

[Da governança comunitária ao atendimento ao cliente e vice-versa](#)

Ethan Zuckerman e Chand Rajendra-Nicolucci

Data da publicação:

Janeiro 2024

Reexaminando modelos pré-Web de governança online para enfrentar a crise de legitimidade das plataformas

Desde os primeiros dias da computação, as pessoas usam a tecnologia da informação para conversar umas com as outras. Quatro anos antes da Internet, Noel Morris e Tom Van Vleck escreveram um sistema de correio eletrônico e um sistema de bate-papo em tempo real para o Sistema Compatível de Compartilhamento de Tempo (*Time-Sharing*), CTSS do MIT, permitindo que usuários que se conectassem ao único computador compartilhado deixassem mensagens para um outro ou enviassem mensagens para o terminal de outro usuário (Van Vleck, 2012). Três anos após a introdução da Internet, o e-mail tornou-se o principal uso de uma rede inicialmente estabelecida para permitir que cientistas da computação executassem programas em máquinas remotas (Sterling, 1993). O serviço Minitel francês, concebido para dar aos usuários acesso a uma lista telefônica eletrônica e a capacidade de efetuar reservas de viagens online, tornou-se rapidamente dominado por serviços de chat, particularmente chat erótico (Tempest, 1989). As pessoas querem conversar e encontrarão maneiras de fazê-lo assim que forem tecnicamente capazes de se conectar umas com as outras.

Infelizmente, assim que as pessoas conseguem falar umas com as outras, elas também podem prejudicar umas às outras. O spam minou a utilidade do e-mail e destruiu em grande parte a Usenet, a plataforma comunitária dominante da Internet acadêmica na década de 1980 e no início da década de 1990. O assédio e o discurso de ódio tornaram-se realidade para os usuários de muitos sistemas online, especialmente para mulheres, pessoas de cor e pessoas LGBTQIA+. As pessoas muitas vezes comportam-se de forma diferente online e offline (Suler, 2004) e o ímpeto para os humanos assediarem uns aos outros através de ferramentas digitais é pelo menos tão forte quanto o impulso para se conectarem.

O surgimento da confiança e da segurança como disciplina profissional reflete a centralidade de questões como moderação de conteúdo, prevenção de spam e fraude, bem como esforços para combater imagens de abuso sexual infantil (CSAM) na operação de plataformas que permitem conteúdo e conversação gerados pelos usuários. Como observa Tarleton Gillespie (2018) em *Custodians of the Internet*, "As plataformas não são plataformas sem moderação". Os esforços recentes para reconhecer a confiança e a segurança como profissão, com a criação da Associação Profissional de Confiança e Segurança (Trust Safety Professional Association) em 2020 e o surgimento de um *Journal of Online Trust and Safety* em 2021, estão atrasados, uma vez que o trabalho de monitoramento de espaços online remonta pelo menos até a década de 1980, se não antes.

Um perigo de perder a história inicial da governança online é o estreitamento dos futuros possíveis, fazendo parecer que o modelo contemporâneo de governança de espaços online, onde os profissionais tomam decisões sobre que comportamento é aceitável, com pouca contribuição dos membros da comunidade, é da maneira que sempre foi feito. Referimo-nos a este modelo como o modelo de "serviço ao cliente" e comparamo-lo com modelos anteriores de governança online, nos quais os membros da comunidade eram decisores importantes, embora nem sempre primários, sobre os espaços online dos quais faziam parte.

Este artigo examina três paradigmas de governança online que precederam o modelo contemporâneo de atendimento ao cliente e sugere que vários graus de governança comunitária podem ser uma opção viável e

socialmente benéfica para muitos espaços online.

O artigo está longe de ser uma história exaustiva dos primórdios da governança online ou do surgimento do modelo de serviço ao cliente, embora ambas as histórias sejam necessárias. Embora tenha havido um excelente trabalho chamando a atenção para as complexidades da confiança e da segurança (Gillespie, 2018; Gray Suri, 2019), este concentrou-se principalmente nas plataformas de mídia social Web 2.0” que surgiram em meados da década de 2000 – a mudança do modelo de atendimento ao cliente começa no final da década de 1980 e é consolidada em meados da década de 1990. Esta também é uma história pessoal e opinativa, já que um dos autores (Zuckerman) construiu o primeiro departamento de moderação de conteúdo do *tripod.com*, um dos primeiros sites de conteúdo gerado pelo usuário, de 1995 a 1999.

Antes da Web

Muito antes de a World Wide Web se tornar a forma dominante de interação nas redes de computadores, o conteúdo gerado pelo usuário era a forma dominante de material compartilhado online. Uma grande atração, se não a principal, dos serviços de Internet discada, como AOL e CompuServe, era a capacidade de interagir com outros usuários em fóruns on-line e salas de bate-papo. Os espaços nativos dos primórdios da Internet acadêmica – grupos Usenet, MUDs e MOOs (mundos virtuais multiusuários) – lidavam inteiramente com conteúdo gerado pelo usuário, já que a Internet acadêmica reunida em torno da NSFNET não estava aberta a usos comerciais antes de 1991 (RFC 1192, 1990) .

Estas comunidades online estavam repletas de mau comportamento: grande parte da linguagem que usamos hoje para discutir mau comportamento online – trolling, flaming e spam – tem origem na cultura inicial da Internet e dos BBS (quadros de avisos em rede). Na maioria dos espaços online, surgiu um sistema de governança para abordar o comportamento anti-social. Vale a pena considerar pelo menos três paradigmas como antecedentes dos modelos contemporâneos de governança online: baseado em normas, baseado na comunidade, e moderação efetiva por operador de BBS (sysop).

Usenet – Governança Baseada em Normas

A Usenet foi criada em 1979 pelos estudantes de graduação Tom Truscott e Jim Ellis na Duke University, usando um protocolo simples (UUCP) para copiar arquivos entre sistemas Unix, permitindo o compartilhamento de software e arquivos de texto. No final da década de 1990, a Usenet conectava pessoas em mais de 200 países e era responsável por um terço da transferência de dados da Internet (M. Smith, 1999).

Os grupos da Usenet foram inicialmente não moderados, e a capacidade de moderar um grupo de notícias foi adicionada apenas em 1984. Arquitetonicamente, a Usenet foi projetada para colocar o maior controle sobre a moderação principalmente nas mãos de um leitor – usuários frequentes da Usenet tornaram-se adeptos do uso de um killfile” preenchido com os nomes dos tópicos e mensagens que se deseja filtrar. (Killfiles – The cure for all that ails you”, 1994)

Além do poder de bloqueio dos killfiles, os usuários da Usenet tinham um grande poder afirmativo ao propor novos grupos de notícias. Usuários de notícias motivados e com conhecimento técnico podiam propor um novo grupo postando uma solicitação de discussão” em *news.announce.newgroups*, que incluía uma justificativa para o novo grupo e um regulamento sobre o que o seria discutido no mesmo. Seguiu-se um período de discussão de no mínimo três semanas, seguido de uma chamada à votação. Os grupos de notícias seriam criados se recebessem dois terços de votos Sim” e pelo menos 100 votos a mais de Sim” do que Não” (Paolillo Heald, 2002). Em última análise, o poder de criar grupos de notícias estava nas mãos dos administradores de servidores, mas normas fortes sobre a participação democrática significavam que os usuários eram poderosos na condução da evolução das discussões da Usenet.

Os limites da arquitetura da Usenet foram testados quando um dos trolls mais prolíficos da Internet, Serdar Argic, começou a usar software de publicação automática (ou seja, um bot”) para responder a qualquer menção à Turquia ou à Arménia na Usenet, com longas argumentações negando o genocídio arménio. Argic postou uma média de 100 postagens por dia, muitas delas direcionadas de forma inadequada (por exemplo, em resposta a uma discussão sobre um peru – turkey” em inglês -- de Ação de Graças). Embora os administradores pudessem ter decidido cancelar” todas as mensagens provenientes da conta de Argic, uma norma forte contra o

cancelamento de terceiros” como forma de censura prevaleceu na Usenet, tornando esta uma solução insustentável. Um administrador da Usenet acabou criando um novo grupo de notícias – *alt.cancel.bots* – que cancelaria automaticamente as postagens de Argic, mas apenas se o usuário optasse por se inscrever no grupo de notícias (DeVoto, 1994).

Como as possibilidades técnicas da Usenet colocaram forte controle nas mãos de do usuário individual, grande parte da governança da Usenet aconteceu por meio de rígidas normas comunitárias. Muitos grupos de notícias da Usenet incluíam um documento de resposta a perguntas frequentes (FAQs), que tinha como objetivo evitar que novos usuários fizessem perguntas repetitivas e respondidas anteriormente. Netiqueta”, uma aglutinação de rede” (net”) e etiqueta”, foi provavelmente cunhada na Usenet como uma forma de falar sobre as regras sociais que tornaram as discussões online menos hostis e mais construtivas (Merriam-Webster).

A dependência da netiqueta pela Usenet reflete-se na história do eterno setembro”. Os usuários da Usenet queixavam-se frequentemente da qualidade do discurso online em setembro, quando novos estudantes chegavam ao campus, passavam a ter acesso à Internet e começavam a irritar os usuários existentes com a sua ignorância sobre netiqueta. Quando a AOL começou a fornecer acesso à Usenet a milhões de seus clientes em setembro de 1993, o experiente usuário da Usenet Dave Fischer declarou que aquele era o setembro que nunca terminou” (Koebler, 2015; E. Smith, 2020). Embora elitista e condescendente, a noção de eterno setembro” capturou algo real e verdadeiro sobre as vulnerabilidades da Usenet: um espaço governado principalmente por normas rígidas era vulnerável aos excessos dos infratores.

Pouco depois de a AOL acolher uma base de usuários muito maior na Usenet, o spam se tornou um problema sério. O escritório de advocacia Canter Siegel publicou um anúncio oferecendo serviços de imigração em 5.500 grupos de notícias, sinalizando a vulnerabilidade da comunidade a anúncios postados de forma cruzada, que sobrecarregavam o tráfego de notícias dos indivíduos e sobrecarregavam os servidores da Usenet. Alguns administradores de sistemas da Usenet, que anteriormente respeitavam as fortes normas de liberdade de expressão da comunidade, iniciaram campanhas para bloquear spammers prolíficos, enquanto outros argumentavam que as normas da Usenet proibiam tais ações em todo o sistema (Miller, 2021).

O poder dos administradores de sistemas para bloquear unilateralmente os spammers foi um lembrete de que todos os espaços mediados tecnologicamente experimentam uma tensão entre as intenções de quem controla tecnicamente o espaço e a comunidade que escolhe usar o espaço. Os administradores da Usenet deram aos indivíduos uma grande responsabilidade e controle sobre o seu comportamento e colocaram em jogo estruturas de votação significativas, mas tiveram que tomar decisões unilaterais quando os violadores das normas começaram a tornar o espaço inutilizável. A Usenet tornou-se um espaço culturalmente menos significativo em meados da década de 1990, tanto devido ao surgimento da Web gráfica quanto a problemas de spam.

MOOs e MUDs – Governança Baseada na Comunidade

Um rito de passagem para a maioria dos programadores é a criação de um jogo. Muitos programadores nas décadas de 1970 e 1980 escreveram jogos, influenciados pelos livros de JRR Tolkien e pelo RPG *Dungeons and Dragons*. Esses jogos são chamados de rastreadores de masmorras” e passaram por um rápido processo de iteração em minicomputadores de acesso compartilhado em universidades no final da década de 1970. Em 1978, um estudante da Universidade de Essex, no Reino Unido, criou uma masmorra que poderia ser explorada por múltiplos usuários, batizada de Multi User Dungeon”, ou MUD (Kelly Rheingold, 1993).

Os MUDs tornaram-se enormemente populares nos campi universitários na década de 1980, com alguns brincando que a sigla significava Multi-Undergraduate Destroyer”, já que alguns estudantes ficaram tão imersos nos mundos virtuais que deixaram de frequentar as aulas (When Studying MUDs [. . .], 1998). Alguns MUDs permaneceram próximos de suas raízes de exploração de masmorras, enquanto outros se concentraram em elementos sociais, permitindo aos usuários conversar, interagir e construir juntos novos recursos de um mundo virtual. Um dos mais populares desses MUDs sociais foi o LambdaMOO, hospedado no Centro de Pesquisa da Xerox em Palo Alto e criado pelo pesquisador Pavel Curtis em 1990.¹ O LambdaMOO tornou-se um ponto de encontro popular para aqueles interessados em novas interações sociais on-line e foi amplamente documentado por autores de tecnologia, incluindo Howard Rheingold (1993) e Sherry Turkle (1995).

LambdaMOO (que ainda existia no final de 2022) se assemelha a um conjunto de salas de bate-papo conectadas, com alguns recursos poderosos escondidos abaixo da superfície. Cada sala possui uma descrição textual, assim

como cada personagem; como disse o jornalista Julian Dibbell (1993), LambdaMOO é uma mansão rústica muito grande e muito movimentada construída inteiramente de palavras”. Além de conversarem entre si, os usuários do LambdaMOO podem construir novas partes do ambiente, criando novos espaços e objetos que podem ser codificados para realizar comportamentos. Por exemplo, uma sala pode ter um alçapão – quando um usuário puxa a alavanca, os usuários da sala caem” para outra sala. A combinação de chat, descrição textual e codificação torna o LambdaMOO um espaço poderosamente envolvente, apesar da falta de gráficos.

O poder sobre a tomada de decisões no LambdaMOO era detido por Curtis e uma pequena equipe de assistentes”, que tinham poderes administrativos e técnicos especiais além dos usuários normais. No início de 1993, Curtis anunciou à comunidade (Mnookin, 1996):

Percebo agora que a comunidade LambdaMOO atingiu um nível de complexidade e diversidade que eu realmente esperava e queria desde que quatro hackers e eu decidimos construir este lugar: esta sociedade saiu do ninho. . . Então, como a última decisão social que tomamos para vocês, e quer vocês, adultos independentes, queiram ou não, os magos estão se retirando do negócio de disciplina/boas maneiras/arbitragem; estamos transferindo o fardo e a liberdade desse papel para a sociedade em geral.

Embora declarasse o seu desejo de entregar o controle aos usuários, Curtis não propôs um mecanismo de governança real e, durante alguns meses, a comunidade ignorou largamente sua decisão.

Em março de 1993, a comunidade LambdaMOO deparou-se um caso flagrante de abuso sexual, no qual um usuário – o Sr. Bungle — usou um trecho de código (um boneco de vodu”) para fazer os jogadores se envolverem em ações sem o seu consentimento. Ele tinha como alvo usuários que se apresentavam como mulheres e forçava seus personagens a situações violentas e sexuais. Num artigo que descreve o ataque e as suas consequências, intitulado “A Rape in Cyberspace”, o jornalista Julian Dibbell (1993) examinou como a comunidade LambdaMOO lutou com as complexidades da violação virtual e da agressão sexual. Uma longa discussão na lista de e-mail principal do LambdaMOO convergiu em torno da decisão de remover o Sr. Bungle do sistema, e um assistente excluiu sua conta.

O caso do Sr. Bungle acelerou a formação de processos de governança no LambdaMOO. Dibbell (1993) descreve discussões contínuas entre aqueles que queriam um sistema legal formal para governar o espaço e aqueles que assumiram posições anarquistas ou libertárias. Jennifer Mnookin (1996) descreve o rápido desenvolvimento de um sistema de governança complexo, coloquialmente referido como Lei Lambda, que inclui um processo de petição para alterações sociotécnicas a serem feitas no ambiente, e um processo detalhado de mediação para conflitos entre usuários. Em 1996, os usuários do LambdaMOO aprovaram 44 dessas petições, fazendo mudanças permanentes nas possibilidades do sistema, incluindo o estabelecimento do processo de mediação (Mnookin, 1996).

Tal como acontece com a Usenet, o controle sobre os sistemas técnicos cabia aos assistentes” capazes de controlar o código e o banco de dados. Mas, numa extensão ainda maior do que com a Usenet, aqueles que detinham autoridade técnica sobre a plataforma optaram por sujeitar esse poder à vontade da comunidade em geral.

BBS e serviços online – o Sysop forte

Kevin Driscoll (2022b) sugere que as primeiras histórias da comunidade da Internet muitas vezes focam demais em comunidades como Usenet e MUDs/MOOs, que eram populares entre o público acadêmico, em detrimento de comunidades hospedadas em quadros de avisos eletrônicos (BBS) – sistemas administrados por indivíduos ou pequenas equipes de computadores domésticos conectados a pequenos bancos de modems. Nestes mundos, o operador do quadro de avisos – o administrador de sistema – era o rei ou a rainha. Mas poucos destes monarcas se consideravam onipotentes: em vez disso, estavam profundamente sintonizados com as necessidades e desejos das suas comunidades, para que os seus usuários não desistissem participar.

Embora esses sistemas fossem frequentemente operados como empresas comerciais, Driscoll argumenta que eles tinham uma vasta gama de modelos de negócios, desde administradores de sistemas que tratavam seus BBSs como um hobby caro, como possuir um carro antigo, até administradores de sistemas que dependiam fortemente de suas comunidades para apoio financeiro e participação. Driscoll (2022a) explica melhor:

Havia também BBSs organizados como um clube social. Os membros pagavam taxas” para manter o disco rígido girando. Outros formaram corporações sem fins lucrativos, solicitando doações isentas de impostos de seus usuários. Mesmo nos BBSs informais, os administradores de sistema às vezes passavam o chapéu virtual, pedindo a todos alguns dólares para comprar um novo modem ou pagar uma grande conta de telefone.

Essa dependência dos usuários para suporte financeiro criou um ciclo de realimentação entre usuários e administradores de sistemas. Um sysop que não ouvisse seus usuários provavelmente não encontraria suporte para expandir seu sistema. Mais provavelmente, os usuários partiriam para outro BBS no mesmo DDD. Como resultado, mesmo os fóruns que eram projetos não comerciais e amadores gastavam tempo e energia ouvindo seus usuários. Driscoll (2022b, p. 148) discute um BBS em Terre Haute, Indiana, chamado TARDIS, administrado por quatro amigos como hobby, que trabalhava para surpreender e encantar seus visitantes”. O TARDIS criou uma comunidade popular entre mulheres (para as quais forneceu uma área somente para mulheres”, como um corretivo útil para espaços online dominados por homens) e para usuários com deficiência visual (mantiveram seu software simples e compatível com leitores de tela). As recompensas por ouvir as necessidades dos usuários foram um grupo central dedicado de usuários que se tornaram extremamente leais e investiram pessoalmente na cultura do BBS” (Driscoll, 2022b, p. 151) e que trabalharam para promover o BBS e recrutar pessoas com ideias semelhantes.

O modelo do sysop todo-poderoso, que escuta atentamente seus usuários, acabou influenciando até os maiores sistemas comerciais. O estudioso da Internet Jonathan Zittrain atribui seu fascínio pelos espaços online ao seu tempo como administrador de sistema na Compuserve, trabalhando para pagar suas substanciais taxas de cobrança online. A Compuserve, explica Zittrain (comunicação pessoal, 22 de novembro de 2022), terceirizou as funções de gerenciamento da comunidade para um conjunto de prestadores de serviços, que eram pagos de acordo com a quantidade de tráfego que seus painéis geravam. No contexto de tornar seus fóruns eletrônicos populares e com forte demanda, os sysops eram muito flexíveis na forma como desempenhavam suas funções. Alguns contratavam auxiliares e os remuneravam nas horas vagas do sistema ou com parcela de seus rendimentos. Outros trabalharam o mínimo possível.

Zittrain (1997) observa que muitos administradores de sistemas descobriram que as formas mais produtivas de gerir as suas comunidades envolviam restringir os seus impulsos autoritários:

Quando ocorre a intervenção de um administrador de sistema, muitas vezes ela ocorre de forma ponderada, com noções como o devido processo ou tolerância a discursos desagradáveis gerados no calor das conversas ou derivados do cenário cultural e jurídico ocidental que ainda paira sobre a Internet de hoje.

Nisso, os administradores de sistemas estavam antecipando o paradigma da liberdade de expressão que muitas grandes plataformas de mídia social professavam inicialmente. Mas também enfatizaram a necessidade de um sistema justificável e defensável graças à dinâmica dos BBS: um fórum governado injustamente é aquele que os usuários abandonarão, afetando os ganhos ou o trabalho do sysop.

A relação entre um sysop forte e a sua comunidade não é um modelo democrático. Está mais perto de uma ditadura benevolente. Mas, tal como Rebecca MacKinnon (2012) sugere em *Consent of the Networked*, uma monarquia que reconhece um conjunto de direitos básicos para os cidadãos é mais saudável e mais estável do que aquela que procura que os seus servos fiquem em silêncio. Os sistemas burocráticos de governança que surgiram rapidamente nos primeiros anos da Web (1994-2000) abandonaram frequentemente o reconhecimento de que os usuários de comunidades online tinham um papel crítico na sua criação, manutenção e governança.

Das Comunidades às Burocracias

A Web inicial cresceu a uma velocidade que muitas vezes chocava aqueles que construíam as suas principais infraestruturas. *Our World In Data* (Roser et al., 2015) estima a população total de usuários da Internet em 1990 (o apogeu da Usenet) em 2,6 milhões, em 1995 (primeiros dias da web comercial) em 44,4 milhões, e em 2000 (o estouro da primeira bolha” da Internet) em 412,8 milhões.

Zuckerman fez parte da equipe fundadora que construiu o *tripod.com*, um site que ganhou popularidade ao oferecer páginas iniciais gratuitas na Internet para usuários registrados. O tráfego para essas páginas iniciais significou que o Tripod foi um dos dez sites com maior tráfego no final da década de 1990, bem como um dos primeiros negócios de conteúdo gerado pelo usuário com suporte de anúncios.² O Tripod estava totalmente

despreparada para os desafios de gerenciar uma comunidade on-line: começou como um site de conteúdo editado, semelhante a uma revista, voltado para recém-formados, e acrescentou a construção de *home pages* como uma reflexão tardia. Descobrir que os usuários usariam suas páginas da Web para compartilhar pornografia, imagens violentas e software pirata foi óbvio em retrospecto, mas surpreendente para aqueles envolvidos com o Tripod à medida que o site crescia.

Zuckerman (que não é advogado) escreveu os termos de serviço do Tripod no final de 1995, para sinalizar quais comportamentos seriam proibidos no Tripod e, na esperança de criar uma forma de devido processo legal para a equipe de abuso” do Tripod (um ramo do departamento de atendimento ao cliente), começou a remover sistematicamente as páginas iniciais dos usuários e a cancelar contas. O Tripod não previu problemas como a ocorrência de CSAM e, quando o CSAM surgiu no site, a equipe da Tripod teve que descobrir como relatá-lo ao FBI. O FBI solicitou que o CSAM não fosse enviado por e-mail, pois levantou questões espinhosas sobre se os servidores Tripod e do FBI fossem usados para transmitir CSAM. Em vez disso, um funcionário do Tripod conduzia periodicamente pilhas de disquetes por 240 quilômetros até o escritório de campo do FBI mais próximo.

Outras empresas iniciais da Web também não estavam preparadas para usos não intencionais de seus softwares. Nicole Wong, que se tornou vice-presidente e consultora jurídica adjunta do Google, disse a Robyn Caplan (*Data Society Podcast*, 2020) que suas primeiras experiências com o que hoje é conhecido como confiança e segurança vieram do trabalho no Craigslist, fundado em 1995. O departamento de atendimento ao cliente do Craigslist naquela época era Craig Newmark, o fundador da empresa, que continuou a responder a maior parte dos e-mails de atendimento ao cliente nos anos seguintes. Quando os assuntos atingiram um nível de complexidade que Craig – essencialmente como administrador de sistemas – se sentiu desconfortável em lidar, eles passaram para Wong, como advogada da empresa (*Data Society Podcast*, 2020).

Wong também descreveu o surgimento de confiança e segurança em seu emprego seguinte, no Google. Wong lembrou (Caplan, 2018): Efetivamente não houve moderação. Havia pessoas de suporte ao cliente que respondiam às perguntas, mas não havia políticas de moderação verdadeiras, por si só, exceto para direitos autorais e pornografia infantil.” Caplan (2018) continua: Com o tempo, [Wong] disse, eventos importantes (como o caso Yahoo v. LICRA, relativo à disponibilidade de parafernália nazista em todo o mundo) desempenharam um papel relevante no desenvolvimento de políticas, com a formalização ocorrendo ao longo do tempo conforme a plataforma crescia.”

A emergência da confiança e da segurança como um departamento profissional dentro de uma plataforma online, abrangendo moderação de conteúdo, conformidade legal e estratégias para minimizar abusos e danos, parece acontecer principalmente como uma reflexão tardia. Os membros da equipe de confiança e segurança da Zoom captam bem esse padrão em sua descrição da criação de uma equipe de confiança e segurança durante a pandemia do coronavírus (Maxim et al., 2022):

As equipes de Confiança e Segurança (C&S) geralmente nascem em uma crise. Com base nas nossas discussões com outras empresas, parece raro que os executivos de tecnologia acordem um dia e pensem: No próximo trimestre devemos formar uma equipe de C&S”. É o que você faz quando algo ruim já aconteceu. Talvez você perceba muitos golpistas de criptomoeda entrando em contato com seus usuários, ou comentários falsos direcionando os usuários para fora da plataforma para entregar credenciais de login, ou seu aplicativo se tornou o ponto de encontro do dia para uma comunidade de zoófilos. Você pega quem puder para resolver o problema imediatamente, e é daí que vêm as equipes de C&S. . . A Zoom ampliou e formalizou rapidamente sua equipe de C&S na primavera de 2020, numa época em que estávamos crescendo em um ritmo alucinante e, ao mesmo tempo, conquistando uma base de usuários global maior. De repente, as pessoas estavam usando o Zoom de maneiras muito além do uso comercial para o qual foi projetado.

Esta evolução de decisões ad hoc tomadas por fundadores e equipes de atendimento ao cliente do tipo sysop para políticas legalmente defensáveis e formalizadas faz sentido. À medida que a base de usuários de serviços aumenta, os riscos aumentam. Empresas como a Craigslist – que começou como uma lista de discussão local administrada como um hobby – tornam-se empresas lucrativas com investidores, funcionários e outras partes interessadas. O conteúdo que viola as regras agora pode atingir centenas de milhares de usuários, em vez de centenas.

O que é menos claro é a razão pela qual as primeiras plataformas Web adotaram tão rapidamente um modelo de confiança e segurança que tinha as suas raízes no serviço ao cliente, e não na governança da comunidade. Embora muitas das pessoas envolvidas na construção das primeiras plataformas Web tenham crescido em

espaços online pré-Web, como Usenet, MOOs, MUDs e BBSs, os modelos de governança comunitária falharam em grande parte em fazer a transição dos espaços pré-Web para os primeiros espaços comerciais da Web .

Aqui estão algumas razões pelas quais empresas como a Tripod – cujo sistema foi construído por programadores experientes no mundo governado pela comunidade de MOOs e MUDs – adotaram inquestionavelmente modelos de governança burocráticos e orientados ao atendimento ao cliente para seus novos espaços online.

- A escala desses novos espaços era esmagadora. Espaços pré-Web como LambdaMOO tiveram algumas centenas de participantes regulares. Os grupos de notícias ativos da Usenet podem ter um núcleo de 50 a 200 postadores ativos e um grupo maior de leitores e postadores ocasionais. As ferramentas para controlar spam e trollagem com esses pequenos públicos podem não ter funcionado com os milhares de novos usuários que migraram para as primeiras plataformas da Web. Pedir aos usuários que ajudassem a definir as regras para um espaço que crescia aos milhares por dia parecia um luxo com o qual essas empresas não podiam lidar.
- Empresas como a Tripod, que obtiveram financiamento de capital de risco, depararam-se com questões sobre a sua capacidade de ampliar a escala de suas operações. Embora as empresas baseadas em conteúdos gerados pelos usuários fossem muito mais baratas de operar do que aquelas que encomendam conteúdos gerados profissionalmente, envolver diretamente os usuários na governança poderia tornar financeiramente inviáveis esses empreendimentos. Demonstrar que as comunidades poderiam ser governadas de forma eficiente com uma pequena equipe – ou terceirizadas – tornou-se parte do processo de limitação de risco dos investidores.
- A responsabilidade legal tornou-se uma grande preocupação para as empresas de conteúdo gerado pelo usuário. Empresas como a Tripod, que começou a hospedar conteúdo antes da Lei de Telecomunicações de 1996 nos EUA, e do porto seguro” da Seção 230 que a acompanha,³ tinham poucas garantias de que estariam isentas de responsabilidade por difamação ou outras reclamações sobre danos causados por conteúdo online. À medida que aumentavam as violações de direitos autorais e as solicitações de remoção oriundas do DMCA⁴ e que o CSAM se tornava um problema sério nas plataformas de conteúdo gerado por usuários, a necessidade de os profissionais responderem rapidamente ao conteúdo infrator tornou-se mais premente.
- Os mecanismos de governança comunitária da Internet pré-Web presumiam pequenas comunidades de usuários com ideias semelhantes. Espaços como a Usenet e os quadros de avisos presumiam um interesse comum. Os primeiros usuários da Internet poderiam ter sido capazes de imaginar que na Internet, ninguém sabe que você é um cão” (Steiner, 1993), mas estatisticamente falando, um usuário da Internet antes de 1990 poderia assumir com segurança que qualquer pessoa com quem estivesse interagindo era um estudante, um recém-formado ou funcionário de tecnologia em uma universidade em um país rico do Norte global. Seria realista acreditar que uma comunidade” de usuários anônimos sem antecedentes ou história comum pudesse ser autogestionada?

Quer estes fatores expliquem adequadamente ou não a mudança de paradigma da governança comunitária para o serviço ao cliente, essa mudança já tinha ocorrido quando a maioria dos usuários passaram a estar online na Internet.

As implicações da mudança da governança comunitária para o atendimento ao cliente

Os modelos de governança que surgiram nos sistemas Usenet, LambdaMOO e BBS, onde aqueles com controle técnico sobre os sistemas comunitários cederam parte desse controle às suas comunidades, ou pelo menos se posicionaram como ouvintes das necessidades de suas comunidades, desapareceram na história da Internet

quando plataformas como Facebook, Twitter e YouTube passaram a dominar. Em 2012, as analogias de Rebecca MacKinnon comparando plataformas de redes sociais a monarquias poderiam ter parecido dramáticas, mas não erradas. Em 2006, quando o Facebook decidiu alterar a experiência dos usuários no site, selecionando um subconjunto de publicações para aparecer no seu feed de notícias”, realizaram testes extensivos, mas não realizaram uma consulta aos usuários nem uma votação (Hempel, 2016).

Esta mudança de uma governança enraizada na participação comunitária para uma governança enraizada em burocracias de cima para baixo levou a uma crise de legitimidade para as plataformas online. À medida que uma plataforma amadurece, ela tenta cada vez mais equilibrar duas prioridades concorrentes – o que Jonathan Zittrain (2019) chama de “direitos” e “saúde pública”. Os direitos referem-se à capacidade dos usuários finais de participarem em espaços online sem interferências indevidas – na maioria das vezes, isto significa liberdade de expressão. A saúde pública pede às plataformas que: [pesem] os benefícios ou danos sistêmicos. . . e pensem sobre quais intervenções sistêmicas podem reduzir seus aparentes excessos” (Zittrain, 2019) – na maioria das vezes isso significa moderar o conteúdo. Quando o sistema de governança de uma plataforma segue o modelo de serviço ao cliente, as suas decisões em nome dos direitos ou da saúde pública são frequentemente examinadas e criticadas como corruptas, arbitrárias e irresponsáveis, em grande parte devido à não participação da comunidade.

O senador norte-americano Ted Cruz capta esse sentimento, referindo-se ao então CEO do Twitter, Jack Dorsey: Sr. Dorsey, quem diabos o elegeu e o encarregou do que a mídia pode reportar e do que o povo americano pode ouvir. . . ?” (CBS News, 2020).

Como devem as plataformas online responder a esta crise de legitimidade? Argumentamos que os processos participativos – um regresso aos modelos anteriores de governança comunitária – são o caminho necessário. Os processos participativos compreendem que todas as opiniões não serão e não podem ser reconciliadas. Mas, idealmente, um processo participativo deixa mesmo aqueles que perderam uma disputa ou debate específico em condição de aceitar uma decisão como legítima.

Os modelos de governança da Internet pré-Web deram aos usuários um sentido de propriedade e autoridade sobre as suas comunidades. Num sentido literal, os usuários muitas vezes ficavam impotentes se os administradores de um site fizessem uma grande mudança, uma vez que o poder de alterar as regras de um sistema estava nas mãos de uma elite técnica. Mas parte desse poder foi transferido para os usuários, quer através do estabelecimento de mecanismos de governança, quer através de métodos mais indiretos, e muitos usuários sentiram que os seus pontos de vista tinham influência sobre o funcionamento de uma comunidade.

Os usuários que se sentem proprietários dos espaços em que participam geralmente estão mais dispostos a apoiá-los e, às vezes, fazem de tudo para fazê-lo. O moderador do Reddit, Robert Peck (2019), explica que sua disposição de moderar comunidades (*subreddits*) como voluntário se baseia em sua percepção da comunidade e nas regras que permitem seu funcionamento:

Você não poderia me pagar para modificar *reddit.com*. Imagine esse trabalho: das 9h às 17h todos os dias atrás de uma tela, eliminando trolls, totalmente anônimo, mas mais vulnerável a cada hora para cada novo racista ou sexista que você banir. Não, insisto em fazer isso de graça.

As regras que Peck impõe em seus *subreddits* são aquelas que ele ajudou a criar. Assim como o LambdaMOO descobriu com seu processo de regulamentação por meio de petições, criar suas próprias regras lhes dá uma legitimidade que é difícil de contestar. Além disso, os juristas Tracey Meares e Tom Tyler (2021) encontraram evidências de que os processos considerados legítimos têm maior probabilidade de serem seguidos. Examinando o Facebook e o Twitter, eles entrevistaram usuários que foram suspensos por violarem os termos de serviço. Aqueles que consideraram que o processo subjacente à sua suspensão foi justo e legítimo tiveram menos probabilidades de reincidir do que aqueles que consideraram que o processo foi injusto ou arbitrário. Tyler, Meares e a sua equipe estão agora experimentando aumentar a percepção de justiça e legitimidade através de mecanismos de governança que enfatizam a participação comunitária.

Além dos benefícios que a participação da comunidade na governança traz para os espaços online, a participação na governança online também pode ajudar os envolvidos a evoluírem como cidadãos numa democracia. Robert Putnam argumentou que o afastamento dos americanos das instituições locais, dos clubes sociais às ligas de boliche e ao governo comunitário, provavelmente prejudicaria a capacidade deles de participarem na vida cívica (Putnam, 2000, p. 410): . . . associações e redes menos formais de envolvimento cívico incutem nos seus membros hábitos de cooperação e espírito público, bem como as competências práticas necessárias para

participar na vida pública.”

Servir numa posição de governança em algo tão cotidiano como uma liga de boliche pode servir de treino para uma participação cívica mais ampla. Os hábitos que aprendemos ao realizar reuniões produtivas, resolver divergências entre pessoas que trabalham num empreendimento comum, articular o nosso ponto de vista e procurar pontos comuns são, sem dúvida, as competências de que necessitamos para participar numa democracia funcional. É possível imaginar o desenvolvimento dessas competências tanto em espaços online como offline. Ao perder a governança das nossas comunidades online, podemos estar perdendo uma educação valiosa em matéria de cidadania democrática. (Zuckerman, 2022).

Trazendo a comunidade de volta à governança da plataforma

A chave para a próxima era da governança online não reside em estabelecer um conjunto de recursos e políticas que sejam corretas ou incorretas na visão de uma pessoa ou empresa, mas sim em saber se esse conjunto de recursos e políticas são legítimos devido à forma inclusiva e deliberativa . . . como foram assentados” (Zittrain, 2019). Os processos participativos oferecem a possibilidade de reconciliar as prioridades concorrentes das plataformas online de uma forma que deixa a maioria dos envolvidos com a sensação de que as decisões tomadas por uma plataforma são legítimas, mesmo que discordem de um resultado específico. Prevemos que tais processos assumam as características de transparência, descentralização e normas profissionais.

Transparência

Há uma ladainha de preocupações sobre a esfera pública digital. Preocupamo-nos que os espaços online possam prejudicar os indivíduos, minando a imagem corporal e a auto-estima das pessoas e empurrando as pessoas vulneráveis para ideologias extremistas. Preocupamo-nos que a Internet esteja aumentando a polarização política, prendendo-nos em câmaras de eco ideológicas e enganando-nos com informações falsas/desinformadas.

Quão válidas são essas preocupações? É complicado, e não apenas porque as ciências sociais são complicadas, mas porque é difícil para investigadores independentes estudarem o que ocorre nos espaços online. Sabemos significativamente mais sobre alguns espaços do que outros – o Twitter tornou-se a drosófila dos estudos de mídia social ao tornar o conteúdo público por padrão e dar aos acadêmicos acesso a APIs – mas um estudo recente de estudiosos de mídia social (Hansen Shapiro et al., 2021) descobriu que ninguém acreditava ter acesso aos dados necessários para responder a questões-chave sobre os efeitos das redes sociais nos indivíduos e na sociedade. Isso incluiu pesquisadores que foram fundamentais na concepção de colaborações acadêmicas e comerciais como o Social Science One.

A transparência permite que investigadores, decisores políticos e o público responsabilizem as plataformas online, compreendam melhor a esfera pública digital e identifiquem oportunidades de melhorias mais facilmente, ao mesmo tempo que legitimam potencialmente as reivindicações e práticas das plataformas. As plataformas deveriam divulgar mais dados sobre moderação de conteúdo, sobre como direcionam o tráfego na Internet e sobre a eficácia das técnicas utilizadas para reduzir a desordem informacional (DiResta et al., 2022). Devem também fornecer aos usuários mais informações e oportunidades de reparação quando são tomadas decisões a nível individual, tais como suspensões, eliminações e desclassificações de conteúdos. A legislação relacionada com a transparência, como a Lei de Responsabilidade e Transparência das Plataformas (Coons, Portman, Klobuchar...”, 2021) pode apoiar tais objetivos.

Descentralização

A descentralização do controle sobre os espaços online pode trazer legitimidade às plataformas, ao incluir mais vozes na construção e operação da esfera pública digital e ao oferecer mais opções aos participantes. Consideramos uma variedade de abordagens para a partilha de poder, não limitadas a abordagens técnicas como protocolos federados, mas incluindo também conselhos consultivos, processos democráticos e experimentação.

Conselhos Consultivos . Os conselhos consultivos que têm poder vinculativo relacionado com uma plataforma e independência suficiente da plataforma são uma forma de descentralizar o controle.

O Conselho de Supervisão da Meta é o exemplo mais proeminente. É composto por ex-líderes políticos, ativistas de direitos humanos e jornalistas de todo o mundo que influenciam as decisões de conteúdo da plataforma (Levy, 2022). As decisões do Conselho de Supervisão (*Appeals Process*, s.d.) de manter ou reverter as decisões de conteúdo da Meta são vinculativas, o que significa que a Meta deve implementá-las, a menos que isso viole a lei. O Conselho de Supervisão também pode fazer recomendações à Meta sobre as suas políticas e como as aplica. A Meta está empenhada em responder publicamente a essas recomendações no prazo de 60 dias (Meta, 2022). Elon Musk manifestou interesse em formar um conselho de moderação de conteúdo” semelhante para o Twitter (Frenkel, 2022).

Os conselhos consultivos permitem que as plataformas online incorporem as contribuições de partes interessadas independentes na sua governança. Idealmente, isto permite à plataforma melhorar a governança do seu espaço, legitimar algumas das suas decisões e reduzir a sua carga operacional. O sucesso dos conselhos consultivos depende provavelmente do fato de o sua contribuição ser vinculativa, do processo de seleção dos membros e da sustentabilidade financeira do conselho. Existem críticas razoáveis tanto ao conceito de um conselho consultivo como à implementação específica da Meta, mas uma verificação do poder absoluto das decisões de confiança e segurança de uma empresa é um passo sábio em direção à descentralização do poder.

Processos Democráticos . Os processos democráticos são uma forma das plataformas online incorporarem a contribuição da comunidade na sua tomada de decisões.

Aviv Ovadya (2021) discute a promessa de processos democráticos num artigo sobre democracia de plataforma”. Ovadya argumenta que os processos democráticos podem criar mandatos populares” independentes para espaços online, um resultado valioso para as populações afetadas, para os governos que são constitucionalmente incapazes de agir com base no discurso, e para as próprias plataformas.

Ovadya destaca o potencial de um conjunto de processos democráticos chamados assembleias de cidadãos” que normalmente envolvem a criação de um minipúblico” demograficamente representativo que é remunerado por um período de tempo fixo para aprender sobre uma questão, deliberar em conjunto e expressar as suas conclusões. Ovadya aponta exemplos bem sucedidos da sua utilização na política convencional em todo o mundo, incluindo na política de aborto na Irlanda e na política de energia nuclear na Coreia do Sul, e argumenta que poderiam ser aplicados com sucesso a questões difíceis de governança de plataformas.

Outra abordagem dos processos democráticos pode ser encontrada na iniciativa *Community Notes* do Twitter (Twitter, 2022). As Notas da Comunidade permitem que os colaboradores proponham uma nota adicionando contexto a qualquer tuite. Se uma nota for classificada como útil por um número suficiente de pessoas de diferentes perspectivas, a nota será exibida junto ao tuite. O projeto piloto desse recurso do Twitter descobriu que: (1) a maioria das pessoas achou as notas úteis, (2) as pessoas tinham 20% a 40% menos probabilidade de concordar com um tuite enganoso depois de ler uma nota sobre ele, e (3) a maioria das notas foram avaliadas com alta precisão por revisores profissionais. Elon Musk apoiou a iniciativa *Community Notes* e até teve os seus tuites sujeitos às notas (Leffer, 2022).

Os processos democráticos podem ser vantajosos para as comunidades e plataformas – as plataformas libertam-se de parte da responsabilidade pela tomada de decisões difíceis, enquanto a comunidade tem mais voz na governança das mesmas. Isto contribui para um sentimento de apropriação comum, ao mesmo tempo que melhora a governança e alivia a pressão sobre a plataforma. Além disso, é uma abordagem que foi testada com sucesso por uma das instituições mais proeminentes da Internet: a Wikipédia. A Wikipédia deixa claro que esses processos são viáveis em escala e bem-sucedidos na prática, fornecendo um modelo inspirador para muitas plataformas desenvolverem.

Federação Técnica . A federação técnica engloba diferentes abordagens que dependem de soluções tecnológicas para descentralizar o controle dos espaços online.

Uma forma de pensar sobre a federação técnica é que ela pretende fazer com que os espaços online funcionem mais como o e-mail (Rajendra-Nicolucci Zuckerman, 2020), com protocolos que facilitam a escolha e a independência. A federação técnica abrange um espectro, desde aplicativos com experiências muito diferentes de interoperabilidade (imagine que o YouTube e o Twitter possam se comunicar entre si) até middleware que permite aos usuários escolher o algoritmo de sua preferência (imagine que, em vez de usar o algoritmo de linha do tempo do Twitter, você poderia escolher entre um conjunto de algoritmos fornecidos por terceiros, como o New York Times ou a Fox News). O interesse por essas tecnologias aumentou nos últimos anos, com projetos construídos

em torno de tecnologia descentralizada como Mastodon (2023), Matrix (2023) e Bluesky (2023), regulamentações que exigem interoperabilidade como a Lei de Mercados Digitais (Lomas, 2022), propostas que enfatizam "middleware" como Gobo (Lane, 2022) e Fukuyama *et al.* (s.d.), e o movimento tecnocultural Web3" (Roose, 2022).

A federação técnica incorpora a federação na tecnologia de um espaço online, proporcionando fortes garantias sobre a descentralização do controle. No entanto, a federação técnica enfrenta uma série de desafios:

- primeiro, a federação técnica pode resultar em experiências de usuário complexas e confusas;
- em segundo lugar, existem questões em aberto sobre a sustentabilidade financeira dos protocolos da federação, muitos dos quais foram construídos como projetos de código aberto por voluntários;
- terceiro, sem mandatos regulatórios ou compromissos vinculativos com a federação técnica por parte dos principais espaços online, esta provavelmente enfrentará grandes barreiras à adoção em massa;
- finalmente, as redes federadas podem enfrentar os mesmos problemas de confiança e segurança que as redes centralizadas, com menos recursos para responder a esses desafios e problemas de coordenação mais complexos.

Experimentação. Uma cultura florescente de experimentação que incentiva a experimentação de diferentes configurações de recursos, normas, políticas e estruturas de governança pode levar a mais escolhas, agência e melhores práticas na esfera pública digital.

Apoiar a experimentação com espaços online significaria que mais pessoas que consideram os espaços existentes inadequados poderiam construir os seus próprios espaços. Casey Fiesler (2021) escreve sobre *Archive of Our Own* (AO3), um arquivo criado por fãs, administrado por fãs, sem fins lucrativos e não comercial para obras de fãs transformadoras" que atende mais de 5 milhões de pessoas. O AO3 foi criado em resposta ao tratamento inadequado dispensado aos autores de fanfic pelas plataformas existentes, incluindo desligamentos inesperados e grandes mudanças em recursos e políticas. Um grupo de fãs convocou um movimento para "possuir os servidores" e o AO3 nasceu. Se outras comunidades pudessem experimentar de forma semelhante os seus próprios espaços, isso aliviaria a pressão sobre os espaços existentes para serem tudo para todos e apoiaria um sentido de agência na esfera pública digital.

No entanto, é difícil personalizar, controlar e executar o software necessário para hospedar um espaço online independente. As soluções comerciais existentes são caras e limitam a capacidade de controlar e personalizar o software e os dados do espaço. O software de código aberto geralmente requer conhecimento técnico para configurar, gerenciar e personalizar. Isto significa que o grupo de pessoas capazes de experimentar espaços online está limitado àqueles que podem pagar pela experimentação técnica e àqueles com conhecimento para experimentar eles próprios. Para realmente permitir o florescimento de espaços online, precisamos de sistemas que permitam que pessoas com pouco conhecimento técnico e dinheiro criem seus próprios espaços personalizados e controláveis.

Reddit é um bom exemplo de cultura de experimentação dentro de uma plataforma importante. Os subreddits têm um controle significativo sobre seus espaços, incluindo a capacidade de definir seus próprios padrões comunitários. Isso significa que os subreddits podem experimentar seus espaços, às vezes em parceria com pesquisadores, para melhorá-los. Por exemplo, Nate Matias (2019) trabalha com subreddits para experimentar formas de tornar os seus espaços mais resistentes à desinformação e mais acolhedores para novos usuários. As descobertas dos experimentos de subreddits individuais informam as melhores práticas compartilhadas que melhoram a plataforma como um todo. A cultura de experimentação do Reddit também significa que espaços com objetivos, normas, políticas e processos muito diferentes podem coexistir. Por exemplo, *r/changemyview*, que tem diretrizes detalhadas para participação que enfatizam a civilidade e a mente aberta e se dedica a ser (Reddit, 2022a) [um] lugar para postar uma opinião que você aceita que pode ser falha, em um esforço para compreender outras perspectivas sobre a questão", pode coexistir com *r/wallstreetbets*, um espaço indisciplinado, autodescrito

(Reddit, 2022b) como o 4chan encontrou um terminal Bloomberg”, dedicado a discussões sobre mercados de ações e que alimentou partes da mania de estoque de memes”.

Nosso próprio trabalho está em pequenas redes sociais para fins cívicos. Construimos uma rede social – Smalltown – que as comunidades podem utilizar para manter conversas cuidadosamente moderadas sobre questões cívicas locais, aumentando o número de pessoas que podem participar. Referimo-nos carinhosamente a ela como a rede social mais chata do mundo, porque as conversas são respeitadas, civilizadas e sobre temas quentes como estacionamento paralelo versus estacionamento de ré em ângulo. Também estamos nos estágios iniciais da construção de uma rede de descoberta musical – Freq – que está tentando criar um espaço fora da órbita do Spotify que coloque o controle de volta nas mãos dos fãs de música. Acreditamos que espaços online construídos com um propósito específico e valores específicos podem nos ajudar a sair da sombra das grandes plataformas. Estes espaços independentes oferecem diferentes formas de reunião online que não funcionam na mesma lógica das grandes plataformas, contribuindo para uma esfera pública digital mais diversificada e fértil.

As alternativas públicas oferecem muitos dos mesmos benefícios. A tradição dos meios de comunicação públicos mostra que as alternativas públicas podem oferecer uma forma de moldar um ecossistema e corrigir falhas de mercado. Um exemplo inicial promissor é o *Public Spaces* (2022), uma coligação de emissoras públicas e instituições culturais nos Países Baixos que tentam diminuir a dependência de software de vigilância das suas instituições e com esperança de construir ferramentas que sejam mais consistentes com os seus valores. Desenvolvedores intimamente ligados aos Espaços Públicos estão lançando uma nova plataforma de mídia social chamada *PubHubs* (2022), projetada para uso por pequenas comunidades geográficas, como sua família, clube esportivo, turma escolar, museu, biblioteca local, bairro ou município”. A plataforma tira partido de um novo sistema de identidade baseado em atributos”, IRMA (*IRMA Technical Documentation*, 2022), que permite às pessoas permanecerem anônimas enquanto apresentam provas de que são membros de um determinado município ou têm filhos numa determinada escola.

Normas Profissionais

Freqüentemente, profissões que têm muito poder desenvolvem normas profissionais de respeito ao público como forma de gerar confiança na sociedade. Os exemplos canônicos são médicos e advogados, mas jornalistas, engenheiros civis e agentes policiais também desenvolveram normas profissionais relativas ao público (Gardner et al., 2001). Os profissionais, especialmente os engenheiros, que trabalham em espaços online devem igualmente adotar um conjunto de normas profissionais relativas ao público. Como diz Jack Balkin (2020):

As empresas de mídia social. . . tornaram-se instituições-chave para promover uma esfera pública saudável. Não podem servir apenas como incentivos econômicos. . . Por analogia, pense no jornalismo. Ele também desempenha um papel crucial na esfera pública. . . Se as normas profissionais do jornalismo forem enfraquecidas ou destruídas e a prática do jornalismo se tornar exclusivamente orientada ao mercado, o jornalismo tornará a esfera pública pior, e não melhor. Escolherá histórias e tratamentos que aumentem a polarização, o tribalismo e a desconfiança social, e gerará ou ajudará a espalhar propaganda e teorias da conspiração.

Existem alguns sinais encorajadores de construção de normas na organização trabalhista em formação (Tech Workers Coalition, s.d.), esforços crescentes para incutir um senso de ética em jovens engenheiros de software (Karoff, 2019) e exemplos inspiradores de responsabilidade pública em ação (Mac Kang, 2021). No entanto, é necessário um esforço mais concertado no desenvolvimento das organizações, da educação e dos padrões necessários para a adoção generalizada de normas profissionais de interesse público. A Associação de Profissionais de Confiança e Segurança é um exemplo promissor de como isso poderia ser na prática. Ela atende a comunidade global de profissionais que desenvolvem e aplicam princípios e políticas que definem comportamento e conteúdo on-line aceitáveis” (Trust Safety Professional Association [TSPA], 2022). Proporciona aos seus membros desenvolvimento de carreira, atividades em rede, formação e partilha de conhecimentos, e trabalha para melhorar a compreensão da sociedade no domínio da confiança e segurança (TSPA, 2022).

Embora o surgimento de um forte sentido de identidade profissional em torno da confiança e da segurança seja inquestionavelmente um desenvolvimento positivo, ele traz à tona a tensão central que exploramos aqui: a governança de uma comunidade de usuários online é responsabilidade desses usuários, ou de profissionais dedicados à segurança dos espaços online? Esperamos que as normas profissionais que emergem mantenham o devido respeito pela vontade dos indivíduos e grupos de usuários, e uma vontade de permitir que esses usuários se autogovernem sempre que possível.

Limitações

Não acreditamos que os processos participativos sejam uma panaceia para a esfera pública digital. Existem muitos desafios que os processos participativos não abordam. Nossa análise e proposta estão focadas em enfrentar a crise de legitimidade que as plataformas enfrentam.

Estamos cientes de que a eficácia dos processos participativos pode variar entre países. Por exemplo, se um país regula rigorosamente quem pode construir redes sociais e o que é mostrado nelas, a experimentação e os processos democráticos têm muito menos probabilidades de serem bem sucedidos, ou mesmo possíveis. Cada um dos tipos de processos participativos que identificamos – transparência, descentralização e normas profissionais – depende do ambiente político. No entanto, especialmente para grandes plataformas utilizadas em vários países, os processos participativos podem ser vistos com bons olhos mesmo por nações que são hostis à democracia. Uma queixa comum é que as grandes plataformas não têm em conta as leis e a cultura locais – os países podem acolher favoravelmente processos dedicados para que os cidadãos e representantes expressem as suas opiniões, bem como soluções que descentalizem o controle. Poderíamos imaginar um país com leis muito mais rígidas sobre nudez, acolhendo a oportunidade de usar APIs disponíveis publicamente para implementar seus próprios filtros de conteúdo. Trabalhos futuros poderiam explorar essas questões com mais detalhes.

Além disso, reconhecemos que os processos participativos podem obscurecer certos pontos de vista e refletir as desigualdades existentes. Os implementadores devem ter o cuidado de conceber os seus processos de forma inclusiva, para que não acabem simplesmente por refletir as opiniões daqueles que têm mais tempo ou dinheiro disponíveis. Trabalhos futuros poderiam explorar como é o design inclusivo na prática.

Conclusão

A importância das plataformas online seguirá aumentando à medida que a esfera pública digital se tornar mais central para a esfera pública contemporânea. No entanto, as plataformas enfrentam uma crise de legitimidade, resultante das suas tentativas de equilibrar as prioridades concorrentes dos direitos individuais e de saúde pública, utilizando um modelo de governança online enraizado em burocracias de cima para baixo, a que chamamos modelo de serviço ao cliente”. Anteriormente, as formas pré-Web de governança online que enfatizavam a participação da comunidade ofereciam um modelo promissor para superar esta crise, combinando as prioridades concorrentes das plataformas de uma forma que deixa a maioria das partes interessadas com a sensação de que as decisões tomadas por uma plataforma são legítimas, mesmo que discordem com um determinado resultado. Estes modelos podem oferecer o benefício adicional de preparar os usuários para formas mais amplas de participação cívica.

Para indivíduos que procuram medidas práticas que possam tomar para apoiar a governança comunitária online, sugerimos:

- voluntariado como moderador ou administrador de um espaço online;
- contribuir para iniciativas de governança comunitária, como as Notas Comunitárias;
- pressionar as plataformas para adotarem processos participativos;
- experimentar novas formas de comunidade online, como a Fediverse;
- contatar os decisores políticos sobre legislação que apoia a transparência e a descentralização online.

--

Declaração de interesses conflitantes

O(s) autor(es) não declararam possíveis conflitos de interesse com relação à pesquisa, autoria e/ou publicação deste artigo.

Financiamento

O(s) autor(es) declararam ter recebido o seguinte apoio financeiro para a pesquisa, autoria e/ou publicação deste artigo: Fundação Ford, Fundação John D. e Catherine T. MacArthur.

ID ORCID

Chand Rajendra-Nicolucci -- <https://orcid.org/0009-0009-9717-8393>

Ethan Zuckerman e Chand Rajendra-Nicolucci

Social Media + Society, julho a setembro de 2023: 1 –12

©O(s) Autor(es) 2023

Diretrizes para reuso de artigos: sagepub.com/journals-permissions

<https://doi.org/10.1177/20563051231196864>

DOI: 10.1177/20563051231196864

journals.sagepub.com/home/sms

Referências

Balkin, J. (2020). How to regulate (and not regulate) social media. Knight First Amendment Institute at Columbia University. <https://knightcolumbia.org/content/how-to-regulate-and-not-regulate-social-media>

Bluesky. (2023). <https://blueskyweb.xyz>

Caplan, R. (2018). Content or context moderation? Artisanal, community-reliant, and industrial approaches. Data Society. <https://datasociety.net/library/content-or-context-moderation/>

CBS News. (2020, outubro 28). Twitter, Facebook and Google CEOs testify before Senate Commerce Committee [Video]. YouTube. <https://www.youtube.com/watch?v=2CO3krTpJ04>

Coons, Portman, Klobuchar announce legislation to ensure transparency at social media platforms. (2021, December 9). <https://www.coons.senate.gov/news/press-releases/coons-portman-klobuchar...>

Data Society Podcast. (2020, julho 8). Origins of trust and safety. <https://datasociety.net/library/origins-of-trust-and-safety/>

DeVoto, J. A. E. (1994). The Zumabot's Tale. <http://www.jaedworks.com/shoebox/zumabot.html>

Dibbell, J. (1993, dezembro 23). A rape in cyberspace. The Village Voice. http://www.juliandibbell.com/texts/bungle_vv.html

DiResta, R., Edelson, L., Nyhan, B., Zuckerman, E. (2022, abril 28). It's time to open the Black Box of social media. Scientific American. <https://www.scientificamerican.com/article/its-time-to-open-the-black-bo...>

Driscoll, K. (2022a, maio 17). The Internet origin story you know is wrong. Wired. <https://www.wired.com/story/Internet-origin-story-bbs/>

Driscoll, K. (2022b). The modem world: A prehistory of social media. Yale University Press.

Fiesler, C. (2021). Gift logic: Labors of love flourish online under fandom's social norms. Knight First Amendment Institute. <https://knightcolumbia.org/blog/gift-logic-labors-of-love-flourish-onlin...>

- Frenkel, S. (2022, outubro 28). Elon Musk moves to form a content moderation council for Twitter. The New York Times. <https://www.nytimes.com/2022/10/28/technology/twitter-elon-musk-content-...>
- Fukuyama, F., Richman, B., Goel, A., Katz, R., Melamed, A., Schaake, M. (s.d.). Middleware for dominant digital platforms: A technological solution to a threat to democracy. https://fsi-live.s3.us-west-1.amazonaws.com/s3fs-public/cpc-middleware_f...
- Gardner, H., Csikszentmihalyi, M., Damon, W. (2001). Good work: When excellence and ethics meet. Basic Books.
- Gillespie, T. (2018). Custodians of the Internet: Platforms, content moderation, and the hidden decisions that shape social media. Yale University Press.
- Gray, M. L., Suri, S. (2019). Ghost work: How to stop Silicon Valley from building a new global underclass. Harper Business.
- Hansen Shapiro, E., Sugarman, M., Bermejo, F., Zuckerman, E. (2021). New approaches to platform data research. <https://www.netgainpartnership.org/resources/2021/2/25/new-approaches-to...>
- Hempel, J. (2016, setembro 6). Everyone hated news feed. Then it became Facebook's most important product. Wired. <https://www.wired.com/2016/09/everyone-hated-news-feed-then-it-became-fa...>
- IRMA Technical Documentation. (2022). <https://irma.app/docs/v0.2.0/overview/>
- Karoff, P. (2019, janeiro 25). Embedding ethics in computer science curriculum. Harvard Gazette. <https://news.harvard.edu/gazette/story/2019/01/harvard-works-to-embed-et...>
- Kelly, K., Rheingold, H. (1993, março 1). The dragon ate my homework. Wired. <https://www.wired.com/1993/03/muds-2/>
- Killfiles—The cure for all that ails you. (1994). https://www.whitman.edu/mathematics/eegtti/eeg_79.html
- Koebler, J. (2015). It's September, forever. VICE. <https://www.vice.com/en/article/nze8nb/its-september-forever>
- Lane, S. (2022, novembro 9). Gobo 2.0: All your social media in one place. Initiative for Digital Public Infrastructure. <https://publicinfrastructure.org/2022/11/09/gobo-2-0-all-your-social-med...>
- Leffer, L. (2022, novembro 28). Elon Musk praises improved community notes" as community flags his fake news post. Gizmodo. <https://gizmodo.com/elon-musk-twitter-community-notes-fake-news-1849827942>
- Levy, S. (2022, novembro 8). Inside Meta's oversight board: 2 years of pushing limits. Wired. <https://www.wired.com/story/inside-metas-oversight-board-two-years-of-pu...>
- Lomas, N. (2022, março 24). Europe says yes to messaging interoperability as it agrees on major new regime for Big Tech. TechCrunch. <https://techcrunch.com/2022/03/24/dma-political-agreement/>
- Mac, R. Kang, C. (2021, outubro 3). Whistle-Blower Says Facebook 'Chooses Profits Over Safety'. The New York Times. <https://www.nytimes.com/2021/10/03/technology/whistle-blower-facebook-fr...>
- MacKinnon, R. (2012). Consent of the networked: The worldwide struggle for Internet freedom. Basic Books.
- Mastodon. (2023). <https://joinmastodon.org>
- Matias, J. N. (2019). Preventing harassment and increasing group participation through social norms in 2,190 online science discussions. Proceedings of the National Academy of Sciences, 116(20), 9785–9789. <https://doi.org/10.1073/pnas.1813486116>
- Matrix. (2023). <https://matrix.org>
- Maxim, K., Parecki, J., Cornett, C. (2022). How to build a trust and safety team in a year: A practical guide from

- lessons learned (so far) at zoom. *Journal of Online Trust and Safety*, 1(4). <https://doi.org/10.54501/jots.v1i4.81>
- Merriam-Webster. (n.d.). Netiquette. In Merriam-Webster. <https://www.merriam-webster.com/dictionary/netiquette>
- Meta. (2022, dezembro 2). Oversight board recommendations. <https://transparency.fb.com/oversight/oversight-board-recommendations/>
- Miller, C. (2021, janeiro 12). Usenet spam: A slice of history. https://fishbowl.pastiche.org/2021/01/12/usenet_spam
- Mnookin, J. L. (1996). Virtual(ly) law: The emergence of law in LambdaMoo. *Journal of Computer-Mediated Communication*, 21(1), JCMC214. <https://doi.org/10.1111/j.1083-6101.1996.tb00185.x>
- Ovadya, A. (2021, outubro 18). Towards platform democracy: Policymaking beyond corporate CEOs and partisan pressure. <https://www.belfercenter.org/publication/towards-platform-democracy-poli...>
- Oversight Board. (s.d.). Appeals process. <https://www.oversight-board.com/appeals-process/> [unreachable]
- Paolillo, J. C., Heald, D. (2002). Democratic participation in the discursive management of Usenet. In *Proceedings of the 35th Annual Hawaii International Conference on System Sciences* (pp. 1040–1049). IEEE. <https://doi.org/10.1109/HICSS.2002.994060>
- Peck, R. (2019, março 21). The punishing ecstasy of being a Reddit moderator. *Wired*. <https://www.wired.com/story/the-punishing-ecstasy-of-being-a-reddit-mode...>
- Porter, J. (2021, fevereiro 9). Reddit to double employees after raising \$ million250. *The Verge*. <https://www.theverge.com/2021/2/9/22274077/reddit-funding-round-250-mill...>
- PubHubs. (2022). <https://pubhubs.net/en/>
- Public Spaces. (2022). <https://publicspaces.net/>
- Putnam, R. D. (2000). *Bowling alone: The collapse and revival of American community*. Simon Schuster.
- Rajendra-Nicolucci, C., Zuckerman, E. (2020). What if social media worked more like email? Knight First Amendment Institute. <https://knightcolumbia.org/blog/what-if-social-media-worked-more-like-email>
- Rheingold, H. (1993). *The virtual community*. Addison Wesley.
- Reddit. (2022a). /r/changemyview. <https://www.reddit.com/r/change-myview/> [unreachable]
- Reddit. (2022b). /r/wallstreetbets. <https://www.reddit.com/r/wall-streetbets/> [unreachable]
- RFC 1192. (1990). <https://www.rfc-editor.org/rfc/rfc1192>
- Roose, K. (2022). What is web3? *The New York Times*. <https://www.nytimes.com/interactive/2022/03/18/technology/web3-definitio...>
- Roser, M., Ritchie, H., Ortiz-Ospina, E. (2015). *Internet*. <https://ourworldindata.org/Internet>
- Smith, E. (2020, outubro 13). No more eternal Septembers. *Tedium*. <https://tedium.co/2020/10/13/eternal-september-modern-impact/>
- Smith, M. (1999). Invisible crowds in cyberspace: Measuring and mapping the social structure of USENET. In M. Smith P. Kollock (Eds.), *Communities in cyberspace: Perspectives on new forms of social organization* (pp. 195–219). Routledge Press.
- Steiner, P. (1993). On the Internet nobody knows you're a dog [Cartoon]. *The New Yorker*. https://en.wikipedia.org/wiki/On_the_Internet_nobody_knows_you%27re_a_d...

- Sterling, B. (1993). A short history of the Internet. <http://textfiles.com/Internet/sterling>
- Suler, J. (2004). The online disinhibition effect. *Cyberpsychology Behavior*, 7(3), 321–326.
- Tech Workers Coalition. (s.d.). <https://techworkerscoalition.org/>
- Tempest, R. (1989, outubro 24). Minitel: Miracle or monster? *Los Angeles Times*. <https://www.latimes.com/archives/la-xpm-1989-10-24-mn-718-story.html>
- Trust Safety Professional Association. (2022). <https://www.tspa.org/>
- Turkle, S. (1995). *Life on the screen: Identity in the age of the Internet*. Simon Schuster.
- Twitter. (2022). Community notes. <https://twitter.github.io/communitynotes/>
- Van Vleck, T. (2012). Electronic mail and text messaging in CTSS, 1965-1973. <https://multicians.org/thvv/anhc-34-1-anec.html>
- Venkatesh, S., Tyler, T., Meares, T., Badiei, F. (2021). In a new light: Social media governance reconsidered. *Yale Journal of Law Technology*, 23. <https://yjolt.org/new-light-social-media-governance-reconsidered>
- When studying MUDs, a form of Role playing game on the Internet, it is clear that there is a society. (1998). <https://mud.co.uk/dvw/mudsasasociety.html>
- Zittrain, J. (1997). The rise and fall of sysopdom. *Harvard Journal of Law Technology*, 10, 495. <https://cyber.harvard.edu/10hjolt495.html>
- Zittrain, J. (2019). Three eras of digital governance. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3458435
- Zuckerman, E. (2022). How social media could teach us to be better citizens. *Journal of e-Learning and Knowledge Society*, 18(3), 36–41.

Biografias dos autores

Ethan Zuckerman é professor de políticas públicas, comunicação e informação na Universidade de Massachusetts Amherst. Seus interesses de pesquisa incluem infraestrutura pública digital, mídia cívica, estudos quantitativos de mídia e governança online. Chand Rajendra-Nicolucci é pesquisador da Universidade de Massachusetts Amherst. Seus interesses de pesquisa incluem infraestrutura pública digital, governança online e design de mídias sociais.

Creative Commons CC BY: Este artigo é distribuído sob os termos da licença Creative Commons Attribution 4.0 (<https://creativecommons.org/licenses/by/4.0/>) que permite qualquer uso, reprodução e distribuição do trabalho sem permissão adicional fornecida o trabalho original é atribuído conforme especificado nas páginas SAGE e Open Access (<https://us.sagepub.com/en-us/nam/open-access-at-sage>).

1 MOO significa "MUD orientado a objetos" – a linguagem de programação no LambdaMOO era orientada a objetos, em oposição às linguagens funcionais usadas nos MUDs anteriores. O fato de o paradigma da linguagem de programação LambdaMOO ter se tornado um elemento central de seu nome nos dá uma ideia de quão *geeks* eram os usuários de tal sistema.

2 O maior concorrente do Tripod, o Geocities, é mais conhecido, embora os dois tivessem modelos de negócios semelhantes.

3 A Seção 230 da Lei de Decência Digital dos EUA determina que "nenhum fornecedor ou usuário de um serviço informático interativo será tratado como editor ou autor de qualquer informação fornecida por outro fornecedor de conteúdos de informação." (47 USC § 230(c)(1)).

[4](#) A Lei de Direitos Autorais do Milênio Digital (DMCA) é uma lei de direitos autorais dos Estados Unidos de 1998 que implementa dois tratados de 1996 da Organização Mundial da Propriedade Intelectual (OMPI). Criminaliza a produção e disseminação de tecnologia, dispositivos ou serviços destinados a contornar medidas que controlam o acesso a obras protegidas por direitos de autor (vulgarmente conhecidas como gestão de direitos digitais ou DRM).

Categoria:

- [poliTICs 37](#)
-