressa em remover um conteúdo (sobre o que constitui, na realidade, um contexto maior de humor, tradição, cultura, arte, protesto, literatura, etc), estende-se rapidamente a uma preferência ética sobre o que está "certo ou errado" para si.

Isto porque, o raciocínio e o julgamento do Moderador sobre manter ou remover determinado conteúdo, geralmente diz mais sobre o conteúdo das suas crenças, valores, desejos, preconceitos, perceções e opiniões a respeito de algumas das questões mais contenciosas e controversas da vida em sociedade (como a terminação voluntária da gravidez e suicídio assistido, etc.) do que do conteúdo outro que é suposto estar a analisar.

### 3.3.1 Políticas: "Carta" ou "Espírito" da Política?

Mas isto pode tornar-se ainda mais grave quando o moderador simplesmente distingue se um conteúdo é (anti)estético e (anti)ético em função da clareza ser empreendida intencionalmente ou não. Como vimos, tipicamente o ónus da prova recai no Autor mas, na moderação de conteúdos online, o fardo é transferido automaticamente para o Moderador e não diretamente do Autor para o Recetor/público (i.e. é o Moderador que decide, como se fosse o Autor, se determinado conteúdo é/será suficientemente claro na eventualidade de ser publicado e posteriormente interpretado pelo Público, ver Anexo B).

Portanto, o Autor do conteúdo não está presente no julgamento, não tem o benefício da dúvida para se defender, e raramente pode/consegue apelar para clarificar a razão porque o publicou de antemão; e o Autor da política idem. Assim, o moderador confronta a 'análise do conteúdo do Autor do público' com a 'análise do conteúdo do Autor da política' e, quando o Moderador obedece somente à "carta da política" está a obedecer à expressão literal da política, mas, se por outro lado, obedece ao "espírito da política" está essencialmente a (tentar) seguir a intenção com que tanto a forma como o conteúdo da política foram formulados de antemão pelos seus autores.

O desenho da Carta de políticas na moderação de conteúdos geralmente é realizado consoante Políticas Baseadas em Evidência (PBE), ou seja, são políticas públicas informadas e redigidas mediante a precedência de evidência objetiva, rigorosa e escrupulosamente estabelecida dos conteúdos do público. Mas como o discurso social é sofisticado e existem tantos ou mais casos em que (como vimos na visão-contentor) a prova daquilo que pode constituir uma violação está indefinidamente suspensa "no ar" porque não é clarividente, torna-

se realmente difícil definir se algo 'realmente é' o que 'aparenta ser' ou 'pode ser' e justificar conteúdos "malformados" com políticas "bem-formadas".

Por isso, muitas vezes o Moderador realiza clareza em função do que pensa que o Autor tentou comunicar com determinado conteúdo, ou seja, a sua intenção. O problema clássico da (não) intencionalidade é vastamente discutido em diversas áreas científicas desde a física à comunicação, mas há uma tremenda dificuldade em avaliá-la em qualquer discurso (mesmo aqueles mais clarividentes em que a evidência da expressão e a clareza da intenção do Autor são mais explícitos, e esses até os algoritmos automáticos de inteligência artificial podem provar-se bastante úteis em detetar).

Mas naqueles em que isso não acontece, a tendência estética e preferência ética do que é claro para o Moderador, deixam algumas dúvidas sobre se qualquer tomada de decisão é realmente válida. Afinal de contas, uma falha de clareza do Autor não é necessariamente uma falha estética ou ética (pelo menos não ainda, e nem sempre). Muitas vezes a falta de clareza do conteúdo do Autor é propositada (ex: para ter piada) ou o resultado de uma articulação incompleta e/ou aleatória de pensamentos e sem carga manifesta e suficiente de intenção/propósito de ambiguidade numa tentativa deliberada de escapar à crítica.

Isto é, a falta de clareza não é necessariamente intencional ou propositada porque pode ser uma questão de lapso, preguiça, literacia, etc. Não obstante, muitos cientistas políticos e especialistas em políticas públicas e em comunicação, entendem que o discurso desajeitado ou descuidado de clareza (i.e. obscuro/ambíguo) é o tipo de linguagem de exclusão que uma democracia não pode tolerar; e é mormente criticado como obscurantista. Contudo, frequentemente se olvidam que, às vezes, a ambiguidade do discurso burocrático das próprias políticas é um alibi que parece servir finalidades semelhantes (éticas, estratégicas ou retóricas), e às vezes pela mesma razão (i.e. falta de jeito). Embora haja, naturalmente, uma maior responsabilidade dos formuladores de políticas serem isentos.

A incerteza ou dúvida é um estado de situações epistémicas (i.e. do conhecimento) em que ambientes estocásticos ou parcialmente observáveis exibem informações imperfeitas, limitadas ou desconhecidas, tornando praticamente impossível descrever o estado existente com exatidão/certeza. A ambiguidade/obscuridade (i.e. falta de clareza) é a incerteza sobre se alguma "evidência" tem um terreno comum com um limite "claro", mas o nosso conhecimento sobre onde existe esse limite é incompleto porque depende do tipo de informação partilhado entre o Moderador e o Autor, e é difícil saber que assunções são partilhadas em comum entre eles com completa precisão (sobretudo na moderação, porque o Autor está ausente).

### 3.3.2 Ideologia: "Evidência do quê?" e "Claro para quem"?

Ao mesmo tempo, parece que um grau construtivo de ambiguidade é uma condição importante na difusão do conhecimento à medida que crescemos e aprendemos (desde a metafísica clássica até à filosofia moderna e à análise de conteúdos contemporânea), e alguma ambiguidade desse tipo não denota necessariamente falta de significância (McCumber, 2003: 64). Aliás, às vezes é precisamente onde as coisas são difíceis de entender e indistintas, mas importantes, que os conceitos polivalentes, instáveis e contestados têm maior probabilidade de emergir, pelo que a flexibilidade na definição e interpretação do discurso pode atender às mais variadas necessidades particulares que de outro modo não seriam atendidas.

Não obstante, tende a existir uma insistência e exigência, social e cultural, generalizada e predominante, de que "transparência" é uma condição indispensável e imprescindível para um discurso verdadeiro e correto, mas alguns autores questionam se essa pressão reflete senão o desejo do domínio imutável e inalterável das realidades e possibilidades às quais as nossas palavras devem permanecer semelhantes conforme a lei da preservação e conservação da forma linguística (McCumber, 2003: 64). Se assim for, exigir que os utilizadores sejam cada vez mais claros no discurso, pode inclusive evitar um desenvolvimento mais sustentável e natural da maneira como comunicamos, pensamos, experienciamos ou experimentamos a nossa liberdade, verdade, justiça e outros valores e conceitos abstratos.

De facto, um discurso opaco não reflete coisa alguma, mas um discurso transparente não se reflete a si próprio, ou algo em si mesmo. O discurso que somente se apresenta como "transparente" é o discurso hierárquico de autoridade na nossa cultura que "expressa sem expressar", e é ainda mais retoricamente eficaz e persuasivo quando obscurece ou ofusca os seus propósitos reais com um grau de "clareza" que é quase tão óbvio que a alternativa de questionar os seus próprios usos parece indigesta (Kreuter, 2013: 11). Por isso, quando falamos de evidência e clareza estamos realmente a falar sobre a definição/ versão de clareza da autoridade em questão, pelo que devemos perguntar sempre: Evidência do quê? Claro para quem?

Na moderação de conteúdos este apelo por mais clareza representa aquilo que Kreuter de alguma forma defendeu como sendo uma tendência estética e uma preferência ética da política conservadora e liberal, e uma epistemologia positivista e racionalista que, desde o século XX, procura tornar o discurso transparente ou invisível para conhecer uma realidade subjacente definitiva, imediata, literal, essencial da intenção das massas (Kreuter: 3,5). Contudo, ao limpar o discurso mediado e pautado nos detalhes, nuances e contexto, ele torna-se igualmente menos verdadeiro e preciso no trabalho ideológico, paralisando o processo de

criatividade, inovação e desenvolvimento, não só da riqueza da linguagem mas do pensamento também.

Por isso, Kreuter menciona que exemplos da estética ou ética, da clareza ou obscurecimento nos conteúdos, não podem ser descobertos por suas qualidades formais, mas apenas pelo trabalho ideológico que eles fazem, ou que evitam (porque muitos apenas fingem ser neutros) (4). O ideal programático de um vocabulário/ linguagem puramente neutros e transparentes, livre de ponderações emocionais, tenta fazer uma totalidade de um fragmento (i.e. obscurecendo, ofuscando ou simplificando partes ou fragmentos importantes de uma situação retórica maior, representando-os como um todo) (7). E isto é particularmente importante para o papel da ciência, para a sustentabilidade do discurso e, logo, e para o futuro da liberdade, não só da 'expressão' mas de 'intenção' também (como veremos neste estudo).

## IV. DESENHO DE PESQUISA

### 4.1 Pergunta de Partida e Objeto Empírico

A questão de partida que se coloca com este trabalho é "de que forma o Ónus da Prova na moderação de conteúdos online está a ajudar a criar comunidades online mais sustentáveis?", na premissa de que alguma moderação é efetivamente necessária para um desenvolvimento mais sustentável em sociedades modernas. Posto isto, o objeto empírico escolhido para esta dissertação é como um grupo de analistas moderaria alguns conteúdos de uma dada categoria para ajudar a alcançar maior sustentabilidade de uma comunidade online hipotética; porém o estudo inicial incluiu ainda outros conteúdos e categorias não incluídos neste trabalho final.

Para efeitos de simplificação e de cumprimento com o limite de páginas preestabelecido para este trabalho, optou-se por incluir aqui apenas 1 categoria designadamente 'Discurso de Ódio, Cruel e Insensível' (porque ajuda a testar melhor algumas questões estéticas e éticas na moderação), com 6 conteúdos em 'formato imagem com texto/símbolos' criados propositadamente para a Internet ou não (porque ajuda a testar melhor a prova daquilo que constitui "evidência" com "clareza" suficiente de um "contexto" de arte, humor, tradição, etc.), a 12 moderadores (porque ajuda a testar melhor se há alguma coerência conceptual de clareza, e alguns usos no trabalho ideológico que ela faz ou evita na sua interpretação pelo analista).

A triagem dos conteúdos foi realizada de acordo com o que a maioria das empresas de redes sociais considera, à data, 'grupos com características comuns protegidas' (definidos como sexo, género, orientação sexual, raça, etnia, nacionalidade, afiliação religiosa, deficiência e doença graves) e ataques contra estes (definidos como qualquer discurso violento de apoio à morte, doença ou dano; qualquer incentivo à exclusão, segregação ou discriminação; qualquer declaração de desprezo, repulsa ou inferioridade [física, mental ou moral]; qualquer comparação desumanizadora [ex.: animais, fezes, bactéria, monstros, criminosos, etc.]; ou qualquer representação ou promoção de organizações, líderes ou símbolos de ódio).

A seleção dos moderadores (que são também utilizadores de redes sociais online) foi orientada para aqueles que partilham também algumas dessas 'características protegidas' no seu conjunto. Por exemplo, este estudo captou 50% de ambos sexos (masculino e feminino) em participantes com idade superior a 18 anos e oriundos de várias nacionalidades – Singapura, Dinamarca, Índia, Brasil, Portugal, Indonésia, Espanha, Bélgica, Suíça – com diferentes posições – Subject-Matter Expert (SME), Líder de Equipa, Analista Sénior (em conteúdos e políticas) – em vários clientes – Google, Youtube, Facebook, Instagram, LinkedIn – ou em várias empresas por eles subcontratadas – Accenture, Arvato/Majorel, Cognizant, Genpact – com escolaridade mínima, mas (curiosamente ou não) nenhum é formado em análise conteúdos.

### 4.2 Definição e Adequação do Método

A pesquisa e análise qualitativa é um dos vários métodos de investigação de base linguístico-semiótica atualmente disponíveis em ciências sociais para reduzir dados a conceitos que descrevam fenómenos de pesquisa (Cavanagh, 1997; Elo & Kyngäs, 2008; Hsieh & Shannon, 2005) e que incorporem questões sobre o provável significado e intencionalidade (manifesto e latente) próprio dos atos e comportamentos humanos. Isto inclui as construções e transformações significativas das suas relações, representações, motivações, perceções, opiniões, crenças e estruturas sociais, por sua vez resultado da conceção e interpretação única de como cada cidadão vive, sente, pensa e concebe o mundo (Schreier, 2012; Bardin, 1977; Della Porta e Keating, 2008).

Como tal, o foco do método qualitativo (e deste estudo) não é examinar a causalidade e representatividade das relações humanas, mas revelar a complexidade e diversidade intrínseca da sua natureza, orientando-se no pressuposto de que diferentes pessoas experienciam (e atuam sobre) a mesma realidade objetiva e material, de maneiras 'claramente' diferentes (Della Porta e Keating, 2008: 28; Miles e Huberman, 1994: 29). Logo, esta abordagem mais natural e interpretativa procura localizar e situar o observador (neste caso, o moderador) no mundo para tornar visível a sua existência e (con)vivência nele. Assim, este trabalho apoia-se sobretudo nas entrevistas por ser o método eleito mais apropriado para estudar o seu objeto.

Uma das características principais da entrevista é o seu *ethos* comparativamente mais forte que os demais métodos porque permite uma maior proximidade e profundidade face aos tópicos discutidos com/pelo entrevistado, tanto no conteúdo substantivo das suas opiniões e interpretações, quanto na forma (verbal ou não) como as expressa utilizando a voz, entoação, ênfase, linguagem corporal, etc. Este nível de descrição mais detalhada destaca-se entre as demais técnicas de pesquisa na medida em que uma resposta, mais aberta e profunda, pode capturar e revelar a inter-relação de emoções e razões dos entrevistados que de outra forma possam estar (e permanecer) ocultas ou escondidas (Weiss, 1994: 122-3).

Além disso, a entrevista facilita uma oportunidade para o esclarecimento direto e imediato de eventuais dúvidas que surjam no decorrer da entrevista e a possibilidade de orientar e adaptar as perguntas a fim de obter uma descrição mais rica e apropriada ao objeto estudado. Neste sentido, foi realizada uma pesquisa intensiva por meio de entrevistas semiestruturadas em profundidade em que cada participante foi convidado, individualmente, a (1) explicar se acredita ou entende que o ónus da prova na moderação de conteúdos é necessário e está a criar comunidades mais sustentáveis online, (2) comentar se todos os usuários devem ter igual 'acesso' e 'tratamento' online no ónus da prova e se tem/deve haver alguma interferência de

outras entidades, (3) visualizar e decidir se cada um de seis conteúdos deve ser mantido ou extraído numa rede social hipotética, em detrimento do que constitui "evidência" e/ou "clareza" de uma violação para si, (4) refletir sobre o futuro do ónus da prova na revisão humana de conteúdos e o que a torna tão especial face à inteligência artificial.

No total, o estudo de observação e análise dos conteúdos teve uma duração de aproximadamente 1 hora e 30 minutos (incluindo eventuais pausas) e foi conduzido em pessoa ou à distância (via Skype ou Google Hangouts), nas línguas Portuguesa, Inglesa ou Espanhola. A transcrição das entrevistas foram posteriormente analisadas e interpretadas por meio da condensação de sentido/significado, num processo que se refinou e resultou no presente documento que descreve o entendimento geral de cada moderador a respeito destas matérias. É importante frisar que, independentemente do rigor e validade das respostas neste trabalho, a identidade dos participantes foi salvaguardada devido à insegurança, sigilo e sensibilidade que esta função acarreta, até porque todos defenderam que quem age nesta capacidade deve permanecer no anonimato.

Em jeito de nota, o conjunto preliminar de conteúdos fornecido neste documento não é exaustivo tão-pouco representativo das centenas de decisões em média que é esperado de cada moderador num dia regular de trabalho, não obstante, é uma amostra e qualquer conteúdo poderá ser considerado atroz, perturbador, obsceno e inadequado para alguns. À data desta publicação, todo o conteúdo facultado ou de outra forma obtido, analisado e retido neste trabalho está disponível publicamente on-line (i.e. optou-se por excluir alguns conteúdos não-moderados, ou seja, "sem filtro"), e é limitado ao mínimo necessário para atender às finalidades legítimas e justas para as quais foi adquirido (i.e. para um estudo científico no âmbito da moderação de conteúdos online, pelo que não declara conflito de interesses algum).

Para efeitos de simplificação, decidiu-se não divulgar quem poderia ter publicado e/ou reportado conteúdo X ou Y, mas ficou explícito que o moderador poderia incluir isso na sua reflexão ou justificação. Na secção que se segue de cada conteúdo, começamos com um breve parágrafo de introdução ao seu contexto de produção (ausente para os analistas, devido à visão-contentor dos conteúdos e ao julgamento incompleto do ónus da prova), e, logo, um breve apanhado geral do que é "evidente" e/ou "claro" para o conjunto de moderadores em cada conteúdo, concluindo com a decisão e razão para mantê-lo ou removê-lo de uma plataforma online hipotética. Por último, na segunda parte da entrevista, problematizamos as suas decisões à luz das outras questões supracitadas.

### 4.3 Hipóteses de Resposta

Antes de continuar, e para compreender de que forma o Ónus da Prova na moderação de conteúdos online está a ajudar a criar comunidades online mais sustentáveis, levanta-se um conjunto de outras perguntas a nível macro, meso e micro, as quais apresentaremos (muito) sucintamente para ensaiar hipóteses de resposta preliminares ao problema enunciado e raciocínio proposto para esta investigação. Há alguma opinião comum sobre o que é sustentabilidade? Hipótese: Talvez; as pessoas não são neutras ou imparciais e têm opiniões diferentes sobre o seu significado. Há alguma necessidade para afirmar estas opiniões? Sim; o desenvolvimento da vida em sociedade deve ser criada e sustentada por cada cidadão.

Há algum meio (online) para satisfazer esta necessidade? Sim; a própria Internet e as redes sociais online como meios de comunicação criados e/ou utilizados por cada cidadão. Há alguma condição para existir este meio? Sim; a habilidade no acesso à informação, à liberdade de expressão, ao pluralismo de opiniões, etc. Há alguma objeção a esta condição? Sim; a injustificada interferência do estado e a desobediência civil aquando do conflito motivado por forças/opiniões ideológicas. Há alguma maneira de evitar essa objeção? Sim; o ónus da prova na moderação de conteúdos com clareza suficiente de evidência num contexto apresentado para julgar de acordo com políticas internas e padrões comunitários da empresa.

Há alguma razão para duvidar dessa maneira? Sim; o autor não está presente no ato da tomada de decisão por isso o ónus da prova é incompleto, e, como muitas vezes a ‘evidência não é clara’ e a ‘clareza não é evidente’, torna-se difícil justificar a decisão mais certa. Há alguma solução para aplacar qualquer suspeição? Sim; o algoritmo poderia ser melhorado para tomar decisões mais rápidas, iguais, e precisas, para aplacar qualquer dúvida sobre a clareza do que foi expresso. Há algum problema ao executar essa solução? Sim; a definição algorítmica de um modelo conceptual de clareza, definiria não só a liberdade na ‘forma’ do que é expresso, como no ‘conteúdo’ do que pode ser expresso, logo a livre-expressão e mesmo o livre-arbítrio.

Há alguma habilidade em contornar esse problema? Sim; partilhar estes desafios e oportunidades no ónus da prova com o público, e recrutar e formar analistas com competências críticas e reflexivas em análise e moderação de conteúdos. Há algumas redes sociais online a suster essa habilidade? Não; o usuário já pode esclarecer (em algumas redes) como/porque escolheu ou (in)tentou expressar-se de tal forma, mas nem sempre é uma questão de forma ou evidência, e é necessário um debate mais esclarecido que partilhe questões do conteúdo ou clareza no que concerne aos seus usos no trabalho ideológico que ela faz (ou evita) tanto na criação e/ou publicação de conteúdos (pelo Autor) como na sua interpretação (pelo Moderador). Sem mais delongas, segue-se a análise e decisões na moderação dos conteúdos pelos analistas.

## V. ESTUDO PRÁTICO 5.1 Entrevista: Parte I

### 5.1.1 Suástica 卐 ou 卍

Na imagem da esquerda é representado um rapaz Hindu no contexto da tradição popular Upanayana, uma cerimónia religiosa que, além de amarrar fios, envolve também a tonsura (i.e. o cabelo da criança é parcialmente rapado) como parte de três ritos de passagem Samskara – deixando frequentemente apenas um tufo de cabelo (sikhã) e o desenho da suástica no topo da cabeça. Na imagem da direita é representado um vaso gigante em frente à Porta da Casa-do-Tesouro (Hōzōmon) de acesso ao Templo budista Sensō-ji Kannon em Tóquio (Japão), com carateres que se traduzem em "doação" (eixo horizontal no topo) e "frente do tesouro" (eixo vertical no centro) que intersetam com a suástica ao meio (que geralmente marca as fachadas e localização de templos/santuários na cartografia deste e outros países do extremo Oriente).

Neste conteúdo (ver Anexo C), para os doze moderadores são evidentes duas imagens, postas lado-a-lado, cada qual representando uma suástica mas com configurações, orientações e aplicações diferentes. Na imagem da esquerda, a suástica a vermelho está orientada para a direita (i.e. sentido relógio) – com hastes ligeiramente angulares nas pontas –, num ângulo picado, sobre o cume da cabeça (reclinada e quase totalmente rapada) de um jovem. Na imagem da direita, a suástica a dourado está orientada para a esquerda (i.e. contra relógio) – com linhas retilíneas parcialmente arredondadas nas pontas – num dos lados de um vaso gigante, maciço, tingido verde-jade ou -esmeralda escuro, com abertura em "boca" de lótus.

Por um lado, sete moderadores mantiveram o conteúdo (Anexo D) porque, embora três tenham associado o símbolo na imagem da esquerda à suástica Nazi, a maioria diz que é claro que o autor o representa num contexto religioso e que criminalizar toda a suástica como um símbolo inconstitucional de ódio apenas devido ao estigma social (sobretudo nestes conteúdos mais ambíguos) "seria negar milhares de anos da sua iconografia sagrada" como símbolo de divindade e espiritualidade nas culturas da Eurásia – desde os Índios Hopi aos Astecas, dos Maias aos Celtas, dos Gregos aos Romanos, dos Hindus aos Jainistas, aos Budistas, etc.

Por outro lado, cinco moderadores removeram o conteúdo (Anexo D) porque não é claro que a imagem da esquerda represente senão a "suástica Nazi, inclusive o autor pode estar a compará-las/ diferenciá-las para celebrar ou promover" uma interpretação diametralmente oposta à de acima quando as coloca lado a lado; mas, mesmo que não seja a suástica hitleriana, é inconciliável o seu significado antigo (como signo de prosperidade e bom auspício até à década de 1930 no mundo ocidental) com o seu significado atual (sobretudo desde o rescaldo da II Guerra Mundial, a partir da qual passou a caracterizar um logotipo propagandista da identidade "raça pura ariana" e, mais recentemente, da supremacia branca no geral).

### 5.1.2 Sátira e Pecado

Em Maio de 2015, a Constituição da Irlanda (uma nação tradicionalmente conservadora e católica) foi emendada com um referendo que aprovou o casamento entre casais do mesmo sexo; tornando-se o primeiro país do mundo a legalizar o casamento homossexual por voto popular. Numa conferência em Roma, o Cardeal Pietro Parolim (Secretário de Estado do Vaticano) descreveu o SIM da Irlanda como "uma derrota para a humanidade" e, como tal, o cartoonista eslovaco Marian Kamensky desenha e publica esta sátira que, nesta dissertação, foi traduzida para o inglês para efeitos de simplificação.

Neste conteúdo (ver Anexo C), para os doze moderadores é evidente um bispo/cardeal da Igreja Católica Romana (talvez pela batina negra e solidéu rosado ou pela cruz latina *imissa* na parede do quarto) que, ao reparar do lado de fora da janela dois homens sorridentes a passear de mãos dadas pela rua em fatos cor-de-rosa, aponta boquiaberto e exclama fumegante e enrubescido "Estes Irlandeses são uma derrota amarga para a humanidade!"; enquanto, debaixo do hábito e entre as suas pernas, parece estar uma outra pessoa, de porte menor, de joelhos, somente com os pés a descoberto. Todos os analistas veem um comentário racista aos irlandeses ("mas claramente homofóbico") e assumem uma comparação do bispo com predadores sexuais.

Por um lado, seis moderadores decidiram manter o conteúdo (ver Anexo D) porque apesar de que possa ser uma crítica de "mau gosto" do autor contra a hipocrisia e promiscuidade moral de alguns membros da Igreja Católica, "não foi criada necessariamente para ser credível, séria ou representativa das ideias que o satirista está a atacar". De facto, se um preconceito romântico exerceu considerável influência até hoje de que a sátira (bem como o humor e tudo o que provoca risos) é indigna de estudo/atenção sérios e credíveis, porém, quando o são, tendem a ser censuradas pelas razões contrárias (i.e. demasiado sérios ou credíveis). "É um pau de dois bicos", diz um moderador, e isso leva-nos às observações do grupo seguinte.

Por outro lado, seis moderadores decidiram remover o conteúdo (ver Anexo D) porque é claro (embora não evidente) que o autor está simultaneamente a sexualizar menores e a injuriar, caluniar e/ou difamar – de forma muito subtil mas acutilante, contundente e cáustica (talvez até blasfema) – "a honra e bom nome de membros de uma congregação religiosa" que são um grupo protegido com características comuns, "talvez até vituperando a dignidade e integridade de um bispo/cardeal qualquer" (como uma figura lasciva e promíscua), e, ainda que o faça indireta ou implicitamente para escapar à remoção (de uma maneira que as críticas mais diretas não fazem), não se torna, porém, mais aceitável, sobretudo quando "é impossível acautelar se um satirista realmente aprova (ou pelo menos aceita como natural) as mesmas coisas que está a atacar".

### 5.1.3 SnapChatice

Em 2017, um jovem – cuja identidade foi salvaguardada devido à idade do adolescente em questão – publicou no SnapChat, sem consentimento, uma série de fotografias de (e com comentários sobre) um outro passageiro Sikh que viajava no mesmo avião (supostamente rumo a Indiana, EUA), que geraram polémica. Neste conteúdo (ver Anexo C), para os doze moderadores são evidentes vários comentários sobrepostos a cada uma de quatro imagens que compõe uma montagem, cuja maioria representa um senhor de barba longa grisalha e envolto num traje rematado com um turbante cor-palha ou -marfim no topo.

Na primeira imagem vemos desenhado um círculo vermelho que assinala o senhor de costas "para nós" mas defronte para o compartimento aberto de malas de cabine, seguido de um comentário em que se lê: "Não se preocupem sou capaz de não chegar a Indy". Na segunda imagem, o mesmo senhor encontra-se sentado de olhos fechados num dos lugares que ocupa a fileira detrás do rapaz, que começa a emergir pela primeira vez e publica a imagem com a seguinte observação: "Atualização ainda estou vivo 'XD" (pictograma a rir de boca aberta com gota de suor frio, que tende a expressar alívio após superar um desafio ou situação difícil).

Na terceira fotografia o enquadramento é semelhante mas menos aproximado, e agora o senhor está com a cabeça ligeiramente inclinada e o jovem aparece "a espreitar" e comenta: "Por favor deus deixa só o homem dormir". Na quarta e última imagem vemos um enquadramento fechado e cortado do nariz para cima do jovem, que reporta: "Ok ele acabou de andar para trás do Avião depois para a frente e depois para o seu lugar :O" (emoji de boca aberta, cuja expressão pode significar desde a mais pequena admiração até à total estupefação, em como a pessoa está positiva ou negativamente surpresa, impressionada ou atordoada).

Para a maioria dos doze moderadores é claro (embora não evidente) que o jovem parece estar, como coloca um deles "a confundir um senhor Sikh, primeiro com um muçulmano e, logo, com um terrorista". No entanto, por um lado, três moderadores decidiram manter o conteúdo (ver Anexo D) porque "podíamos remover afirmando que o autor foi insensível ao assumir uma ameaça, mas não podemos negar que o estamos a assumir também e sem evidência para justificá-lo; será que todo o público o assume tal-qual?". De facto, dos doze moderadores, somente um identificou que o senhor era provavelmente Sikh, os restantes abstiveram-se ou confundiram com um senhor muçulmano também.

Por outro lado, nove moderadores decidiram remover o conteúdo (ver Anexo D) porque, como diz um moderador mais concretamente, "o jovem parece insinuar que o passageiro de turbante denota comportamentos suspeitos e constitui uma ameaça" (ler comentários na quarta, terceira e segunda imagem respetivamente) e isso, embora não seja evidente, pode ser

justificado pelo preconceito social que se gerou com as comunidades islâmicas "sobretudo após os ataques terroristas no 11 de Setembro". Caso contrário, se nada mais, "no limite, está a invadir a sua privacidade ao publicar fotografias sem o seu expresso consentimento".

### 5.1.4 Pensa Positivo!

Freddie Mercury foi um cantor, pianista e compositor britânico que ficou mundialmente conhecido como fundador e vocalista da banda de rock Queen, que ele integrou de 1970 até falecer vítima de uma broncopneumonia em 1991 acarretada pela SIDA que lhe tinha sido diagnosticada anos antes (i.e. Freddie era seropositivo). Cerca de vinte e quatro horas após Freddie fazer o comunicado a confirmar que era portador do vírus, Freddie morre. A imagem neste conteúdo é a fotografia da sua capa de álbum (em 1992, atribuído a Mercury Songs) que parece ter sido adaptada para esta espécie de MEME.

Neste conteúdo (ver Anexo C), para os doze moderadores é evidente uma rúbrica em que podemos ler "Estás a pensar sair do armário? Pensa Positivo!" seguida da imagem de um homem (que alguns reconheceram como Freddie Mercury). Para a maioria dos moderadores, à primeira vista não há (nem parece haver) violação alguma a apontar neste conteúdo, aliás pelo contrário porque o autor escreve literalmente que se alguém está a "sair do armário" (i.e. a assumir/ revelar/ publicar a sua sexualidade) que pensem "positivo!". No entanto nem todos os moderadores acharam o conteúdo inteiramente inocente.

Por um lado, onze moderadores decidiram manter o conteúdo (ver Anexo D) mas por razões diferentes: para dois o conteúdo é clarividente (uma vez que um não conhece a expressão "sair do armário" e outro não reconhece o cantor); para outros nove, a evidência é suficientemente clara e, embora possa ser efetivamente uma piada com a SIDA ou outra coisa qualquer, essa clareza não é suficientemente evidente no conteúdo para justificar o contrário, pelo que uma remoção deste tipo de conteúdos seria "assumir demais e tolerar de menos" ou "aplicar sanções demasiado" porque "não reflete o espírito da política".

Por outro lado, uma analista decidiu remover o conteúdo (ver Anexo D) porque, embora a evidência seja suficientemente clara, pode ser claro também que o autor está a "gozar com a morte do cantor e/ou com a SIDA e as suas vítimas, e/ou a desencorajar aqueles que de facto estão a pensar sair do armário de realmente o fazerem", e, se sim, coloca-se ainda a questão se se pode/deve brincar ou não com qualquer uma dessas coisas; o moderador reclama então o ónus da clareza alegando que de outra forma pode ser recebido com uma crítica "zero positiva" por aqueles que o poderão interpretar como discurso cruel e insensível com a "orientação sexual" e "doenças sérias" (que são características protegidas).

### 5.1.5 Quem ganha?

Nos países lusófonos designa-se por Pedra-Papel-Tesoura um jogo em que cada um de três gestos com a mão representa um símbolo (geralmente definido como: punho fechado = pedra; mão aberta = papel; dois dedos esticados = tesoura) e, caso os jogadores façam o mesmo gesto, dá-se um empate. Se este jogo é relativamente simples, a análise do conteúdo que se segue não é, porém, para amadores, sobretudo quando o conjunto das decisões dos moderadores chegam quase a um empate também e mormente gesticularam para tentar explicar o que é claro (mas não evidente) aqui.

Neste conteúdo (ver Anexo C), para os doze moderadores é evidente uma rúbrica em que podemos ler "a jogar/jogando pedra, papel ou tesoura com o esquadrão" seguido de uma imagem com o que aparentam ser malformações e/ ou deformações fisiológicas ou anatómicas cujas causas (desconhecidas neste conteúdo) poderíamos especular que vão desde anomalias congênitas (como sin-/ oligo-/ ectro-/ clino- dactilia, etc.) até amputações cirúrgicas; entre outras. À primeira vista não parece haver violação alguma a apontar neste conteúdo, aliás pelo contrário porque o autor escreve literalmente que está "a jogar" pedra-papel-tesoura, porém nem todos os moderadores acharam o conteúdo inofensivo.

Por um lado, seis moderadores decidiram manter o conteúdo (ver Anexo D) porque não é evidente (embora possa ser claro) que o autor está a fazer uma piada com (esses) portadores de deficiência, dizendo que "pela aplicação do presente contínuo/ progressivo na rúbrica subentende-se que possam ser os próprios que tenham criado e posteriormente publicado o conteúdo" (i.e. um tempo verbal do Presente usado em Inglês para descrever uma ação ou ocorrência no agora e/ou momento enunciado). Além disso, diz um moderador, "quais as regras do jogo? Afinal de contas, não diz em lado algum que são precisos uma mão e cinco dedos para jogar e ganhar e, se houvesse, isso sim seria claramente discriminação".

Por outro lado, seis moderadores decidiram remover o conteúdo (ver Anexo D) porque é claro (embora não seja evidente) que o autor faz uma piada com (esses) portadores de deficiência (que são um "grupo protegido"), e mesmo que sejam os próprios envolvidos na criação e/ou posterior publicação do conteúdo "não escusa estarem a gozar com características comuns que partilham com outros portadores de deficiência que podem ser sensíveis a isso". Além disso, diz um moderador, seria um argumento do "absurdo contra o senso comum manter somente porque o significado dos gestos pode variar em alguns países ou entre comunidades de jogadores".

### 5.1.6 Análise crónica...

Esta foto da agência noticiosa Caters (que parece ter sido adaptada por um utilizador para este MEME), representa Curtis McDaniel de New Jersey EUA que tinha apenas 11 anos quando começaram a aparecer manchas brancas na pele causadas pelo vitiligo e, logo, começaram a aumentar e espalhar-se sobretudo de modo generalizado e simétrico pelo rosto e corpo todo, tornando-se vítima de bullying e de uma depressão profunda durante 5 anos, até finalmente superar ambos e abraçar a sua diferença (inclusive tornando-se modelo). Vitiligo é uma condição crónica caracterizada por áreas de despigmentação cutânea cujas áreas afetadas tornam-se "manchas" brancas com margens geralmente bem delimitadas.

Neste conteúdo (ver Anexo C), para os doze moderadores é evidente uma rúbrica em que podemos ler "Quando ficas uma semana sem roubar" seguida da imagem de um homem com vitiligo espalhado pelo rosto e corpo – de olhos e boca semicerrados, cabelo rapado, e bigode desenhado –, a usar uma t-shirt branca justa e arrepanhada pela postura com que está sentado e encostado a uma boca de incêndio metalizada cinza. Embora vitiligo seja mais recorrente em pessoas de pele escura, é de antecipar desde já que nenhum moderador achou o conteúdo um ataque (direto) à pessoa representada e tão-pouco à sua condição, mas um ataque geral à raça negra.

Por um lado, três moderadores decidiram manter o conteúdo (ver Anexo D) porque a evidência não é suficientemente clara e, embora possa ser claro que o autor está a comparar assaltantes e pessoas de raça negra em jeito de piada "posso estar a dizer isso porque sou branco", e "não posso assumir que é uma pessoa de raça negra somente pela quantidade de manchas, além disso quem anda a contar ou a medir?" e outro moderador diz "é uma questão de tempo, afinal de contas a pessoa em breve poderá ficar completamente branca e aí teríamos de questionar se é de facto uma questão de raça ou cor"; por isso todos comentaram que se removermos o conteúdo podemos estar a acentuar mais ainda o racismo.

Por outro lado, nove moderadores decidiram remover o conteúdo (ver Anexo D) porque a evidência é suficientemente clara de uma pessoa com vitiligo mas "é óbvio" que o autor está a fazer uma comparação entre comportamentos criminosos e pessoas de raça negra; porém não souberam explicar exatamente o porquê: um diz "é uma pessoa negra pela fisionomia e menos quantidade de manchas brancas", outro diz "é uma pessoa de raça negra de ressaca ou em desmame porque já não rouba à uma semana e por isso está a ficar branca", outra diz "é contra uma pessoa de raça negra porque racismo contra brancos não existe!"; por isso comentaram de um modo geral que se não fizermos nada estamos a ignorar que esse estigma social sequer existe.

## 5.2 Entrevista: Parte II

Alguns moderadores defendem que estamos efetivamente a criar comunidades online mais sustentáveis "comparativamente ao que a Internet era antes", mas a maioria argumenta o contrário, por exemplo que "estamos o tempo todo a tentar apagar um incêndio sem tempo para preveni-lo", e que portanto "é só uma questão de estabilidade e risco" e "nem sempre conseguimos traçar um limite claro entre moderação e censura" porque é somente o que "as empresas creem ser importante fazer e não pensar, aliás não nos dão tempo nem nos pagam para isso". Não obstante, o total dos inquiridos concordou que o princípio base para criar comunidades online mais sustentáveis é a proteção na igual oportunidade de acesso à liberdade de expressão, como pedra basilar e angular do Estado de direito democrático nas sociedades modernas. Porém, não houve consenso sobre o tipo de intervenção estatal legítima para o efeito, aquando da sua colisão com outros direitos.

De grosso modo, e embora não os tenham conseguido qualificar como tal, os analistas concordaram que "a Internet e as redes sociais online são meios de comunicação" (cuja maioria são empresas privadas também), portanto "idealmente seriam independentes" ou "arriscamos formas de restrições" prévias do Estado "no acesso à informação do/pelo público geral" em detrimento de determinadas culturas, línguas, símbolos, crenças, valores, leis, etc. No entanto, disseram que "é inevitável não haver interferência alguma" porque essas empresas desempenham (direta ou indiretamente) um "papel cívico preponderante no tratamento" de conteúdos-gerados-por-utilizadores (em massa) "de sociedades em diferentes fases de desenvolvimento", em matérias como privacidade, vida, ordem, saúde (e pudor) públicos; reparemos no caso do Discurso de Ódio, Cruel ou Insensível.

Embora a 'igualdade' e a 'equidade' sejam importantes no panorama geral das coisas, os analistas lamentam ser difícil (talvez impossível sequer) proteger ambas simultaneamente numa única decisão sem comprometer uma ou outra, quando se trata do tratamento "igual" ou "diferente" da diversidade e pluralismo de expressões que advém (do exercício ou abuso) dessa liberdade. Por um lado, alguns analistas comentaram que, ao abrirmos um precedente para proteger um (sub)grupo, não há limite para a multiplicação de tantos outros como cor, casta, escolaridade, classe económica, beleza, inteligência, etc.; e provavelmente teríamos de estender isso a subgrupos como espetros de cor, estratos de castas, níveis de inteligência, graus de educação, escalões salariais, padrões de beleza, etc. "à semelhança do que atualmente fazemos com a extensão do acrónimo 'LGBTQIA+' ou com a definição de 'nível grave ou sério de doença ou deficiência"; "por isso temos de tratar todos iguais" ou discriminamos sempre quem não é abrangido nessa proteção.

Por outro lado, uns analistas mencionaram como os debates sobre exclusão e segregação surgiram exatamente de "questões históricas, estruturais e sistémicas no ataque a minorias, não a maiorias", por isso devemos colmatar a lacuna, "preencher o vazio deixado pelo potencial da justiça" para compensar os infortúnios daqueles em maior desvantagem e colocados à margem da discussão, "mesmo que isso implique uma distribuição desigual de crescente proteção no tratamento privilegiado a membros de grupos protegidos ao ponto de extravasar novamente no futuro" (ex.: ação afirmativa), porque, diz uma analista, "mais vale discriminação positiva do que negativa" e "para haver igualdade teríamos de optar em poder 'expressar tudo ou expressar nada', mas se escolhemos exprimir umas coisas e omitir outras, temos de tratar de maneira distinta os desiguais".

O curioso nestas duas abordagens é que, exceto uma moderadora, nenhum outro interveniente foi coerente em sustentar os seus argumentos com as decisões sobre os conteúdos, caso contrário poderíamos dizer que qualquer analista teria removido todos ou nenhum (i.e. igualdade), ou só aquele conteúdo com essas características (i.e. equidade); em rigor, nenhum dos doze tomou sequer o mesmo conjunto de decisões. Mas isso faz sentido, porque, embora a maioria deles tenha deduzido possíveis violações na maioria dos conteúdos, a clareza em dados contextos nem sempre foi suficiente para estabelecer uma lógica para todos. Isso significa que o tratamento da liberdade de expressão, como vimos, não é somente uma questão sobre o tipo de evidência que pode ser tratado igual ou diferente para todos, mas sobre o maior ou menor grau de clareza com que essa evidência (ou a sua ausência) pode ser interpretada por alguns.

Uma moderadora removeu todos os conteúdos incluindo aquele que assumiu ser uma piada contra a homossexualidade e/ou SIDA e/ou as suas vítimas, e outro moderador manteve tudo exceto o que aparenta ser (para si) a representação/ promoção da suástica Nazi. Um moderador removeu o que julgou ser uma crítica dolosa a membros da Igreja Católica como predadores sexuais de menores, mas manteve aquele MEME por não ser evidente uma comparação de pessoas de raça negra com assaltantes. Outro moderador removeu o que considerou ser uma piada de "mau gosto" com portadores de deficiência, mas manteve aquele SNAP por não ser evidente uma equivalência entre muçulmanos (ou sikhs) e terroristas. E por aí em diante. Quer dizer que algo muito estranho, mas de suma importância, parece emergir no reverso da discussão sobre a liberdade de expressão quando discutimos o que (define) verdadeira ou corretamente 'arte', 'tradição', 'ciência', 'protesto', 'humor', 'crítica' e outras temas em disputa, como 'notícias' (ex: no debate atual sobre a moderação de *fake news*):

O motivo, propósito, sentido, vontade, intenção com que determinados conteúdos poderão ter sido criados e/ou publicados de antemão, num contexto. Porém, os moderadores atestam que, fora os casos mais clarividentes, "raramente temos contexto suficiente" para construir um caso de defesa ou oposição sólidos com vista a sustentar uma declaração de inocência ou culpa, sem incorrer em algumas polarizações ou generalizações toleráveis "em função do que é mais ou menos claro" do que o autor pode ter escolhido ou (in)tentado expressar (sobretudo devido à Visão-Contentor que discutimos antes). Nesse sentido, clareza seria "o que algo pode ser" (i.e. intenção), evidência é "o que algo é" (i.e. expressão), e no contexto entre eles está "o que algo é sobre" (i.e. relação). Assim, a questão tantalizante que se coloca é a liberdade com que algo pode ter sido expressado pelo Autor.

Essa ideia estende-se rapidamente a todas as outras categorias (algumas inclusive foram testadas neste estudo também, embora não incluídas nesta dissertação). Por exemplo, o quadro L' Origine du Monde (1866), do pintor realista e ativista oitocentista francês Gustave Courbet, mostra o fragmento anatómico – que compõe os quadris, o ventre protuberante, e a fenda cavernosa da vulva abastada com pelos pubianos – de uma mulher deitada, mas emoldurada de modo que nada mais visto abaixo das coxas afastadas ou acima dos seios tapados (parcialmente) por um lençol branco. Para todos os analistas isso é evidente, mas para alguns não basta o arsenal de boas intenções ao insistir que tudo aquilo que se representa como tal "é" arte, quando não nos convida a evocar valores do seu significado ou a divertir a ficção de um mundo que "pode ser" fora da pintura, senão aquela representação concreta e exata das aparências na realidade nua, crua e desarmante rapidamente identificada na superfície da tela, que, comentam uns analistas, ataca o "pudor público".

Hobbes e Locke apontaram ser um erro categórico atribuir liberdade à intenção porque esta não é 'livre' nem 'não-livre', mas todos os moderadores parecem atribuir-lhe algum grau de liberdade (se não à intenção diretamente, então à interpretação de como ela é expressa consoante o grau de clareza). De facto, há ataques explícitos e implícitos, ativos e passivos, evidentes e latentes, e um ataque aparente ou iminente (mas claro) pode ser tão frutífero quanto um ataque evidente, mesmo alguns mais descuidados. Contudo, se assumirmos que a intenção é "determinada" (i.e. não-livre), então a expressão reflete a intenção e responsabilizamos o autor de tudo. Porém, se assumirmos que a intenção é "livre", então a expressão não reflete necessariamente coisa alguma e responsabilizamos o autor de coisa nenhuma. É algures aqui que reside tanto uma dificuldade quanto uma oportunidade em explicar porque designamos então esta prática de moderação de "Conteúdos" quando tudo o que tendemos a discutir é ao nível da (liberdade de) "Expressão".

Em parte, isso pode ser explicado talvez com a diferença que os moderadores fazem entre a revisão humana e a inteligência artificial. Quando confrontados com a pergunta "o que significa ou torna ser Ser humano tão especial na prática de moderação de Conteúdos?", as respostas cobriram conceitos tão abstratos como "contexto", "cultura", "identidade", "condição", "sentimento", "compaixão", "empatia", "sentido", "consciência", "ponderação", "subjetividade", "interpretação", "abstração", mas, sobretudo (e talvez não surpreendentemente), "clareza", "significado" e "intenção". Logo, parece que ao falarmos de "clareza" estamos essencialmente a referir-nos ao plano do conteúdo, da intenção, dos significados; enquanto ao falarmos de "evidência" estamos realmente a discutir o plano da forma, da expressão, dos significantes; e ao falarmos do "contexto" estamos possivelmente a sondar o plano do uso, da relação, e do que se torna significativo.

É em meio da interseção destes mundos que os moderadores tomam/devem tomar as suas decisões (tríade de uma análise e moderação mais sustentável, Anexo E); e talvez não seja por acaso que todos eles tenham respondido, em seguida, que a inteligência artificial pode ser efetivamente muito útil em moderar, cada vez mais e melhor, a "forma" (não os conteúdos), a "evidência" (não a clareza), a "expressão" (não a intenção), os "significantes" (não os significados). Assim, se não conseguimos precisar exatamente o Conteúdo, no limite, o ser Humano deve carregar o ónus da Clareza e realizá-la sempre respeitando o maior ou menor grau de liberdade com que um autor pode ter intentado expressar algo. De alguma maneira, é aceitar que, devido à Visão-Contentor, algumas coisas são ainda mais difíceis de enxergar com clareza de contexto suficiente para decidir justamente, mas que existe responsabilidade tanto na articulação (pelo Autor) como na realização (pelo Moderador) na melhor imputação possível do significado com que cada conteúdo poderá ser interpretado pelos demais.

Lamentavelmente, a tese funcionalista ambiciosa que vem sendo inoculada por algumas teorias das ciências da comunicação e da informação (e da própria psicologia) afeta ao modelo mecânico do cérebro e suas analogias fascinantes com um computador, anima comparações erróneas e "portáteis" (apesar das semelhanças curiosíssimas) entre os sofisticados agrupamentos de neurónios do cérebro com o 'hardware' e o serpentear de um emaranhado de fios que compõe o circuito elétrico do 'software' com a mente. Todavia, tendem a ignorar um (todo) sistema vivo, dinâmico e infinitamente mais complexo com que a nossa mente está intima e inextricavelmente relacionada e posicionada em todo o lado e em lado nenhum, num diálogo criativo mas discreto com o mundo, numa inteireza indivisível, irredutível, intangível, irreproduzível, e num abismo intransponível a observações ou medições físicas e concretas que a definam como 8 polegadas ou 1.1 quilogramas.

O potencial para esclarecimento estaria algures aí, colocado delicadamente numa "linha" divisória muito crítica, num equilíbrio muito ténue e subtil, entre o estático e o caótico, o tédio e a confusão, o silêncio e o ruído, o bom senso e o absurdo, o comum e o extraordinário, o esperado e o acidente. Percorrer essa linha é um ato autêntico de funambulismo e, malgrado nosso medo vertiginoso de perder o equilíbrio (i.e. a falência das nossas certezas), é importante reter que esse equilíbrio é um constante contrabalanço natural que nos obriga a reconhecer possíveis limites à nossa liberdade para descobrirmos um maior controlo de formas criativas e alternativas de crescermos com ela. Se removermos o valor dos significados da equação (i.e. a clareza do ónus), o quociente da inteligência humana pouco difere daquela artificial.

Devido às falsas equivalências que negam este potencial, aliadas ao impacto que as decisões dos moderadores têm tido nas liberdades dos utilizadores ou aos efeitos psicológicos que os conteúdos têm tido em alguns moderadores, alguma parte da opinião pública pende no desejo de substituir a revisão (da consciência) humana por completo com inteligência artificial. Vários moderadores lamentam esse facto, "até certo ponto já trabalhamos como máquinas", "tomamos decisões cada vez mais rápidas e com menos tempo para as questionarmos", "somos cúmplices desta transição", "somos um meio para um fim, ou um fim para um meio", "se calhar até já estamos a ser moderados, para esse efeito". Porém, se há um desejo e tendência em tornar um discurso tão isento de ambiguidades quanto possível, devemos, por outro lado, questionar o que seria da moderação de conteúdos se um dia não tivermos outra possibilidade senão ceder toda a riqueza discursiva à transparência de um todo-poderoso algoritmo, código e métrica.

Nesse dia, talvez reconsideremos o que é mais insidioso, o analista ser tendencioso ou o algoritmo ser neutro. "Uma tragédia talvez", mas apropriando as palavras de Aristóteles, a tragédia suscita sentimentos de horror e piedade, mas não de culpa. Até certo ponto, parece difícil imaginar (responsabilizar) um algoritmo na moderação de um excerto de comédia *stand-up* protagonizado por Sarah Silverman, ou um trecho do protesto de Thích Quảng Đức através da sua autoimolação, ou o vlog no encontro inesperado de Paul Logan com um cadáver em Aokigahara, [entre tantos outros conteúdos controversos que foram objeto deste estudo (mas não incluídos aqui)], sem incorrer em erros de proporções hercúleas do que define 'arte', 'protesto', 'humor', etc. É precisamente desarranjando as fronteiras entre o estético e inestético, ético e antiético, que o conteúdo instala um desconforto situacional que questiona a posição segura (e aceite) do moderador quando confrontado com o problema moral do seu voyeurismo ao tornar-se cúmplice (logo, responsável) do que está a acontecer, e esse ónus parece necessário estender ao público, ou todos os moderadores não teriam dito que a sua opinião a respeito destas questões foi negociada desde o momento em que começaram a moderar conteúdos.

Assim, independentemente da entidade e tipo de intervenção, todos estes analistas reiteraram ser "importante moderar conteúdos" e ser "humanos a fazê-lo", e que, por todas estas razões, devemos explicar "porquê que o fazemos" mesmo que não expliquemos exatamente "como o fazemos" na prática. Isto porque o desenho e constantes calibrações e clarificações exatas do conteúdo das políticas estão sempre em atualização e não refletem "exatamente" os conteúdos do público, e, se qualquer interpretação já gera discórdia entre os próprios moderadores, é incalculável a proporção de mal-entendidos que uma transparência total na visibilidade das políticas internas pode criar nos utilizadores, que, logo, "procurariam formas de contorná-las para enganar a máquina" ou "minaria a intenção deles sequer criarem ou publicarem conteúdo algum".

De facto, à data desta publicação, em muitas redes sociais online as políticas internas já estão parcialmente visíveis ao público e inclusive já é possível (embora difícil) recorrer/ apelar na eventualidade de alguns conteúdos serem removidos para clarificar o intuito na sua criação e/ou publicação. Contudo, diz um moderador "não investiga o ónus da clareza que é filtrado dos utilizadores, pelos moderadores, para os formuladores e empresas" e, se o faz, "não é partilhado connosco". Porém, como vimos, se não devemos (e provavelmente não podemos nem conseguimos) apontar qualidades formais particulares que nos permitam um entendimento estável e concreto do que constitui o fenómeno da clareza, devemos, no entanto, questionar o trabalho ideológico que ela faz ou evita nos conteúdos, quando o que está em jogo é a liberdade de expressão e o livre-arbítrio, ou censura de ambos.

Em suma, o ónus da prova (sobretudo da clareza) deve ser partilhado (não inteiramente retido pelas empresas, ou transferido por completo para o público), porque como diz uma moderadora "são precisos dois para dançar o tango". É preciso um 'porteiro' de sentinela e um 'olheiro' de visita ou as 'portas' não fazem sentido, quando se trata de salvaguardar os direitos de todos e de uma ponderação/apreciação valorativa mais sensível da matéria-do-sujeito (um anagrama gateKeePing e gatePeeKing, está em curso). Vários moderadores expressaram inclusive que se deve/ tem de haver alguma interferência na supervisão, então melhor que fosse de uma "comissão internacional, independente, especialista, mas isenta ao processo", que não tão-somente "aceita e gere recomendações/sugestões de todas as partes", como "investiga também essas questões mais profundas em estudos análogos a este", "executa auditorias", "estabelece requisitos de entrada para moderadores (ex.: psicotécnicos, formação nas áreas das ciências da comunicação e da informação sobretudo naquelas conexas à análise de conteúdos, semiótica, do discurso, etc.)" e a "cria um tanque/ repositório de conteúdos 'com' e 'sem' filtro para problematizar algumas destas questões".

## CONCLUSÃO

### Moderação, Sustentabilidade e Kairós: Signos do tempo

Esta dissertação procurou compreender de que forma o Ónus da Prova na moderação de conteúdos está a ajudar a criar comunidades online mais sustentáveis. Assim, introduzimos este trabalho com uma breve (r)evolução histórica e cultural da "grande" ideia e mito de progresso, para situar o leitor em alguns debates sobre o conhecimento, desde a obscuridade da "Idade das Trevas", ao esclarecimento da "Idade das Luzes", até uma certa busca por clareza formal na atualidade que abriu a "porta" para a abordarmos o gatekeeping online, e o contexto histórico e atual, da análise e moderação, da forma e conteúdo.

Esse enquadramento permitiu uma incursão nos campos da investigação e especulação científica fundamentada sobre o fenómeno designado aqui de "ClarEvidência" para problematizar *quanta* evidência com clareza suficiente de contexto, se manifesta com carga analisável de prova na prática de moderação. Isso ajudou a diferenciar o ónus da prova online dos típicos procedimentos tradicionais offline, na medida em que o campo de análise do Moderador é limitado à clareza da evidência num contexto apresentado e não consegue construir um caso de defesa ou oposição sólidos para sustentar uma declaração de inocência ou culpa porque o Autor está ausente no ato da tomada de decisão (Visão-Contentor).

Assim, na secção seguinte, foram abordadas algumas questões sobre a estética, ética e dialética da clareza e na falsa equivalência de "transparência" com "verdade" e "retidão", sugerindo que, para haver um maior e melhor esclarecimento, é indispensável procurar constantemente um equilíbrio saudável entre uma omissão ou exclusão natural necessária na própria comunicação e evitando sempre que a transparência de umas coisas paralise as possibilidades e oportunidades do público, de interpretar e contestar, classes ou categorias inteiras, de informações potencialmente úteis em prole do discurso mais claro.

Tudo isto foi importante para testar "de que forma o Ónus da Prova na moderação de conteúdos online está a ajudar a criar comunidades online mais sustentáveis", com um grupo de moderadores que, em parte, confirmou que é difícil falar verdadeiramente de uma sustentabilidade do discurso em comunidades online porque, entre outras coisas, o ónus é incompleto e torna-se difícil desenhar um limite claro entre "liberdade" e "censura" naqueles casos menos clarividentes, em que a "evidência nem sempre é clara" e a "clareza não é evidente" para antecipar como alguns autores escolheram/poderão ter escolhido ou (in)tentado expressar-se e, se isso é "verdade" e/ou "correto". Nenhum dos doze moderadores tomou o mesmo conjunto de decisões sobre os mesmos seis conteúdos, cada qual demonstrando diferentes níveis de sensibilidade, tolerância e coerência sobre a gravidade de alguns tópicos.

Até certo ponto, todos defenderam que a Internet e as redes sociais online são meios de comunicação que veiculam conteúdos-gerados-pelos-utilizadores, cuja liberdade (em escolher) expressar-se deve ser protegida no 'acesso' e no 'tratamento', mas não houve consenso sobre o tipo de tratamento ou interferência que são expectáveis aquando da colisão desta liberdade com os direitos de outros, naqueles casos menos clarividentes. Contudo, todos concordaram que o ónus da prova na moderação de conteúdos é necessário, conquanto nem sempre o estado da prova em si seja suficientemente evidente e/ou claro, mas que é exatamente por isso que é preciso que ser 'humanos' a moderar conteúdos criados e/ou publicados por humanos, sob pena de que um algoritmo avançado (apesar da sua real utilidade neste âmbito) mine a quantidade e/ou qualidade do que é expresso (forma, evidência, significante) ou do que pode ser expresso (conteúdo, clareza, significado).

Assim, concluímos com algumas das suas considerações e sugestões de que o ónus da clareza deve ser partilhado (não retido pelas empresas ou transferido para o público), sem nunca definir conceptual ou exatamente, mas questionando os seus usos no trabalho ideológico que ela faz (ou evita) tanto em como o Autor a articula (concreta) na criação e/ou publicação de conteúdos, como o Moderador a realiza (subjetiva) na análise e moderação destes. Até lá, parece realmente difícil responder se estamos de facto a criar gerações online mais livres atualmente, sem comprometer as liberdade de gerações vindouras em escolher e expressar-se desse jeito também. Um trabalho mais aprofundado sobre a ideologia na moderação, está em curso.

Em jeito de conclusão, o significado de um desenvolvimento sustentável não é fundamentalmente uma questão científica ou técnica inequívoca ou neutra, mas uma questão intrinsecamente complexa e inerentemente normativa de um processo social inevitavelmente dinâmico e experimental do comportamento humano. Este, por sua vez, é condicionado pela perceção, opinião e posição única de cada cidadão e sobre a maneira apropriada de conceber a sua relação com o meio envolvente num contexto específico-concreto, sob condições de contingência e incerteza profundas em questões de valor e juízos morais sobre os futuros preferidos (Elliott, 2000: 9).

De facto, sustentabilidade quis dizer coisas diferentes para diferentes pessoas em tempos diferentes. Houve um período em que beber álcool fazia-te mais sexy, fumar fazia-te mais saudável, terapia de conversão fazia-te mais normal, pesticidas fazia-te mais preocupado com o meio ambiente, etc. (Anexo F). O que quer dizer que o próprio conceito de sustentabilidade baseia-se na necessidade de se pensar em escalas temporais e naquilo que é mais oportuno em determinado momento mediante as condições de necessidade e suficiência do público. Nesse aspeto, o conceito de clareza (num contexto) é análogo ao conceito de *Kairós* aplicado no

âmbito da mitologia e teologia, que – ao contrário do tempo sequencial, linear e quantitativo de *Khrónos* –, significa a experiência do "tempo certo" ou do "momento oportuno" (embora indeterminado) em que algo especial acontece. Até certo ponto, podemos dizer que aquelas decisões tomadas num momento mais provavelmente adequado à recetividade do público têm, geralmente, um bom senso de Kairós. Mas, assim como com a "clareza", não podemos, de boa fé, dizer aos utilizadores ou aos moderadores precisamente qual é o tempo mais apropriado, certo ou ideal, pois essas qualidades são (e provavelmente sempre serão) dependentes de circunstâncias específicas que são potencialmente infinitas em número.

Em todo o caso, muito embora seja necessário o desenvolvimento e refinamento conceptual/ teórico e metodológico para alcançar um desenvolvimento mais sustentável em comunidades online (e mesmo offline), a incerteza e a dúvida só pode ser superada com devida análise, e novas formas de exploração e aprendizagem social, "pois o teste será o modo como as coisas acontecem nas ruas no dia-a-dia" (Robinson, 2004). E para isso é preciso uma boa noção de tempo. E, mais do que nunca, parece oportuna uma discussão sobre a partilha do ónus da prova e sobre algum grau de liberdade de intenção que é esperado e respeitado mediante esclarecimento dos conteúdos do/pelo público.

Ao mesmo tempo, a divisão disciplinar e hiperespecialização do conhecimento no sistema universitário significa que muitas questões transversais (ex: sobre o clareza, evidência, contexto, e mesmo sobre a liberdade de expressão ou livre-arbítrio) se perdem nos "espaços em branco" ou no "vazio" entre as disciplinas e, como tal, uma abordagem do conceito de sustentabilidade face à moderação de conteúdos pode ter um papel catalisador em ajudar a colmatar algumas dessas lacunas, de modo a repensar o novo paradigma para um desenvolvimento sustentável à luz das teorias multiparadigmáticas do ramo epistémico das ciências da comunicação também.

Esta dissertação procurou testar e compreender como a sustentabilidade do discurso em comunidades online poderá estar ligada a uma questão mais profunda sobre como carregamos o Ónus da Prova na moderação de conteúdos, e o que esta tensão significa para o futuro da liberdade de expressão e do livre-arbítrio (na medida que os experienciamos). Assim, o estudo prático explorou alguns significados que uns analistas atribuem à "expressão" e aos "conteúdos" quando confrontadas com decisões de moderação difíceis, em que a "evidência nem sempre é clara" e a "clareza nem sempre é evidente".

No entanto, à data, a questão da moderação de conteúdos é deixada num impasse, logo, irresoluto: Num mar tempestuoso de infinitas possibilidades e probabilidades de interpretação, viveremos eternamente assolados pela dúvida e fadados a enxergar apenas sombras ou imagens turvas e toldadas do real significado de cada conteúdo? O ónus da prova na moderação de conteúdos é irremediavelmente confuso e fatalmente comprometido ou oferece alguma esperança para navegar no sentido de uma maior entendimento? Haverá então alguma nitidez ou clareza fixa por que esperar no desenho das políticas públicas online para um desenvolvimento mais sustentável para todos?

Se nada mais, então este trabalho conclui que se calhar não há uma única resposta para estas e outras questões porque têm de ser constantemente atendidas, e a qualidade de qualquer abordagem começa pela qualidade da pergunta e da nossa capacidade de conceber respostas alternativas. "O que significa ou torna ser Ser Humano tão especial na moderação de conteúdos" pode ser um ponto de partida, e se calhar a resposta está na clareza desta pergunta. Independentemente disso, este trabalho é baseado (mas não determinado) na ClarEvidência, é informado (mas não resolvido) nos debates sobre a liberdade de expressão e do livre-arbítrio, é inspirado (mas não paralisado) no atual contexto histórico e sociocultural em que se insere, e sugere (mas não conclui) que o ónus da prova na moderação de conteúdos é de uma natureza profundamente moral, política e discursiva que exige deliberação ponderada e resolução coletiva, assim como moderadores emancipados e formados nestas áreas, e uma transformação mais fundamental dos valores e atitudes subjacentes dos públicos para criar/ motivar mudanças substanciais no comportamento ou na prática em sociedade.

## BIBLIOGRAFIA


Bakari, M. (2017). Mapping the 'Anthropocentric-Ecocentric' Dualism in the History of American Presidency: the Good, the Bad, and the Ambivalent. *Consilience: The Journal Of Sustainable Development*, *17*(1), 1-32. Retrieved from https://www.jstor.org/stable/26188780?seq=1&cid=pdf-reference#references_tab_contents

Bardin L. *Análise de conteúdo*. Lisboa: Edições 70; 1977.

Berelson, B. (1952). *Content analysis in communication research*. New York: Hafner.

Bury, J. (1920). *The Idea of Progress: an Inquiry into its Origin and Growth* (p. 2). London: Macmillan and Co.

Campos, C. (2004). Método de análise de conteúdo: ferramenta para a análise de dados qualitativos no campo da saúde.

Cassirer, E., Kristeller, P., & Randall, J. (1987). *The renaissance philosophy of man*. Chicago: University of Chicago Press.

Cavanagh S. (1997) *Content analysis: concepts, methods and applications*. Nurse Researcher 4, 5–16.

Citizen Lab. (2018). Planet Netsweeper: an investigation into the global proliferation of internet filtering systems manufactured by canadian company Netsweeper, Inc. (pp. 1-60). Toronto: University of Toronto.

Cole F.L. (1988) *Content analysis: process and application*. Clinical Nurse Specialist 2(1), 53–57.

Creswell, J. (2003). *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches* (2nd ed.). Thousand Oaks, California: Sage Publications.

Della Porta, D., & Keating, M. (2008). *Approaches and methodologies in the social sciences*. Cambridge, N.Y.: Cambridge University Press.

Desjardins, J. (2019). What Happens in an Internet Minute in 2019? [Blog]. Retrieved from https://www.visualcapitalist.com/what-happens-in-an-internet-minute-in-2019/

Elliott, J. (2000). *An introduction to sustainable development*. New York: Routledge.

Elo, S., & Kyngäs, H. (2008). The qualitative content analysis process. *Journal Of Advanced Nursing*, 62(1), 107-115. doi: 10.1111/j.1365-2648.2007.04569.x

European Commission. (2016). *Code of Conduct on countering illegal hate speech online: First results on implementation*. Retrieved from http://ec.europa.eu/information_society/newsroom/image/document/2016-50/factsheet-code-conduct-8_40573.pdf

Freedom House. (2018). The Rise of Digital Authoritarianism (p. 13). Washington DC. Retrieved from https://freedomhouse.org/report/freedom-net/freedom-net-2018

Heuer, R. (1999). *Psychology of intelligence analysis*. Washington D.C.: Center for the Study of Intelligence, Central Intelligence Agency.

Holsti, O.R. (1968). Content Analysis. In G.Lindzey & E.Aronson (Eds.), *The Handbook of Social Psychology* (2nd ed.) (Pp.596-692), Vol.II, New Delhi: Amerind Publishing Co.

Hsieh H.-F. & Shannon S. (2005) *Three approaches to qualitative content analysis*. Qualitative Health Research 15, 1277– 1288.

Janeira, A. (1972). A técnica de análise de conteúdo nas ciências sociais, natureza e aplicações. *Análise Social : Revista Do Instituto De Ciências Sociais Da Universidade De Lisboa*, (Vol. 9.), 370-399. Retrieved from http://analisesocial.ics.ul.pt/documentos/1224260109P6yXY4bm6Vt51JF8.pdf

Jensen, K. (2002). *A Handbook of Media and Communication Research*. USA: Taylor and Francis.

Kemp, S. (2019). Q2 Global Digital Statshot [Blog]. Retrieved from https://wearesocial.com/blog/2019/04/the-state-of-digital-in-april-2019-all-the-numbers-you-need-to-know

Kracauer, Siegfried. (1952-1953). *The challenge of quantitative content analysis*. Public Opinion Quarterly, 1 6, 631-642.

Kreuter, N. (2013). The Ethics of Clarity and/or Obscuration. Composition Forum, 27, 14. Retrieved from https://files.eric.ed.gov/fulltext/EJ1003970.pdf

Krippendorff, K. (2004). *Content analysis: An introduction to its methodology* (2nd ed.). Los Angeles, CA: Sage.

Lasswell, H. (1927). *Propaganda techniques in the world war*. New York: Knopf.

Madrigal, A. (2018). Inside Facebook's Fast-Growing Content-Moderation Effort. *The Atlantic*. Retrieved from https://www.theatlantic.com/technology/archive/2018/02/what-facebook-told-insiders-about-how-it-moderates-posts/552632/

Macnamara, J. (2003). Mass media effects: a review of 50 years of media effects research. Retrieved from http://www.archipelagopress.com/jimmacnamara

Mccumber, J. (2003). The Metaphysics of Clarity. In J. Culler & K. Lamb, Just Being Difficult?: Academic Writing in the Public Arena (pp. 58-70). USA: Stanford University Press.

Miles, M., & Huberman, M. (1994). *Qualitative data analysis*. California: Sage.

Neuendorf, K. (2002). *The Content Analysis Guidebook*, Thousand Oaks, CA: Sage Publications.

Neuman, W. (1997). *Social research methods: qualitative and quantitative approaches*. Needham Heights, MA: Allyn & Bacon.

Newig, J., Schulz, D., Fischer, D., Hetze, K., Laws, N., Lüdecke, G., & Rieckmann, M. (2013). Communication Regarding Sustainability: Conceptual Perspectives and Exploration of Societal Subsystems. *Sustainability*, *5*(7), 2976-2990. doi: 10.3390/su5072976

Nisbet, R. (1980). *History of the idea of progress* (p.224-29). New Brunswick, NJ: Basic Books, Inc.

PEN America. (2018). F*orbidden Feeds: Government Controls on Social Media in China* (p. 21). New York. Retrieved from https://pen.org/wp-content/uploads/2018/03/PENAmerica_Forbidden-Feeds-3.13-3.pdf

Pool, I. (1959). *Trends in content analysis*. Urbana: University of Illinois Press.

Roberts, S. (2014). *Behind the Screen: The Hidden Digital Labor of Commercial Content Moderation* (Mestre). University of Illinois.

Robinson, J. (2004). Squaring the circle? Some thoughts on the idea of sustainable development. *Ecological Economics*, *48*(4), 369-384. doi: 10.1016/j.ecolecon.2003.10.017

Ryan, A. (2012). *The making of modern liberalism* (p. 25). Princeton: Princeton University Press.

Schreier, M. (2012). *Qualitative content analysis in practice*. Thousand Oaks, CA: Sage.

Shoemaker, P. & Reese, S. (1996). *Mediating the message: theories of influences on mass media content*. White Plains, NY: Longman.

Statista. (2019). Media usage in an internet minute as of March 2019 [Image]. Retrieved from https://www.statista.com/statistics/195140/new-user-generated-content-uploaded-by-users-per-minute/

United Nations General Assembly. (2018). Report of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression (pp. 3-20). Special Rapporteur to the Human Rights Council. Retrieved from https://documents-dds-ny.un.org/doc/UNDOC/GEN/G18/096/72/PDF/G1809672.pdf?OpenElement

Vala, J. (1986). A análise de conteúdo. In A. Silva & J. Pinto, *Metodologia das ciências sociais* (pp. 101-128). Porto: Edições Afrontamento.

Weber, R. (1990). *Basic content analysis* (2nd ed.). Newbury Park, CA: Sage.

Weiss, R. (1994). *Learning from strangers: The Art and Method of Qualitative Interview Studies*. New York, NY: Free Press.

Wray, D. (2018). The Companies Cleaning the Deepest, Darkest Parts of Social Media. *Vice*. Retrieved from https://www.vice.com/en_us/article/ywe7gb/the-companies-cleaning-the- deepest-darkest-parts-of-social-media

**-- imagens e vídeos --**

Al Jazeera. (2016). *2016-09-16 Sarah T. Roberts, Al Jazeera English Tiziana Cantone* [Video]. Retrieved from https://www.youtube.com/watch?v=2VYzhmSJbwo

AP. (2003). *U.S. Military In Iraq* [Image]. Retrieved from https://www.cbsnews.com/pictures/us-military-in-iraq/2/

BBC iWonder. (2014). *Swastika use* [Image]. Retrieved from https://twitter.com/bbciwonder/status/525239433327828992?lang=ca

Camel, R. J. Reynolds Tobacco Company. (1946). *More Doctors Smoke Camels* [Image]. Retrieved from http://tobacco.stanford.edu/tobacco_main/images.php?token2=fm_st001.php&token1=fm_img0002.php&theme_file=fm_mt001.php&theme_name=

Canadian Patriotic Fund. (1916). *Moo-che-we-in-es. Pale Face, My skin is dark but my heart is white for I also give to Canadian patriotic fund* [Image]. Retrieved from http://www.begbiecontestsociety.org/Racism.htm

Chase & Sandborn. (2019). *If your husband ever finds out* [Image]. Retrieved from https://www.dailymail.co.uk/news/article-2254806/Didnt-I-warn-serving-bad-coffee-Outrageously-sexist-ads-1950s-shocking-domestic-scenes-subservient-women-carrying-domestic-duties-husbands.html

Courbet, G. (1866). L'Origine du monde [Image]. Retrieved from https://www.artlyst.com/news/facebook-courbet-censorship-lawsuit-day-paris-court/corbet-mash/

dopl3r. (2019). *Soldiers overseas* [Image]. Retrieved from https://en.dopl3r.com/memes/dank/safe-to-say-this-guy-is-not-going...?page=90

E. Kealey. (1915). *Women of Britain say - GO! (Military recruitment poster)* [Image]. Retrieved from https://commons.wikimedia.org/wiki/File:7_Collection_Eybl_Great_Britain_-_E._Kealey_-_Women_of_Britain_say_%E2%80%93_GO.jpg

Elliott's Paint. (1930). *White Veneer* [Image]. Retrieved from http://vintagenewsdaily.com/racist-sexist-and-dishonest-vintage-advertisements-that-seem-shocking-today/

Heather Halls. (1957). *Cocaine Candy* [Image]. Retrieved from https://issuu.com/vaneatongalleries/docs/crumpissuu/21

iFunny. (2018). *When you haven't robbed in a week* [Image]. Retrieved from https://ifunny.co/picture/when-you-haven-t-robbed-for-a-week-bCTw2ARF6

Jockey Junior Brief. (1955). *They keep their fit!* [Image]. Retrieved from https://atomic-flash.tumblr.com/post/98089249294/1955-jockey-junior-briefs-advert-what-does-this

Kamensky, M. (2015). *Cartoon* [Image]. Retrieved from https://www.cartoonstock.com/directory/i/irish_referendum.asp

*Kid and a boot*. (2018). [Image]. Retrieved from https://twitter.com/cnnpolitics/status/1079943567378915328

Lewis County. (1855). *Great Sale of Slaves* [Image]. Retrieved from https://www.reddit.com/r/pics/comments/8bozdz/great_sale_of_slave_advertisement_from_1855/

Mckay, H. (2018). *Prince William* [Image]. Retrieved from https://awarenessact.com/the-truth-hurts-15-examples-of-how-the-media-manipulates-the-truth/

Moral Majority Report. (1983). *AIDS: Homosexual Diseases Theraten American Families*[Image]. Retrieved from https://www.gayinthe80s.com/2016/08/1983-aids-and-the-moral-majority/

PainMagnetGaming. (2009). *North Korean soldiers* [Image]. Retrieved from http://istoryporn/comments/9g23px/north_korean_soldiers_escort_kim_jongun_to_safety/

*Rock-Paper-Scissors*. [Image]. Retrieved from https://me.me/i/playing-rock-paper-scissors-with-the-squad-tjustweirdothings-looks-8190452

*SnapChat*. (2017). [Image]. Retrieved from https://www.independent.co.uk/news/world/americas/snapchat-racist-twitter-sikh-terrorism-muslim-islamophobia-airplane-passenger-outrage-racism-hate-a7807161.html

Splitpics.uk. (2019). *They've cracked it* [Image]. Retrieved from https://www.thesun.co.uk/news/8689562/it-may-take-you-a-while-to-work-out-whats-going-on-in-these-mind-bending-photos/

Time Magazine. (1947). *DDT is Good for Me* [Image]. Retrieved from https://enviroethics.org/2011/06/18/animation-ddd-is-good-for-me/

Van Heusen. (1952). *The world's smartest shirts* [Image]. Retrieved from http://vintagenewsdaily.com/racist-sexist-and-dishonest-vintage-advertisements-that-seem-shocking-today/

Winston Cigarettes. (1962). *What's up front that counts* [Image]. Retrieved from https://www.vintage-adventures.com/vintage-tobacco-ads/4896-1962-winston-cigarettes-ad-what-s-up-front.html

WPA War Services of LA. (1942). *Censored, Let's censor our conversation about the war*[Image]. Retrieved from https://www.publicdomainpictures.net/en/view-image.php?image=77568&picture=vintage-war-censorship-poster

Wuebker, E. (1940). *Say No to Prostitutes* [Image]. Retrieved from https://www.vice.com/en_ca/article/gq8dq4/evolution-of-safe-sex-propaganda

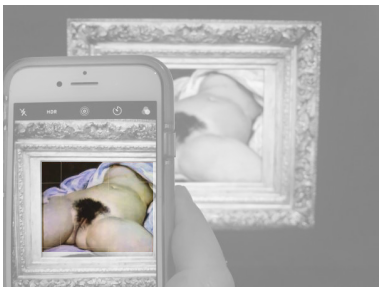
## ANEXO A: EXEMPLOS (CONTEÚDOS – VISÃO-CONTENTOR)

**McKay, 2018**

**AP, 2003**

**[kid and a boot], 2018**

**Courbet, 1866**

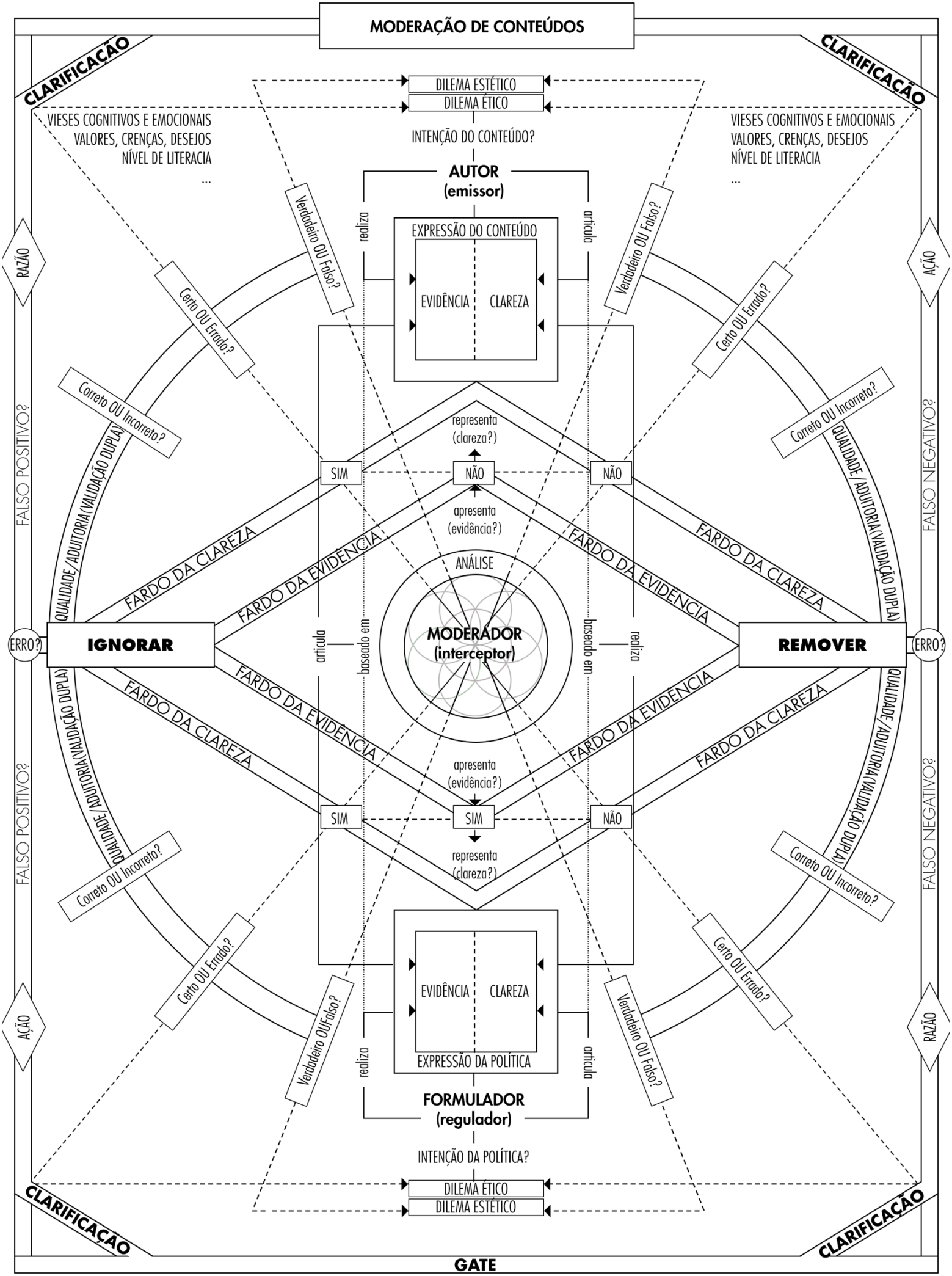


**Splitpics.uk, 2019**

**PainMagnetGaming, 2009**

**Dopl3r, 2019**

# ANEXO B: QUADRO CONCEPTUAL (MODERAÇÃO DE CONTEÚDOS)



## ANEXO C: CONTEÚDOS (ANALISADOS)

### 1. Suástica 卐 ou 卍

### 2. Sátira e Pecado


### 3. SnapChatice


### 4. Pensa Positivo!


### 5. Quem ganha?


### 6. Análise Crónica

## ANEXO D: CONTEÚDOS (RESPOSTAS I)

X = Remover

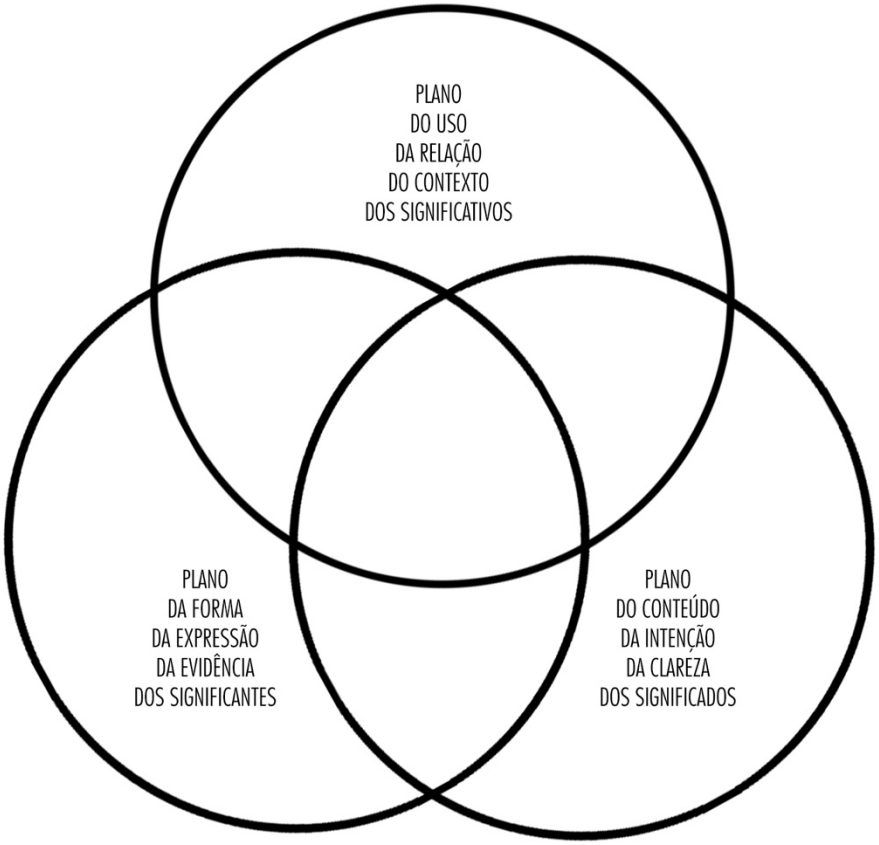
V = Manter

**CONTEÚDOS**

| MODERADORES | 1. Suástica 卐 ou 卍 | 2. Sátira e Pecado | 3. Snap Chatice | 4. Pensa Positivo! | 5. Quem Ganha? | 6. Análise Crónica... |
|---|---|---|---|---|---|---|
| 1 | X | X | V | V | V | X |
| 2 | V | V | X | V | V | X |
| 3 | V | X | X | V | V | V |
| 4 | X | V | X | V | X | X |
| 5 | X | X | X | X | X | X |
| 6 | V | X | X | V | X | X |
| 7 | V | V | X | V | X | X |
| 8 | X | X | X | V | V | X |
| 9 | V | V | V | V | X | V |
| 10 | V | V | X | V | X | X |
| 11 | X | V | V | V | V | V |
| 12 | V | X | X | V | V | X |
| **TOTAL** | X = 5 V = 7 | X = 6 V = 6 | X = 9 V = 3 | X = 1 V = 11 | X = 6 V = 6 | X = 9 V = 3 |

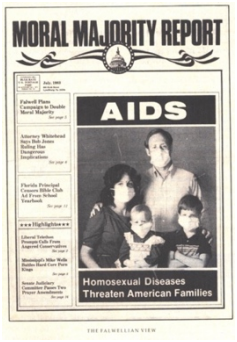
## ANEXO E: QUADRO CONCEPTUAL (MODERAÇÃO + SUSTENTÁVEL)

**MODERAÇÃO + SUSTTENTÁVEL**



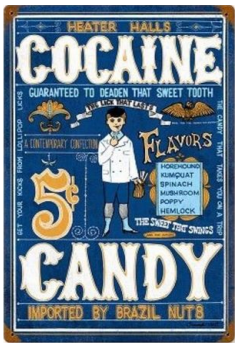
## ANEXO F: EXEMPLOS (CONTEÚDOS ANTIGOS)

**Moral Majority Report, 1983**




**Van Heusen, 1952**


**Heather Halls, 1957**




**Camel, 1946**


**Jockey Junior Brief, 1955**

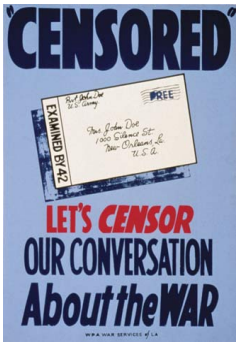

**Winston Cigarettes,**


**Elliot's Paint, 1930s**


**Time Magazine, 1947**

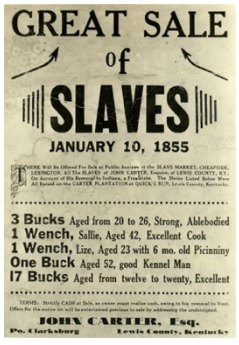

**Kealey, around 1915**


**WPA War Services, 1955**

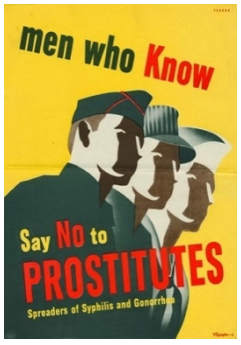



**Canadian Patriotic Fund,**


**Lewis County, 1855**




**Wuebker, 1940s**




**Chase & Sandborn,**


**Moral Majority Report,**