Ano VII, v.1 2026 | submissão: 29/10/2026 | aceito: 05/11/2026 | publicação: 10/11/2026

A expressão como reorganização de uma condição comum: liberdade, restrição e o critério da participação

Expression as a reorganization of a common condition: freedom, restriction, and the criterion of participation

La expresión como reorganización de una condición común: libertad, restricción y el criterio de la participación

Pablo Dietrich — Mestrando em Direito pela Universidade Estácio de Sá (UNESA). Advogado (OAB/SC 41.212). ORCID: https://orcid.org/0009-0003-5447-0550. E-mail: pablodietrichh@hotmail.com

Resumo: O debate sobre os limites da liberdade de expressão costuma percorrer um de dois atalhos: o que passa da constatação de que uma fala alterou algo para a autorização de contê-la, e o que passa da proteção da expressão para a conclusão de que nada se deve a quem ela alcança. Os dois são recusados a partir de um único critério, o da participação: a expressão que reorganiza condições de que outros vivem exige justificação perante eles; a intervenção que restringe essa expressão também reorganiza condições de participação e responde a exigências próprias de justificação, competência e controle. O critério é desdobrado em três perguntas — a do diagnóstico, a da justificação e a da resposta e da competência — que separam a exigência de resposta da autorização de restringir e submetem o restritor ao mesmo exame que o falante. A partir de três trabalhos anteriores sobre recepção, mito político e autorização da ação, mostra-se o que a reorganização sozinha não autoriza, como as respostas possíveis se comparam e quais exigências cada intervenção deve satisfazer. A parte final aplica o critério ao direito brasileiro vigente, registrando o que ele já assegura e o que permanece em aberto, em especial no regime do art. 19 do Marco Civil da Internet após a tese final do Supremo Tribunal Federal nos Temas 987 e 533. Uma comparação com o Regulamento dos Serviços Digitais da União Europeia e com o Online Safety Act britânico mostra que as lacunas do regime brasileiro — razões para a redução de alcance e para a moderação por condições de uso, e uma instância de revisão independente do provedor — têm correspondentes normativos em outros ordenamentos, que ampliam as alternativas institucionais em discussão.
Palavras-chave: liberdade de expressão; participação; justificação; moderação de conteúdo; direito comparado; Marco Civil da Internet.

Abstract: Debates on the limits of free speech tend to follow one of two shortcuts: from the finding that speech has altered something to the authorization to contain it, or from the protection of expression to the conclusion that nothing is owed to those it reaches. This article refuses both on the basis of a single criterion, that of participation: expression that reorganizes conditions others live from calls for justification before them; the intervention that restricts such expression also reorganizes conditions of participation and must meet its own requirements of justification, competence, and control. The criterion is unfolded into three questions — of diagnosis, of justification, and of response and competence — which separate the duty to answer from the authority to restrict and subject the restrictor to the same examination as the speaker. Drawing on three earlier studies on reception, political myth, and the authorization of action, the article shows what reorganization alone does not authorize, how the available responses compare, and what each intervention must satisfy. The final part applies the criterion to Brazilian law, recording what it already secures and what remains open, especially in the regime of Article 19 of the Marco Civil da Internet after the Supreme Federal Court's final ruling in Themes 987 and 533. A comparison with the European Union's Digital Services Act and the United Kingdom's Online Safety Act shows that the gaps in the Brazilian regime — reasons for reductions in reach and for moderation under terms of service, and a review body

Este é um artigo publicado em acesso aberto (Open Access) sob a licença CreativeCommons Attribution, que permite uso, distribuição e reprodução em qualquer meio, sem restrições desde que o trabalho original seja corretamente citado.

Ano VII, v.1 2026 | submissão: 29/10/2026 | aceito: 05/11/2026 | publicação: 10/11/2026

independent of the provider — have normative counterparts elsewhere, which widen the institutional alternatives under discussion.
Keywords: freedom of expression; participation; justification; content moderation; comparative law; Brazilian Internet Civil Framework.

Resumen: El debate sobre los límites de la libertad de expresión suele seguir uno de dos atajos: el que pasa de constatar que un discurso alteró algo a autorizar su contención, y el que pasa de la protección de la expresión a la conclusión de que nada se debe a quienes alcanza. Este trabajo rechaza ambos a partir de un único criterio, el de la participación: la expresión que reorganiza condiciones de las que otros viven exige justificación ante ellos; la intervención que restringe esa expresión también reorganiza condiciones de participación y responde a exigencias propias de justificación, competencia y control. El criterio se despliega en tres preguntas — la del diagnóstico, la de la justificación y la de la respuesta y la competencia — que separan la exigencia de respuesta de la autorización para restringir y someten al restrictor al mismo examen que al hablante. La parte final aplica el criterio al derecho brasileño vigente, registrando lo que ya asegura y lo que sigue abierto, en especial en el régimen del art. 19 del Marco Civil de Internet tras la tesis final del Supremo Tribunal Federal en los Temas 987 y 533. Una comparación con el Reglamento de Servicios Digitales de la Unión Europea y con la Online Safety Act británica muestra que las lagunas del régimen brasileño — razones para la reducción de alcance y para la moderación por condiciones de uso, y una instancia de revisión independiente del proveedor — tienen correspondientes normativos en otros ordenamientos, que amplían las alternativas institucionales en discusión.
Palabras clave: libertad de expresión; participación; justificación; moderación de contenidos; derecho comparado; Marco Civil de Internet.

1. Introdução

Dois movimentos se repetem no debate sobre os limites da liberdade de expressão. Um apresenta um dano e conclui pela contenção; o outro apresenta a proteção constitucional e conclui que nada é devido a quem a fala alcança. Nenhum dos dois pertence a um só campo político, e cada um aparece, conforme a ocasião, em ambos. Os dois movimentos são atalhos, e são simétricos. O primeiro salta da constatação de que uma fala alterou algo para a autorização de contê-la, como se identificar um dano fosse o mesmo que demonstrar quem pode responder a ele e como. O segundo salta da proteção da fala para a dispensa de qualquer justificação, como se proteger um ato fosse o mesmo que declará-lo sem destinatários. Os dois atalhos atravessam os regimes que, na última década, redesenharam a responsabilidade das plataformas — o Regulamento europeu dos Serviços Digitais, o Online Safety Act britânico, a tese do Supremo Tribunal Federal brasileiro e o decreto que a regulamenta —, e em cada um deles têm praticantes que leem o outro como adversário político. O que segue não escolhe um lado dessa disputa; dirige às duas posições as mesmas perguntas, e a simetria entre elas deve aparecer na aplicação, não na declaração.

Este é um artigo publicado em acesso aberto (Open Access) sob a licença CreativeCommons Attribution, que permite uso, distribuição e reprodução em qualquer meio, sem restrições desde que o trabalho original seja corretamente citado.

Ano VII, v.1 2026 | submissão: 29/10/2026 | aceito: 05/11/2026 | publicação: 10/11/2026

O instrumento é um critério formulado em outro lugar e já aplicado à ação em geral: o da participação. Uma ação que reorganiza condições de que outros vivem tem essas pessoas dentro de si pela posição que elas ocupam, e por isso deve resposta a elas (DIETRICH, 2026a). Aplicado à fala, o critério produz um resultado que nenhum dos dois atalhos aceita: a expressão que reorganiza condições de participação exige justificação perante quem delas vive, mas essa exigência não se converte, por si, em autorização para restringir; e a intervenção que restringe a expressão é, ela própria, uma reorganização de condições de participação, sujeita às mesmas perguntas que dirige ao falante.

Três trabalhos anteriores, sobre a recepção nas redes, o mito político e a autorização da ação, prepararam o terreno sem percorrê-lo (DIETRICH, 2026b; 2026c; 2026d); nenhum derivou do dano uma autorização de restringir, e ela tampouco será derivada deles aqui. O que se acrescenta é a distinção que os três não fizeram: entre a resposta que uma fala deve e a medida que alguém pode impor-lhe, há três passos — o diagnóstico, a justificação e a escolha da resposta pela competência de quem a adota — e cada um exige argumento próprio. O método é o da reconstrução conceitual submetida a teste: o critério é formulado, confrontado com as objeções mais fortes e aplicado a dois casos documentados, a um caso hipotético com variações, ao direito brasileiro vigente, lido no texto das normas e da tese judicial, e, por comparação, aos regimes europeu e britânico. O Brasil é o caso principal porque nele o redesenho foi o mais recente e o mais abrupto, feito por um tribunal e completado por decreto.

O terreno a percorrer vai, assim, do que já está estabelecido ao que ainda não está: dos três resultados anteriores ao critério e suas perguntas; do que a reorganização sozinha não autoriza à comparação entre as respostas e às exigências de cada intervenção; da intervenção como reorganização ao direito brasileiro vigente — à Constituição, aos tipos penais de honra e incitação, à Lei 14.197/2021 e à tese final do Supremo Tribunal Federal nos Temas 987 e 533 — registrando o que esse direito já assegura e o que permanece em aberto —, e, por fim, à comparação com os regimes europeu e britânico.

2. Três resultados anteriores

Os três trabalhos que precedem este investigam dimensões de um mesmo fenômeno — como se recebe, o que circula, o que se autoriza — e cada um fundamenta uma peça distinta do argumento. Tomá-los em bloco apagaria a diferença e emprestaria ao conjunto uma força que nenhum deles tem isoladamente.

Este é um artigo publicado em acesso aberto (Open Access) sob a licença CreativeCommons Attribution, que permite uso, distribuição e reprodução em qualquer meio, sem restrições desde que o trabalho original seja corretamente citado.

Ano VII, v.1 2026 | submissão: 29/10/2026 | aceito: 05/11/2026 | publicação: 10/11/2026

O primeiro descreveu a recepção. Quando o sucesso de alguém nas redes é tomado como aptidão para um cargo, não é apenas o candidato que demonstra competência de um modo novo; é o eleitor que passa a acreditar possuir competência para julgá-lo. A comunicação eficaz é tomada como esclarecimento quando confirma antecipações, oferece uma resposta imediatamente reproduzível e produz a sensação de domínio, sem exame suficiente (DIETRICH, 2026b). Esse resultado é hermenêutico, não lesivo. Ninguém é tratado como não participante por receber uma mensagem desse modo, e nenhuma posição é obstruída. O que o trabalho fundamenta é que a fala pode formar no receptor um juízo sem exame; o que ele não fundamenta é qualquer resposta além da que se dirige ao próprio exame. Suas hipóteses, ademais, descrevem um mecanismo e não valem como evidência empírica de que ele opere em um caso determinado.

O segundo tratou do conteúdo. O mito político, no sentido que Barthes, Edelman, Flood e Bottici permitem reconstruir, integra acontecimentos diferentes a uma narrativa de sentido político compartilhado, e o rótulo que dele deriva substitui o exame pelo pertencimento (DIETRICH, 2026c). Mas a conclusão daquele trabalho foi, deliberadamente, contrária à restrição: a interpretação mais ofensiva de uma frase ambígua não basta, por si só, para fechar a praça. O Tribunal Constitucional alemão, ao decidir sobre a frase “soldados são assassinos”, não certificou que a frase era verdadeira nem que era justa; decidiu que a praça em que ela podia ser examinada devia permanecer aberta (BVerfGE 93, 266). A proteção jurídica da fala não certifica seus juízos, e a degradação do debate pelo mito não retira de ninguém a condição de participante. O que o segundo trabalho fundamenta é um resultado negativo: há uma classe de falas que degradam a praça sem a fechar, e a resposta devida a elas é o exame.

O terceiro tratou da autorização. Um entendimento formado sob influência pode converter uma resposta que o agente tinha por proibida em uma resposta que ele passa a ter por exigida, e o agente pode fazer o mal sem o reconhecer como tal. Os casos de Southport, do Capitólio e de Brasília mostraram o ciclo — influência, compreensão, ação — e mostraram onde ele pode ser interrompido com menor dano: na correção pública tempestiva por quem detém a informação, na responsabilidade concentrada nos poucos que formam o entendimento de muitos, e em uma medida estreitamente delimitada de redução de alcance para conteúdo já verificado falso durante uma crise declarada (DIETRICH, 2026d). Aquele trabalho reconheceu a objeção de que a evidência de que a desinformação transforma cidadãos em arruaceiros é fraca e de que um poder de demover conteúdo em crise será usado fora dela (REILLY, 2024), e confinou sua proposta a limites que a objeção não alcançava. O que ele fundamenta é que

Este é um artigo publicado em acesso aberto (Open Access) sob a licença CreativeCommons Attribution, que permite uso, distribuição e reprodução em qualquer meio, sem restrições desde que o trabalho original seja corretamente citado.

Ano VII, v.1 2026 | submissão: 29/10/2026 | aceito: 05/11/2026 | publicação: 10/11/2026

uma fala pode reorganizar a garantia pública de que um grupo não será atacado pelo que é (WALDRON, 2012) e que há respostas a isso que não entram no ciclo. O que ele não fundamenta é que os limites daquela intervenção sejam condições de toda resposta legítima à expressão: uma incitação pode usar informação verdadeira, uma ameaça pode ocorrer fora de crise, e exigir crise declarada entrega à autoridade que a declara parte do poder de definir quando a proteção diminui.

A ordem em que os três trabalhos foram escritos sugere uma sequência: a recepção sem exame tornaria o rótulo eficaz, e o rótulo tornaria a autorização possível. A sequência é útil para organizar a análise e enganosa se for lida como cadeia causal. Nenhum dos três demonstrou que o primeiro momento produz o segundo ou que o segundo produz o terceiro; cada um descreveu um mecanismo e reuniu evidência de circulação — visualizações, repetições, convocações — que não é evidência de efeito. Quatrocentas e vinte mil visualizações de um nome falso e dois hotéis atacados nos dias seguintes são fatos do mesmo episódio, não uma relação causal demonstrada entre uma publicação e um ataque. O que segue mantém, por isso, três coisas separadas: a interpretação do mecanismo, a evidência de circulação e a demonstração de efeitos. Só a última autorizaria afirmações causais sobre uma publicação específica, e ela raramente está disponível.

3. O critério: participação, justificação e competência

Uma pessoa recebe uma mensagem informando que o local de votação mudou. A informação é falsa e foi divulgada para conduzi-la ao endereço errado. Ninguém declarou que ela perdeu o direito de votar. Seu nome continua no cadastro, sua condição de eleitora permanece reconhecida e as regras que asseguram sua participação não foram alteradas. Ainda assim, a mensagem pode impedir o exercício dessa posição. O reconhecimento de um direito e as condições de exercê-lo são coisas distintas.

Considere-se agora uma denúncia de irregularidade na apuração. Sua divulgação pode reduzir a confiança no procedimento e provocar contestação. Isso não basta para aproximá-la da mensagem anterior. A denúncia pode revelar um defeito real, exigir uma investigação ou apresentar uma suspeita que não se confirma. A confiabilidade de um procedimento inclui a possibilidade de examinar seus erros. Preservá-la não significa conservar intacta a confiança que as pessoas depositam nele. Há, porém, uma denúncia de outro tipo: a que se mantém idêntica depois da apuração, da perícia e do recurso, e que serve para convocar, não para examinar. Essa denúncia fecha a possibilidade que invoca. A diferença entre as duas não está

Este é um artigo publicado em acesso aberto (Open Access) sob a licença CreativeCommons Attribution, que permite uso, distribuição e reprodução em qualquer meio, sem restrições desde que o trabalho original seja corretamente citado.

Ano VII, v.1 2026 | submissão: 29/10/2026 | aceito: 05/11/2026 | publicação: 10/11/2026

no conteúdo, que pode ser o mesmo, mas na relação que cada uma mantém com o exame que pede: uma o abre, a outra o substitui.

Os dois exemplos mostram a dificuldade de tomar a alteração de uma condição comum como fundamento imediato de restrição. A expressão modifica entendimentos, expectativas e relações. Algumas dessas mudanças permitem corrigir injustiças; outras dificultam a participação de quem já ocupava uma posição. Identificar uma transformação é o começo da avaliação.

Uma condição comum, no domínio aqui considerado, é aquela da qual depende o exercício da participação de várias pessoas: a possibilidade de acessar informações operacionais de uma eleição, contestar uma decisão ou circular sem estar submetido a ameaças dirigidas. Sua existência não exige consenso sobre o que é verdadeiro ou justo. Pessoas que discordam profundamente podem depender das mesmas condições para sustentar sua discordância. Tampouco se presume que toda condição compartilhada seja legítima. Uma prática excludente pode ser estável, antiga e amplamente aceita; essas características não justificam sua preservação. O que a distingue de uma condição legítima é precisamente o que o critério examina: a prática excludente trata parte de quem dela depende como se não dela participasse, e por isso a segunda pergunta, adiante, a alcança do mesmo modo que alcança a fala.

A participação designa uma relação determinada com a condição examinada. O eleitor participa do procedimento eleitoral; o residente de um lugar depende das condições que tornam possível habitá-lo; quem apresenta uma contestação participa da prática em que razões podem ser oferecidas e respondidas. Identificar essas relações exige descrição e argumento. Não basta chamar alguém de participante, nem dizer que toda pessoa alcançada por uma consequência ocupa a mesma posição. O ponto de partida é o do livro de que este critério deriva: ninguém começa produzindo aquilo de que vive, e a recepção de uma condição é um modo de participar dela, não um grau menor de participação (DIETRICH, 2026a, caps. 1, 4 e 8); e o campo em que as posições se distribuem precede o ponto de vista de quem age sobre ele (DIETRICH, 2026a, cap. 15).

A primeira pergunta é diagnóstica: o que a fala fez, ou criou razões suficientes para recear que fizesse, com uma condição de participação? A resposta precisa identificar o ato, a condição e as pessoas cuja posição está em questão. Essas pessoas estão dentro do ato pela posição que ocupam, não pela intenção de quem fala: a direção da ação não nasce da intenção do agente, mas do lugar sobre o qual ela recai (DIETRICH, 2026a, cap. 11). A resposta também deve distinguir a transformação já ocorrida do risco de transformação. Uma mensagem que não

Este é um artigo publicado em acesso aberto (Open Access) sob a licença CreativeCommons Attribution, que permite uso, distribuição e reprodução em qualquer meio, sem restrições desde que o trabalho original seja corretamente citado.

Ano VII, v.1 2026 | submissão: 29/10/2026 | aceito: 05/11/2026 | publicação: 10/11/2026

chega a seus destinatários, uma ameaça que altera imediatamente sua conduta e uma convocação que encontra uma audiência mobilizada apresentam problemas diferentes. A possibilidade abstrata de produzir consequências não permite tratá-las como equivalentes.

Essa descrição não contém ainda o juízo de injustiça. Uma denúncia pode tornar inviável a continuidade de uma prática porque revelou seus abusos. Uma informação verdadeira pode frustrar expectativas cuja manutenção não era devida. Uma crítica pode abalar a autoridade de quem dependia da ausência de contestação. A transformação das relações anteriores exige exame; não estabelece uma prioridade geral de sua conservação.

A segunda pergunta é normativa: por que a transformação identificada é justificável ou injustificável perante aqueles cuja participação alcança? O fundamento assumido é que quem intervém nas condições de participação de outros não pode considerar essa intervenção assunto exclusivamente seu (DIETRICH, 2026a, cap. 13). A premissa é assumida, e sua defesa pode ser dada em poucas linhas. Quem age sobre uma condição comum vive dela tanto quanto aqueles cuja posição altera; tratar a alteração como assunto só seu seria servir-se do que é de todos como se fosse próprio, reivindicando para si um uso que nega aos demais. A premissa recusa essa assimetria, e quem a rejeita precisa explicar por que a condição de que todos vivem pode ser reorganizada por um só sem razões dirigidas aos outros. Ela exige razões dirigidas às posições atingidas. Não exige que toda fala receba autorização prévia, nem que cada pessoa afetada disponha de veto sobre sua circulação. A formulação aproxima-se do princípio do discurso de Habermas, segundo o qual são válidas as normas de ação a que todos os possíveis atingidos poderiam dar assentimento como participantes de discursos racionais, e dele se afasta em dois pontos. Habermas define o atingido pelos interesses que as consequências de uma regulação afetam; aqui, o destinatário é definido pela posição que ocupa na condição alterada. A diferença não é apenas verbal: o anunciante cuja receita cai com a remoção de um conteúdo tem interesse afetado e nenhuma posição na condição de participação em causa; o eleitor que nunca recebeu a mensagem falsa não tem interesse individual atingido e ocupa, ainda assim, posição no procedimento cuja confiabilidade ela afetou. E o princípio do discurso submete normas a um assentimento hipotético, enquanto a premissa aqui assumida exige razões dirigidas às posições atingidas sem fazer de seu assentimento condição da ação. A proximidade tem uma consequência que importa adiante: para Habermas, o discurso que valida normas pressupõe a livre circulação de temas, informações e argumentos num espaço público (HABERMAS, 1997, v. 1, p. 142), e a intervenção que restringe a expressão incide sobre essas mesmas condições.

Este é um artigo publicado em acesso aberto (Open Access) sob a licença CreativeCommons Attribution, que permite uso, distribuição e reprodução em qualquer meio, sem restrições desde que o trabalho original seja corretamente citado.

Ano VII, v.1 2026 | submissão: 29/10/2026 | aceito: 05/11/2026 | publicação: 10/11/2026

A proximidade é maior com Forst, para quem não deve haver relações políticas ou sociais de governo que não possam ser adequadamente justificadas àqueles por elas afetados, por razões que satisfaçam reciprocidade e generalidade, e para quem o núcleo da injustiça é o arbítrio, o domínio sem fundamento legítimo (FORST, 2012, p. 1-2, 20-21). Um traço o separa do critério aqui adotado, e é ele que o argumento precisa defender: Forst funda o direito à justificação na dignidade do ser que pode pedir e dar razões; a premissa aqui assumida o funda na posição ocupada numa condição comum. Sua noção de ordem de justificação — todo contexto em que normas e instituições governam a vida em comum — tem, porém, uma consequência que importa ao argumento: a plataforma que fixa e aplica regras de circulação para milhões de pessoas governa, e quem governa deve razões aos governados.

A distinção entre negar uma pretensão e negar a condição de participante ajuda nesse exame. Contestar uma proposta não equivale a afirmar que seu autor deixou de ser alguém a quem razões são devidas. Entretanto, a primeira operação também pode ser injustificada: uma pretensão pode ser recusada por razões falsas, irrelevantes ou discriminatórias. E a exclusão pode ocorrer sem declaração expressa de inferioridade. A mensagem sobre o endereço eleitoral obstrui um exercício enquanto conserva formalmente o reconhecimento de quem deveria exercê-lo.

Por isso, a negação de status e a obstrução do exercício são indicadores de problemas que precisam ser examinados. Nenhum deles determina sozinho a medida cabível. A avaliação depende do que foi negado, do modo da intervenção, das razões apresentadas e das evidências disponíveis. Uma razão coletiva pode alcançar cada participante, mas a invocação de um benefício geral não dispensa explicar a posição concretamente imposta a ele.

A terceira pergunta concerne à resposta e à competência. Reconhecer uma fala como injustificada não confere a qualquer pessoa o poder de suprimi-la. Contestar, corrigir, reparar, reduzir alcance, remover conteúdo e punir são respostas diferentes. Cada uma intervém de modo próprio nas relações existentes e exige uma justificação correspondente. A escolha deve considerar sua aptidão para enfrentar o problema, as alternativas disponíveis, sua extensão e duração, além das condições de contestação e revisão. Um jurista reconhecerá nessas exigências o exame de proporcionalidade — adequação, necessidade, ponderação em sentido estrito (ALEXY, 2008) — e a semelhança é real. O critério da participação não propõe um teste diferente, e o controle das restrições a direitos é, há muito, objeto da teoria constitucional (ALEXY, 2008; SILVA, 2002). O que ele oferece é uma fundamentação comum para examinar a fala e a intervenção: identifica os destinatários de ambas — as pessoas cujas posições cada

Este é um artigo publicado em acesso aberto (Open Access) sob a licença CreativeCommons Attribution, que permite uso, distribuição e reprodução em qualquer meio, sem restrições desde que o trabalho original seja corretamente citado.

Ano VII, v.1 2026 | submissão: 29/10/2026 | aceito: 05/11/2026 | publicação: 10/11/2026

uma altera — e exige, para ambas, a possibilidade de contestação efetiva em tempo útil. Se essa leitura torna visível algo que uma aplicação limitada ao conteúdo da fala deixaria de examinar, isso tem de ser mostrado em um caso. Há ainda um limite que o livro formula e que vale para a resposta mais do que para a fala: a resposta não pode destruir aquilo mesmo a respeito do qual é devida (DIETRICH, 2026a, cap. 18). A resposta que elimina a prática em que razões podem ser oferecidas e respondidas destrói a condição que invocava para justificar-se.

Quem restringe a expressão também reorganiza uma condição de participação. O bloqueio de um canal pode interromper uma ameaça e impedir outras manifestações; a remoção pode proteger alguém e eliminar material necessário à crítica da própria decisão. Esses efeitos não tornam toda intervenção inadmissível. Tornam insuficiente justificá-la apenas pela identificação de um defeito na fala anterior.

A exigência de justificação alcança, assim, o falante e o restritor, sem tornar idênticas suas posições. Uma autoridade que pode impor decisões dispõe de competências e meios que um participante isolado não possui. Uma plataforma que controla acesso e circulação pode instituir dependências práticas mesmo quando suas regras anunciam garantias. Importa verificar se a contestação funciona em tempo útil e se quem decide permanece sujeito a controle efetivo.

As perguntas valem em dois níveis, que convém não confundir. No nível do regime — as regras legais, judiciais ou contratuais que organizam a expressão e sua moderação —, pergunta-se como essas regras distribuem as possibilidades de falar, de receber, de contestar e de obter revisão. No nível da intervenção particular, pergunta-se se a aplicação dessas regras ao caso é justificada, competente e contestável em tempo útil. A distinção acompanha a que Rawls traçou entre justificar uma prática e justificar uma ação que ela define (RAWLS, 1955, p. 3-7), sem reduzir um nível ao outro: uma regra pode ser injusta e corretamente aplicada, e uma regra justificável pode ser aplicada de modo errado. Na moderação em escala, o exame das decisões isoladas não basta, porque políticas, sistemas automatizados e a distribuição dos erros também reorganizam a participação (DOUEK, 2022); tampouco se dispensa, porque é na decisão particular que o erro alcança alguém. O mesmo vale para quem fala: a permissão jurídica de uma fala não a torna, por si, justificada perante quem ela alcança, e o dever de justificação não se converte em prestação de contas de cada frase antes de sua circulação. Ele se torna exigível quando a fala alcança alguém e altera, ou cria razões suficientes para recear que altere, uma condição de participação.

O critério da participação organiza essas exigências sem substituir os argumentos necessários em cada etapa. A descrição de uma transformação não prova sua injustiça; a

Este é um artigo publicado em acesso aberto (Open Access) sob a licença CreativeCommons Attribution, que permite uso, distribuição e reprodução em qualquer meio, sem restrições desde que o trabalho original seja corretamente citado.

Ano VII, v.1 2026 | submissão: 29/10/2026 | aceito: 05/11/2026 | publicação: 10/11/2026

injustiça não determina automaticamente uma sanção; a finalidade legítima da resposta não demonstra a competência de quem a impõe. Entre a fala e sua restrição há pessoas cujas posições precisam permanecer identificáveis, inclusive quando se conclui que uma pretensão sua deve ser recusada.

4. O que a reorganização não autoriza

Identificar que uma fala alterou uma condição comum, e que por isso deve resposta a quem dela vive, não basta para autorizar uma intervenção coercitiva. A afirmação parece modesta e não é: a maior parte dos argumentos restritivos consiste exatamente em omiti-la. Convém, por isso, percorrer o que fica aquém das duas primeiras perguntas e o que, mesmo cruzando-as, ainda não indica medida alguma.

Fica aquém, primeiro, a opinião que afirma sem atribuir crime, sem convocar e sem obstruir, ainda que seja falsa e ainda que seja odiosa. Uma pessoa que sustenta que determinada política é injusta, que determinado grupo está errado ou que determinada instituição não merece confiança nega pretensões; não nega, por esse ato, a ninguém a condição de participante. A resposta devida a essa fala é a contestação, e quem a contesta reconhece, pelo próprio ato, que seu autor continua sendo alguém a quem razões são devidas. A falsidade e o caráter ofensivo de uma afirmação não demonstram, isoladamente, que ela tenha obstruído a participação de alguém; demonstram que há o que contestar. Nesse ponto o critério se separa de Waldron, que defende restringir a difamação de grupo publicada e duradoura justamente porque ela corrói, pela permanência, a garantia pública que aqui serve para descrever o dano (WALDRON, 2012). O critério acolhe a descrição e recusa a passagem: a corrosão da garantia é uma alteração a diagnosticar, e só a demonstração de que ela ocorreu, numa forma e num alcance determinados, autoriza passar à pergunta sobre a medida. Kaufmann formula o mesmo ponto como limite da tolerância: ela não é ilimitada, mas daí não se segue que a intolerância deva ser sempre combatida com intolerância, e, quando a fala intolerante não põe em perigo a liberdade garantida, falta fundamento para negar-lhe a liberdade (KAUFMANN, 2004, p. 494-495). Se a ofensa é reiterada, dirigida a uma pessoa determinada e organizada de modo a afastá-la de um espaço, o que está em causa já não é a ofensa, e sim a obstrução, e o diagnóstico da seção anterior se aplica.

Fica aquém, em segundo lugar, o mito e o rótulo político enquanto rótulo. Chamar um adversário de genocida, de ladrão ou de comunista substitui o exame pelo pertencimento, e a praça em que o exame poderia ocorrer fica mais pobre. O que o segundo trabalho estabeleceu

Este é um artigo publicado em acesso aberto (Open Access) sob a licença CreativeCommons Attribution, que permite uso, distribuição e reprodução em qualquer meio, sem restrições desde que o trabalho original seja corretamente citado.

Ano VII, v.1 2026 | submissão: 29/10/2026 | aceito: 05/11/2026 | publicação: 10/11/2026

vale aqui: a degradação do debate pelo mito pede exame, e o exame é a resposta que não reproduz o mito com sinal trocado (DIETRICH, 2026c). Mas o emprego de um rótulo não é, por si, um diagnóstico. A mesma palavra pode funcionar como crítica política, como imputação determinada de um crime, como instrumento de assédio reiterado ou como componente de uma campanha de exclusão, e o que ela fez com a participação de alguém depende do uso, dos destinatários e do contexto. Poder responder formalmente não demonstra que o exercício da participação permaneceu desobstruído — a eleitora do primeiro exemplo também podia responder, e nem por isso chegou ao local certo. O que se pode afirmar é mais estreito e mais seguro: falsidade, caráter ofensivo ou emprego de um rótulo não demonstram, isoladamente, obstrução da participação nem autorizam restrição; o que demonstram ou deixam de demonstrar é matéria do diagnóstico, caso a caso.

Fica aquém, em terceiro lugar, a ambiguidade. Uma frase que admite uma leitura ofensiva e uma leitura crítica não reorganiza, pelo simples fato de admitir a primeira, condição alguma; a primeira pergunta exige que o ato tenha feito algo, ou criado razões suficientes para recear que fizesse, e a possibilidade de ser lido de certo modo não é nem uma coisa nem outra. O Tribunal Constitucional alemão não decidiu que soldados não são assassinos; decidiu que a interpretação mais ofensiva de uma frase ambígua não autoriza o Estado a escolhê-la contra o falante (BVerfGE 93, 266). A regra tem um corolário que o critério acolhe: onde o diagnóstico não consegue distinguir a degradação da destruição, a resposta devida é a de menor grau, porque a dúvida sobre o que a fala fez não pode ser resolvida contra quem falou. Hegel já via a dificuldade pelos dois lados. Recusava definir a liberdade de imprensa como liberdade de dizer e escrever o que se quer, definição que tinha por tão superficial quanto a da liberdade como fazer o que se quer; e reconhecia que a indeterminação da expressão impede as leis de alcançar a precisão que se exige da lei, de modo que o juízo sobre ela tende a aparecer como decisão subjetiva (HEGEL, 1997, § 319, p. 290-291). O primeiro lado recusa a imunidade; o segundo é, desde então, o problema de quem restringe.

Fica aquém, por fim, o que o primeiro trabalho descreveu: o mecanismo pelo qual o eleitor julga sem exame e acredita possuir competência para fazê-lo não demonstra, isoladamente, que alguém o tenha lesado. O mecanismo descreve um modo de receber, não uma posição reorganizada contra ele, e a resposta a um modo de receber é a oferta de exame, não a contenção de quem o provocou.

Há, depois, o que cruza as duas primeiras perguntas e ainda assim não indica a medida. Uma fala que verbaliza a exclusão de um grupo — que diz, em termos, que seus membros não

Este é um artigo publicado em acesso aberto (Open Access) sob a licença CreativeCommons Attribution, que permite uso, distribuição e reprodução em qualquer meio, sem restrições desde que o trabalho original seja corretamente citado.

Ano VII, v.1 2026 | submissão: 29/10/2026 | aceito: 05/11/2026 | publicação: 10/11/2026

são destinatários de nada — demonstra a injustificação. Não demonstra se a resposta devida é a contestação pública, a correção, a reparação ou a contenção. Isso depende do que a terceira pergunta examina: do contexto em que a fala foi dita, do alcance que teve, do risco que criou e das alternativas à medida que se pretende adotar. Mill já o observava: a opinião de que os comerciantes de cereais matam os pobres de fome pode circular livremente pela imprensa e ser legitimamente punida quando dita a uma multidão exaltada diante da casa de um deles (MILL, 2006, cap. III). O conteúdo é o mesmo; o ato não é.

Tampouco indicam a medida, por si, os limites que o terceiro trabalho fixou para a intervenção que propôs. Falsidade verificada, crise declarada e o conjunto de fatores que caracterizam a fala perigosa — falante, audiência, mensagem, contexto e meio (BENESCH, 2013) — delimitaram uma medida específica, a redução temporária de alcance de conteúdo já desmentido durante os dias de Southport. Como condições daquela medida, são razoáveis. Como requisitos de toda resposta legítima à expressão, não se sustentam. A incitação de Connolly usava uma informação verdadeira sobre a existência de hotéis que abrigavam requerentes de asilo (CONNOLLY v. R, 2025); a ameaça dirigida a uma pessoa determinada não precisa de crise para alterar imediatamente a conduta de quem a recebe; e a exigência de crise declarada entrega a quem a declara o poder de definir quando a proteção da fala diminui — um poder que o critério, aplicado ao restritor, não pode conceder sem exame. Os três limites permanecem como caso, não como regra.

E não indica a medida, por último, a descrição que a torna leve. Dizer que reduzir o alcance de uma publicação não é censura, porque não a remove e apenas recusa recompensá-la, descreve o mecanismo e não o justifica. A redução de alcance é uma intervenção sobre a condição de participação de quem publicou e de quem leria; sua legitimidade depende de quem a determina, dos efeitos que produz e das possibilidades de contestá-la, e essas possibilidades são, na prática, menores do que as que acompanham a remoção, porque a redução é menos visível. Uma intervenção invisível não é uma intervenção menor; é uma intervenção com menos controle.

O que se segue do diagnóstico e da justificação é, portanto, que uma resposta é devida. Não se segue qual resposta, nem por quem.

5. Resposta, competência e controle

As respostas a uma fala que alterou injustificadamente uma condição comum não são equivalentes entre si, e a diferença entre elas não é apenas de intensidade. Contestar

Este é um artigo publicado em acesso aberto (Open Access) sob a licença CreativeCommons Attribution, que permite uso, distribuição e reprodução em qualquer meio, sem restrições desde que o trabalho original seja corretamente citado.

Ano VII, v.1 2026 | submissão: 29/10/2026 | aceito: 05/11/2026 | publicação: 10/11/2026

publicamente, corrigir, reparar, reduzir alcance, remover e punir intervêm de modos distintos nas relações existentes, exigem de quem as adota títulos distintos e admitem controles distintos. O que se pode dizer de todas é que cada uma deve ser comparada concretamente com as alternativas, segundo sua extensão, sua duração, sua reversibilidade, a urgência que a justifica, seus efeitos sobre terceiros e a possibilidade de revê-la; e que quem adota uma delas deve explicar por que uma alternativa menos onerosa para a participação do falante, e suficientemente eficaz, não bastava. Não se pode dizer que elas ocupem sempre a mesma ordem. Uma indenização elevada pode ser mais onerosa do que a retirada pontual de uma ameaça; acrescentar contexto a uma publicação não tem o mesmo peso que reduzir seu alcance; uma ameaça urgente pode exigir remoção sem passar antes pela correção pública. A preferência pela intervenção menos onerosa é uma exigência; uma sequência fixa de degraus não é.

Convém, ainda assim, nomear as respostas, porque cada uma tem uma natureza própria. A contestação pública é a resposta que qualquer participante pode dar e que não exige título algum; o direito de resposta, quando exigido institucionalmente, já impõe uma obrigação a um terceiro — o veículo, a plataforma — e por isso não é a mesma coisa que contestar livremente. A correção pública tempestiva por quem detém a informação é a resposta que só uma autoridade determinada pode dar: quando uma atribuição falsa circula e quem poderia desmenti-la se cala, o silêncio não é neutralidade. Shklar chamou de injustiça passiva a omissão de quem podia e devia impedir um mal e não o fez, e lembrou que a autoridade pública que nada faz está entre os casos mais graves dessa omissão (SHKLAR, 1990, p. 40-46); a inspeção britânica sobre Southport apontou o vácuo informativo como falha, o que não dispensa distinguir o que a lei permitia dizer (HMICFRS, 2025). A reparação civil responde a um ato passado sem conter um ato futuro. A redução de alcance, a rotulagem e o contexto adicionado mantêm a fala disponível e alteram sua circulação, cada qual de um modo. A remoção retira a publicação; a sanção penal atinge o falante. Essas diferenças de natureza são o que a comparação concreta tem de considerar, e são também o que torna cada resposta uma reorganização distinta das condições de quem fala, de quem lê e de quem precisará examinar a decisão.

As exigências que acompanham qualquer dessas respostas podem ser formuladas como seis. A primeira é um diagnóstico demonstrado: a conduta expressiva identificada, a condição identificada e evidência suficiente do dano ou do risco que a medida se propõe a enfrentar, segundo a natureza da intervenção. Três coisas têm de ser distinguidas aqui, porque costumam ser confundidas: a prova de que a fala ocorreu e do que ela disse; a prova do dano consumado

Este é um artigo publicado em acesso aberto (Open Access) sob a licença CreativeCommons Attribution, que permite uso, distribuição e reprodução em qualquer meio, sem restrições desde que o trabalho original seja corretamente citado.

Ano VII, v.1 2026 | submissão: 29/10/2026 | aceito: 05/11/2026 | publicação: 10/11/2026

ou do risco relevante que ela criou; e o fundamento e os requisitos da medida escolhida. Uma ameaça pode exigir interrupção antes da agressão anunciada, e exigir a consumação do dano final como condição de toda medida equivaleria a proibir a prevenção; o que se exige é que o risco seja demonstrável, e demonstrado, na medida da intervenção. A circulação, isoladamente, não demonstra nenhum desses elementos. Quando a exigência falha, a intervenção responde a uma classe de conteúdo, não a um ato — e responder a classes é o modo mais seguro de alcançar quem não fez nada. A segunda é uma justificação enunciada: a negação de status ou a obstrução de exercício nomeadas, e a razão da medida dita a quem a sofre. Uma medida cuja razão não pode ser dita ao atingido é uma preferência com força.

A terceira exigência é a competência. Quem decide precisa ter título para reorganizar a condição de participação do falante, e a pergunta não é se há um título, mas qual, com que limites e sob que controles. Um tribunal o tem por mandato; uma autoridade administrativa, pela norma que o fixa; uma plataforma, por um conjunto mais variado de fontes — a lei que lhe impõe deveres, mas também as condições contratuais de uso, a organização de um fórum temático, o combate ao que degrada o serviço — e a pergunta relevante é quais poderes essas regras legitimam e a que exame elas próprias se submetem. O Regulamento europeu de Serviços Digitais admite decisões fundadas tanto na ilegalidade do conteúdo quanto em sua incompatibilidade com os termos do serviço, e submete umas e outras a deveres de fundamentação (UNIÃO EUROPEIA, 2022, art. 17). Isso não resolve a questão normativa; mostra que excluir toda política própria como fundamento seria um compromisso forte, e que o problema está na fundamentação e no controle, não na origem da regra. A exigência parece óbvia e é a mais frequentemente omitida, porque a gravidade do dano tende a ser tomada como título suficiente para qualquer um responder a ele. Não é. O trabalho sobre Coase mostrou, em outro domínio, que um administrador pode ter deveres formais e conservar o controle prático unilateral da reorganização que administra (DIETRICH, 2026e); a observação vale para quem remove, bloqueia ou reduz alcance. A quarta exigência é a de garantias e controle: contraditório, recurso, prazo, publicidade da decisão e do padrão que ela aplicou, e a possibilidade de contestar a redução de alcance tanto quanto a remoção. Solum sustentou, a propósito do processo judicial, que a legitimidade de uma decisão vinculante depende de que os atingidos possam participar num momento em que a participação ainda importa, e que esse valor não se reduz ao efeito da participação sobre o resultado (SOLUM, 2004, p. 275). Solum restringe a tese às decisões finais e vinculantes e exclui expressamente as temporárias. A extensão à moderação é, portanto, argumento próprio, não de Solum: a remoção definitiva aproxima-se do caso que ele examina,

Este é um artigo publicado em acesso aberto (Open Access) sob a licença CreativeCommons Attribution, que permite uso, distribuição e reprodução em qualquer meio, sem restrições desde que o trabalho original seja corretamente citado.

Ano VII, v.1 2026 | submissão: 29/10/2026 | aceito: 05/11/2026 | publicação: 10/11/2026

e mesmo uma medida temporária pode ter efeito definitivo quando atinge a participação num debate datado — a oportunidade perdida na véspera de uma eleição não se recupera depois dela. Sem controle, a resposta se torna exceção, e a exceção é o sinal de que a resposta entrou no ciclo que pretendia interromper.

Uma consequência dessa abertura precisa ser dita. Um fórum dedicado à literatura pode retirar propaganda eleitoral que nada tem de injusta, apenas porque é estranha ao tema que organiza aquele espaço. A retirada não se justifica pela injustiça da fala, que não houve, e sim pela finalidade declarada do espaço, que os participantes conheciam ao entrar e da qual também vivem. O caso separa duas perguntas que o argumento até aqui tendeu a aproximar: por que uma fala é injustificada, e por que sua exclusão de um espaço determinado é justificada. A segunda pode ser respondida sem a primeira quando a regra do espaço é pública, anterior ao caso e aplicada de modo uniforme, e quando a fala conserva outros lugares onde circular. A resposta muda à medida que o espaço deixa de ser um entre muitos e passa a ser aquele por onde passa a maior parte da circulação: a finalidade temática de um fórum não se transfere a uma plataforma que organiza o debate público de um país, e quanto mais a participação depende de um espaço, mais a regra que o organiza se aproxima de uma regra do próprio regime e mais justificação exige. A distinção acompanha a de Post entre o discurso público, cuja proteção serve a que as pessoas experimentem governar a si mesmas, e a expressão que circula fora dele, regulada por outras finalidades (POST, 2011, p. 483-486). Post formula a distinção para a proteção constitucional contra o Estado; estendê-la aos deveres das plataformas é argumento próprio, que se apoia no critério e não na classificação de Post: o fórum temático regula a expressão segundo sua finalidade, e a plataforma que se tornou o lugar do debate público, por reorganizar a condição de que esse debate depende, já não pode fazê-lo sem as razões que se exigem de quem o regula.

A quinta exigência é a de condições de revisão. Uma medida que não deixa ao falante a possibilidade de rever o que disse sem confessar culpa total fecha um caminho que a resposta tinha razões para manter aberto; a psicologia da dissonância sugere que a pena que o fecha pode reforçar a autoimagem que pretendia corrigir (TAVRIS; ARONSON, 2007), mas a reação a uma medida pode reforçar convicções, produzir revisão ou apenas interromper a conduta, e a legitimidade de uma resposta não depende de conseguir modificar a compreensão do agente. O que se exige é mais modesto: que a medida preserve condições de revisão, não que as garanta. A sexta é a distribuição da responsabilidade. Alcance e posição são fatores relevantes — quem forma o entendimento de muitos responde por algo que quem apenas repete não fez, embora

Este é um artigo publicado em acesso aberto (Open Access) sob a licença CreativeCommons Attribution, que permite uso, distribuição e reprodução em qualquer meio, sem restrições desde que o trabalho original seja corretamente citado.

Ano VII, v.1 2026 | submissão: 29/10/2026 | aceito: 05/11/2026 | publicação: 10/11/2026

quem repete, à medida que repete, não possa dizer-se alheio ao que repete (DIETRICH, 2026a, cap. 22) — mas não substituem o conhecimento, a intenção, a contribuição efetiva e as possibilidades de agir de outro modo, que o critério de responsabilidade do terceiro trabalho já reunia: acesso efetivo à informação que desmentiria a fala, capacidade de examiná-la, razões para duvidar e oportunidades de exame (DIETRICH, 2026d). Uma resposta que atinge a massa e poupa quem originou inverteu a distribuição; uma que atinge quem originou sem examinar o que ele sabia e podia inverteu-a de outro modo.

Dois casos documentados mostram as exigências em operação, com resultados desiguais e, em parte, inconclusivos; um terceiro fica apenas delimitado. No caso de Connolly, a condição era a garantia pública de segurança dos residentes de hotéis que abrigavam requerentes de asilo; a conduta, uma incitação inequívoca a incendiá-los, publicada na noite de 29 de julho de 2024, horas depois do ataque e antes dos distúrbios que começariam no dia seguinte; o risco tem de ser avaliado pelas circunstâncias daquele momento — a comoção pelo ataque e a circulação de falsidades sobre seu autor —, sem transportar para a audiência a violência posterior; a medida, a sanção penal. A competência era a de um tribunal ordinário, e o controle existiu: o juiz de primeiro grau tratou como agravante o momento da publicação, numa noite de especial sensibilidade, e a Court of Appeal examinou a pena de trinta e um meses, fundamentou sua adequação e recusou autorização para recorrer dela, de modo que a pena permaneceu inalterada (CONNOLLY v. R, 2025). O caso ilustra, portanto, a fundamentação judicial e o controle recursal — as exigências terceira e quarta. Não ilustra, por si, as demais: o acórdão não compara a prisão com alternativas menos onerosas no sentido aqui exigido, e nada diz sobre as condições de revisão deixadas à condenada. Examinado exigência por exigência, o caso é concludente quanto a duas e inconclusivo quanto às outras, e é assim que um caso pode testar um critério. O caso do nome falso atribuído ao agressor de Southport é mais complexo, e por isso mais útil. A condição era a mesma; a conduta, uma atribuição falsa de crime a uma categoria; o nome falso foi visto cerca de quatrocentas e vinte mil vezes, número que mede alcance potencial, não pessoas convencidas (HOUSE OF COMMONS, 2025a, §8). A resposta de quem detinha a informação estava limitada de modos diferentes. O nome do acusado, que tinha dezessete anos, estava protegido por restrições legais — as regras de desacato ao tribunal e as que resguardam a identidade de acusados menores; outras informações corretivas podiam ser divulgadas e foram, como o local de nascimento e o desmentido do nome falso; sobre outras ainda, como a religião, houve orientações divergentes entre instâncias do Ministério Público; e a chefe de polícia de Merseyside não soube dizer à comissão parlamentar se a divulgação da religião teria

Este é um artigo publicado em acesso aberto (Open Access) sob a licença CreativeCommons Attribution, que permite uso, distribuição e reprodução em qualquer meio, sem restrições desde que o trabalho original seja corretamente citado.

Ano VII, v.1 2026 | submissão: 29/10/2026 | aceito: 05/11/2026 | publicação: 10/11/2026

evitado novos distúrbios. A comissão registrou que essas restrições criaram um vácuo em que a desinformação cresceu (HOUSE OF COMMONS, 2025b). O caso não é, portanto, de silêncio simples nem de silêncio imposto. É um conflito entre a garantia pública da comunidade ameaçada e a proteção processual do acusado, ambos participantes da condição em jogo, e o critério exige que cada categoria de informação seja examinada perante uns e outros. A injustiça passiva de que fala Shklar alcança apenas a parte da omissão não coberta por um dever concorrente; a própria Shklar mede a responsabilidade de quem tem deveres concorrentes à luz deles (SHKLAR, 1990, p. 44-45). O que o caso permite dizer é estreito. Onde a correção era juridicamente possível e tardou, houve uma resposta devida que não veio a tempo. Onde a lei impunha reserva, a pergunta desloca-se para o regime — as regras de anonimato de menores — e para o modo de preservá-las sem deixar o vácuo ocupado pela falsidade. Em nenhum dos dois planos o caso demonstra que a publicação, ou a falta de correção, tenha causado os ataques. No caso das convocações que antecederam o 8 de janeiro de 2023, a condição era institucional; a conduta, a convocação a impedir o funcionamento das instituições; a resposta, a responsabilização de quem convocou, em processos perante o Supremo Tribunal Federal. O caso precisa, porém, de uma precisão que a fórmula esconde. Em janeiro de 2023 o procedimento eleitoral estava concluído; a condição em jogo era o respeito ao resultado proclamado e o funcionamento das instituições que dele decorrem, e as convocações que antecederam a invasão das sedes dos três Poderes dirigiam-se contra esse funcionamento. Avaliar a resposta pelas seis exigências pediria o exame de decisões identificadas, uma a uma, e a existência de processo e de decisões publicadas não demonstra, por si, que elas tenham sido satisfeitas. O caso fica registrado como aquele em que o critério teria de ser aplicado com mais cautela, porque nele a simetria é mais difícil de sustentar.

As seis exigências podem ser postas como cinco perguntas a qualquer intervenção. Que condição foi alterada, por que conduta, com que evidência do dano ou do risco? A alteração nega status ou obstrui exercício, e por quê? Qual resposta, e por que uma alternativa menos onerosa e suficientemente eficaz não bastava? Quem adota a medida tem título para reorganizar a condição do falante, e qual? Que garantias e que controle acompanham a medida, inclusive quando ela é a redução de alcance? Postas ao regime, as mesmas perguntas recaem sobre as regras: que condições elas permitem alterar, que razões as sustentam, que medidas autorizam, a quem dão título e que revisão garantem — inclusive sobre a distribuição dos erros que nenhuma regra aplicada em escala consegue evitar. As quatro primeiras perguntas estavam, de modo disperso, nos três trabalhos anteriores. A quinta não estava em nenhum deles.

Este é um artigo publicado em acesso aberto (Open Access) sob a licença CreativeCommons Attribution, que permite uso, distribuição e reprodução em qualquer meio, sem restrições desde que o trabalho original seja corretamente citado.

Ano VII, v.1 2026 | submissão: 29/10/2026 | aceito: 05/11/2026 | publicação: 10/11/2026

6. A intervenção como reorganização

Quando um perfil é bloqueado, uma publicação removida ou seu alcance reduzido, algo foi reorganizado. Não apenas a situação do falante: também a de quem o lia, a de quem responderia a ele, a de quem precisaria do material removido para avaliar se a remoção foi justa. A restrição da expressão é uma ação sobre uma condição comum — a praça em que razões podem ser oferecidas e respondidas — e, como toda ação desse tipo, tem dentro de si, pela posição que ocupam, as pessoas que dessa condição vivem. O critério que foi dirigido ao falante dirige-se, por isso, a quem o restringe. Não porque as posições sejam idênticas, mas porque o campo é o mesmo.

A diferença entre as posições é real e deve ser dita. Uma autoridade que pode impor decisões dispõe de competências e meios que um participante isolado não tem; uma plataforma que controla o acesso e a circulação pode instituir dependências práticas que nenhum falante institui. Essa diferença não atenua a exigência de justificação; agrava-a. Quem pode mais reorganiza mais, e a quem reorganiza mais, mais se pergunta. O livro de que o critério deriva distingue três modos de estar dentro de uma mesma reorganização — o de quem age, o de quem detém a faculdade de reorganizar e a sustenta, e o de quem a repete — e observa que a resposta de quem detém não pode consistir em suas boas intenções (DIETRICH, 2026a, cap. 22). O restritor ocupa o segundo modo. As três perguntas do critério, aplicadas ao restritor, tomam a seguinte forma. Que condição comum a intervenção reorganiza, e com que evidência de que a reorganização era necessária? Perante quem a intervenção deve ser justificada — e a resposta inclui o falante, que não deixa de ser participante por ter falado mal? Quem controla a decisão de quem interveio, em que prazo e com que publicidade?

A última pergunta é a decisiva, e é a que o trabalho sobre Coase preparou. Ali se mostrou que a existência de deveres formais — de informar, de consultar, de prestar contas — não impede que quem administra uma reorganização conserve o controle prático e unilateral dela, se os deveres não dão a quem os invoca um meio de alterar a decisão em tempo útil (DIETRICH, 2026e). O mesmo vale aqui. Uma tese judicial pode impor às plataformas deveres de transparência, canais de atendimento e relatórios periódicos; a pergunta do critério não é se esses deveres existem, mas se eles dão a quem sofre uma remoção ou uma redução de alcance um meio de contestá-la que funcione antes que o dano à sua participação se consume, ou se apenas documentam, depois, o controle de quem a adotou. Um relatório anual sobre remoções não é um recurso; é uma estatística.

Este é um artigo publicado em acesso aberto (Open Access) sob a licença CreativeCommons Attribution, que permite uso, distribuição e reprodução em qualquer meio, sem restrições desde que o trabalho original seja corretamente citado.

Ano VII, v.1 2026 | submissão: 29/10/2026 | aceito: 05/11/2026 | publicação: 10/11/2026

A distinção entre justificar uma prática e justificar uma ação particular que nela se insere ajuda a precisar o ponto. Rawls a formulou a propósito da pena: a quem objetava que certa doutrina autorizaria condenar um inocente para dissuadir os demais, respondeu que era preciso dizer quem é o "nós" que decidiria fazê-lo, com que autoridade e segundo que regras, porque essa autoridade teria de estar escrita na instituição; e uma instituição que concedesse a seus agentes o poder de decidir, caso a caso e segundo seu juízo sobre o bem comum, quem seria condenado dificilmente se justificaria (RAWLS, 1955, p. 10-13). A pergunta vale para a restrição da expressão. Não basta mostrar que, neste caso, seria bom que alguém retirasse esta fala. É preciso descrever a prática que autoriza alguém a retirá-la — quem, segundo que regra, com que controle — e perguntar se essa prática se justifica perante quem dela participa, inclusive os falantes cujas falas ela alcançará por engano.

A redução de alcance é o caso mais agudo. Ela não é invisível por natureza: pode ser comunicada a quem a sofre, acompanhada de razões e submetida a contestação, e o Regulamento europeu de Serviços Digitais exige fundamentação para determinadas restrições de visibilidade, incluída a despriorização (UNIÃO EUROPEIA, 2022, art. 17). O problema está na despriorização opaca — a que não é notificada, não expõe o padrão aplicado e não oferece meio de contestação. Nessa forma, a medida que parece menor é a que menos controle admite. Quem a descreve como mera recusa de recompensar descreve a mecânica; a pergunta continua a ser quem decide o que recompensar, perante quem justifica a decisão e quem a revê.

O teste de simetria a que o critério se submete tem, então, dois eixos. No primeiro, ele deve produzir o mesmo resultado quando o falante pertence ao campo político com que quem aplica o critério se identifica. Dois atos de fala com a mesma estrutura — uma atribuição falsa de crime a uma categoria de pessoas, uma convocação a impedir um procedimento — recebem as mesmas cinco perguntas, e as respostas não mudam com a cor do falante. No segundo eixo, que os três trabalhos anteriores não percorreram, o critério deve alcançar quem restringe: o Estado que bloqueia um perfil, a plataforma que reduz um alcance, o tribunal que impõe uma medida. Se as perguntas valessem para a convocação e não para a exceção que a ela responde, o critério deixaria de examinar o campo comum e passaria a servir a um dos lados.

Há uma objeção a esse ponto que merece ser enfrentada em seus termos. Ela diz que submeter o restritor ao mesmo exame que o falante equipara o incendiário e o bombeiro, e que a simetria é uma forma de paralisia. A réplica é que a simetria não está nas medidas, mas nas perguntas. O incendiário e o bombeiro não recebem a mesma resposta; recebem a mesma exigência de justificar o que fizeram com a casa, e o bombeiro, que entrou com um título que o

Este é um artigo publicado em acesso aberto (Open Access) sob a licença CreativeCommons Attribution, que permite uso, distribuição e reprodução em qualquer meio, sem restrições desde que o trabalho original seja corretamente citado.

Ano VII, v.1 2026 | submissão: 29/10/2026 | aceito: 05/11/2026 | publicação: 10/11/2026

incendiário não tinha, tem mais a dizer, não menos. Um restritor que pode responder às cinco perguntas não é paralisado por elas; é legitimado. O que a simetria paralisa é o restritor que não pode respondê-las, e esse é exatamente o que deve ser paralisado.

Um caso hipotético permite ver o que essa leitura acrescenta. Na véspera de uma eleição, circula em grupos de uma cidade uma mensagem anunciando a transferência de determinadas seções eleitorais para um endereço distante. Uma análise limitada ao conteúdo pergunta se a mensagem é falsa e, sendo, se pode ser removida. A leitura pela participação faz outras perguntas, e na ordem em que importam. Quem está em causa: os eleitores daquelas seções, cuja condição de eleitor permanece intacta e cujo exercício pode ser obstruído; mas também os eleitores de alguma seção efetivamente transferida, se houver, para os quais um aviso verdadeiro de teor semelhante é indispensável. Qual resposta enfrenta o risco no tempo em que ele existe: a correção pela Justiça Eleitoral, que detém a informação sobre os locais de votação e pode publicá-la pelos próprios canais antes da abertura das urnas, atinge os destinatários sem depender de remoção. Se a remoção for adotada, quem a decide e com que controle: uma plataforma que remove em massa, por semelhança textual, mensagens sobre mudança de seção pode retirar também o aviso verdadeiro e obstruir, por excesso, o mesmo exercício que pretendia proteger; e uma decisão só contestável depois da eleição não é contestável em tempo útil. O que a remoção deve preservar: a prova da conduta, para eventual responsabilização de quem embaraçou o exercício do sufrágio (Código Eleitoral, art. 297). Nenhuma dessas perguntas é estranha ao exame de proporcionalidade. O que a leitura pela participação torna visível é que o destinatário da fala e o destinatário da medida devem ser identificados antes que a medida seja escolhida, e que o mesmo erro — obstruir um exercício — pode ser cometido pelo falante e por quem o restringe.

O caso muda quando as condições mudam, e é nas variações que o critério trabalha. Se a correção oficial não alcança os destinatários — porque a mensagem circula em grupos fechados, porque falta tempo, porque a autoridade eleitoral não tem a confiança deles —, a alternativa menos onerosa deixa de ser suficiente, e a remoção ou a limitação do encaminhamento podem tornar-se a resposta devida, sem que as perguntas sobre título e revisão desapareçam. Se a comunicação é aberta, a correção ao lado da mensagem pode bastar; se é cifrada, nem a plataforma vê o conteúdo — e o regime brasileiro exclui dos deveres de cuidado e de notificação as comunicações interpessoais em serviços de mensageria (Decreto 8.771/2016, art. 16-O) —, e a resposta possível desloca-se para quem pode falar aos destinatários. Se o risco é remoto — a mensagem circula semanas antes, quando há tempo para

Este é um artigo publicado em acesso aberto (Open Access) sob a licença CreativeCommons Attribution, que permite uso, distribuição e reprodução em qualquer meio, sem restrições desde que o trabalho original seja corretamente citado.

Ano VII, v.1 2026 | submissão: 29/10/2026 | aceito: 05/11/2026 | publicação: 10/11/2026

desmentir —, falta a urgência que justificaria remover. E se o aviso era verdadeiro, a remoção é ela mesma a obstrução, e o critério volta-se inteiro contra quem restringiu. Em nenhuma das variações o critério decide sozinho; em todas, indica quem deve ser considerado e o que precisa ser demonstrado antes da medida.

Segue-se, por fim, um limite que vale para o restritor com mais força do que para o falante: a resposta não pode destruir aquilo mesmo a respeito do qual é devida (DIETRICH, 2026a, cap. 18). O limite precisa ser lido com cuidado, porque a imagem de fechar a praça confunde três coisas de alcance distinto. Retirar uma publicação não é excluir uma pessoa de um espaço, e nenhuma das duas coisas é eliminar a possibilidade de contestar a própria intervenção. Uma publicação pode ser retirada com preservação da prova, decisão fundamentada e recurso efetivo, e nada disso destrói a praça. A exclusão de uma pessoa altera sua participação de modo mais extenso e duradouro, e exige mais. A intervenção que suprime os meios de contestá-la — porque não é notificada, porque elimina o material de que a crítica dependeria, porque o recurso chega depois que o dano à participação se consumou — é a que atinge a condição que invocava para justificar-se. É nesse terceiro sentido que a resposta fecha a praça. Por isso a quinta pergunta — quem controla, em que prazo, com que publicidade — não é um acréscimo procedimental; é a forma que o critério assume quando se volta para quem o aplica.

7. O direito brasileiro perguntado pelo critério

O direito brasileiro distingue, sem o nomear, o que o critério distingue. Protege a manifestação do pensamento e a expressão, independentemente de censura ou licença (CF, art. 5º, IV e IX; art. 220); prevê a resposta e a reparação antes de qualquer contenção (CF, art. 5º, V e X); e responde a atos — a incitação, a discriminação, a abolição violenta do Estado democrático — e não a opiniões. O exame que segue não conclui sobre esse direito antes de lê-lo; dirige a cada fonte as perguntas do critério, registra o que o texto normativo já responde e delimita o que permanece em aberto.

A primeira fonte é a própria Constituição, e a primeira pergunta é se ela já contém o critério. A vedação do anonimato (art. 5º, IV) costuma ser lida como condição da responsabilização posterior (art. 5º, V e X); pode também ser lida — e é a leitura aqui proposta — como reconhecimento de que a manifestação tem destinatários perante os quais seu autor responde. O direito de resposta (art. 5º, V) e a inviolabilidade da honra (art. 5º, X) fixam a resposta de primeiro grau — a fala contra a fala — antes de qualquer outra. E a decisão que

Este é um artigo publicado em acesso aberto (Open Access) sob a licença CreativeCommons Attribution, que permite uso, distribuição e reprodução em qualquer meio, sem restrições desde que o trabalho original seja corretamente citado.

Ano VII, v.1 2026 | submissão: 29/10/2026 | aceito: 05/11/2026 | publicação: 10/11/2026

afastou a Lei de Imprensa fixou a regra de que a liberdade de expressão tem precedência, com responsabilidade posterior e não restrição prévia (STF, ADPF 130, 2009). A pergunta a fazer à doutrina constitucional é se essa combinação — destinatários perante os quais se responde, resposta antes de contenção, responsabilidade depois e não antes — equivale à comparação entre respostas antes exigida, ou se a precedência da liberdade é lida como dispensa de justificação. A resposta determina se o critério acrescenta algo à Constituição ou apenas a lê.

A segunda fonte são os tipos penais, e aqui a pergunta é dupla. Os crimes contra a honra (CP, arts. 138 a 140) respondem, em regra, a falas que atingem a pretensão de alguém à consideração alheia, e a ação penal depende, na maior parte dos casos, da iniciativa do ofendido. A classificação, porém, não é uniforme: a injúria qualificada do art. 140, § 3º, do Código Penal e a injúria racial, hoje tipificada na Lei 7.716/1989 (art. 2º-A, incluído pela Lei 14.532/2023) e processada por ação pública, atingem precisamente o status. O Supremo Tribunal manteve, para os crimes contra a honra, a exigência de ordem judicial como condição da responsabilidade da plataforma, e cabe perguntar se esse regime alcança também as figuras que negam status. A incitação pública a crime (CP, art. 286) e a incitação à discriminação (Lei 7.716/1989, art. 20) respondem a atos de convocação e de negação de status, e a decisão que confirmou a condenação de um editor por obras antissemitas afirmou que a liberdade de expressão não alcança a incitação ao racismo, em termos que se aproximam da garantia pública de que um grupo não será atacado pelo que é (STF, HC 82.424, 2003; WALDRON, 2012). A fronteira entre defender a mudança de uma norma e incitar ao crime foi traçada pelo Supremo Tribunal no julgamento da "marcha da maconha" (STF, ADPF 187, 2011), e é ali, mais do que em formulações abstratas, que a distinção entre negar uma pretensão e convocar recebe seu teste no direito brasileiro. A pergunta do critério a esses tipos é se sua aplicação em rede exige evidência suficiente da conduta e do dano ou do risco que o tipo pressupõe, e não só da circulação, ou se a gravidade do tipo tem sido tomada como dispensa do diagnóstico. A pergunta não supõe que o resultado danoso deva ocorrer: tipos de incitação existem precisamente para alcançar o risco, e o que o critério pede é que o risco seja demonstrado na medida da resposta.

A terceira fonte é a proteção do procedimento eleitoral e da ordem democrática. A Lei 14.197/2021 tipificou a abolição violenta do Estado democrático de direito e o golpe de Estado, ambos dependentes de violência ou grave ameaça, e a interrupção do processo eleitoral mediante violação indevida dos mecanismos de segurança do sistema eletrônico de votação (CP, arts. 359-L a 359-N). Nenhum desses tipos descreve a convocação enquanto tal. A mesma

Este é um artigo publicado em acesso aberto (Open Access) sob a licença CreativeCommons Attribution, que permite uso, distribuição e reprodução em qualquer meio, sem restrições desde que o trabalho original seja corretamente citado.

Ano VII, v.1 2026 | submissão: 29/10/2026 | aceito: 05/11/2026 | publicação: 10/11/2026

lei, porém, acrescentou ao art. 286 do Código Penal um parágrafo único que pune a incitação pública de animosidade entre as Forças Armadas, ou delas contra os poderes constitucionais, as instituições civis ou a sociedade; por essa via, pela incitação ao crime em geral (caput) ou pela participação nos atos, a fala que convoca entra no campo penal, e a distinção importa, porque a resposta a uma convocação depende do que ela convocou e de como. A mesma lei continha um tipo dirigido a quem promovesse ou financiasse campanha para disseminar fatos sabidamente inverídicos capazes de comprometer a higidez do processo eleitoral (art. 359-O); o dispositivo foi vetado. A ausência desse tipo não deixa sem proteção a eleitora do primeiro exemplo: o Código Eleitoral pune quem impede ou embaraça o exercício do sufrágio (art. 297), e pune também a divulgação, na propaganda ou durante a campanha, de fatos sabidamente inverídicos sobre partidos ou candidatos (art. 323), tipo que, por seu objeto, provavelmente não alcança a mensagem sobre o local de votação. O Supremo Tribunal já examinou, além disso, um poder de remoção rápida de desinformação sobre o processo eleitoral atribuído à Justiça Eleitoral (Resolução TSE 23.714/2022) e o manteve, primeiro ao referendar o indeferimento da cautelar, em 26 de outubro de 2022, depois ao julgar improcedente a ação (STF, ADI 7261); é nesse julgamento que as perguntas sobre título, prazo e controle foram efetivamente enfrentadas no direito brasileiro, e o critério deve ser medido contra ele. A pergunta do critério torna-se, assim, mais estreita: se o embaraço por engano — a mensagem falsa sobre o local de votação — está ao alcance desse tipo e dos instrumentos da Justiça Eleitoral, e com que exigências de diagnóstico, ou se a obstrução de exercício praticada em massa ficou sem resposta proporcional depois do veto.

A quarta fonte é o regime de responsabilidade das plataformas, e é a que mais exige leitura. O art. 19 do Marco Civil da Internet condicionava a responsabilidade civil do provedor ao descumprimento de ordem judicial específica. Em junho de 2025, o Supremo Tribunal Federal declarou a inconstitucionalidade parcial e progressiva do dispositivo e fixou um regime escalonado (STF, RE 1.037.396 e RE 1.057.258, Temas 987 e 533); em 17 de junho de 2026, ao julgar os embargos de declaração, consolidou a tese final, com trânsito em julgado imediato, modulação a partir de 5 de agosto de 2025 e prazo de sessenta dias para as obrigações estruturais (STF, 2026). A tese final distingue regimes: para crimes e atos ilícitos em geral, a notificação extrajudicial basta, e a responsabilidade do provedor que não remove é solidária, salvo se ele demonstrar dúvida razoável quanto à ilicitude após análise diligente e qualificada; para os crimes contra a honra, mantém-se a ordem judicial como condição da responsabilidade, ressalvada expressamente a possibilidade de remoção por notificação extrajudicial; para um rol

Este é um artigo publicado em acesso aberto (Open Access) sob a licença CreativeCommons Attribution, que permite uso, distribuição e reprodução em qualquer meio, sem restrições desde que o trabalho original seja corretamente citado.

Ano VII, v.1 2026 | submissão: 29/10/2026 | aceito: 05/11/2026 | publicação: 10/11/2026

taxativo de ilícitos graves, impõe-se um dever de cuidado com monitoramento ativo, cuja falha sistêmica responsabiliza; para anúncios, impulsionamentos pagos e redes artificiais de distribuição, presume-se relativamente a culpa, afastável por diligência em tempo razoável. Ao lado da tese, dois decretos de 20 de maio de 2026, em vigor desde 20 de julho, compõem o quadro normativo, com objetos distintos: o Decreto 12.975 altera o Decreto 8.771/2016, que regulamenta o Marco Civil, e o Decreto 12.976 trata da proteção das mulheres no ambiente digital. O primeiro interessa diretamente às perguntas sobre o restritor: atribui à Agência Nacional de Proteção de Dados a fiscalização dos deveres dos provedores (Decreto 8.771/2016, art. 19-A, na redação do Decreto 12.975/2026) e reserva à Advocacia-Geral da União um poder de notificação estreito, limitado à publicidade enganosa, abusiva ou fraudulenta relacionada a políticas públicas (art. 16-N, § 1º). Sua validade é questionada no Supremo Tribunal (ADI 7981).

A esse regime, lido no texto do decreto que o regulamenta, as perguntas do critério recebem respostas parciais, e convém registrá-las antes de apontar o que falta. Quanto ao diagnóstico, o dever de cuidado não responde a publicações isoladas: o provedor responde por falha sistêmica na indisponibilização de um rol de condutas típicas — terrorismo, induzimento ao suicídio, incitação à discriminação, crimes contra a mulher, crimes sexuais contra vulneráveis, tráfico de pessoas e os tipos de incitação e de atentado ao Estado democrático (CP, art. 286, parágrafo único, e arts. 359-L, 359-M, 359-N, 359-P e 359-R) —; a existência isolada de conteúdo ilícito não caracteriza a falha; e é vedado à autoridade responsabilizar o provedor apenas pela manutenção ou remoção isolada de um conteúdo, bem como notificá-lo para moderar conteúdos isolados (Decreto 8.771/2016, arts. 16-B, § 3º, e 16-I, na redação do Decreto 12.975/2026). O regime separa, assim, os dois níveis antes distinguidos: o sistema é examinado pela autoridade; a decisão particular, pelo provedor, sob notificação. Quanto à justificação, a remoção por notificação deve ser comunicada ao usuário com o fundamento específico e os meios de contestá-la, a manutenção deve ser comunicada ao notificante nos mesmos termos, e a reconsideração após contestação deve ser fundamentada às duas partes (art. 16-E); os termos de uso devem prever o devido processo e relatórios anuais de transparência (art. 20-A). Nesse ponto o regime atende à exigência de razões dirigidas a quem a medida alcança. Quanto à competência, o provedor pode manter o conteúdo quando, após análise diligente e fundamentada, concluir haver dúvida razoável sobre seu caráter criminoso, considerada a proporcionalidade entre a dúvida e a gravidade do crime, e deve levar em conta o contexto e a finalidade informativa, crítica, satírica ou paródica da publicação (art. 16-G, §§ 1º e 2º). O grau

Este é um artigo publicado em acesso aberto (Open Access) sob a licença CreativeCommons Attribution, que permite uso, distribuição e reprodução em qualquer meio, sem restrições desde que o trabalho original seja corretamente citado.

Ano VII, v.1 2026 | submissão: 29/10/2026 | aceito: 05/11/2026 | publicação: 10/11/2026

de escrutínio cresce com o que está em jogo, como na responsabilidade epistêmica descrita por Amaya (2011, p. 316-318); e, como nela, a diligência demonstrada no processo serve para avaliar o decisor — o provedor —, enquanto a correção da decisão de manter ou remover permanece questão distinta (AMAYA, 2011, p. 318, nota 11). O provedor deve ainda coibir o uso abusivo das notificações contra a liberdade de expressão (art. 16-F), dever dirigido ao próprio restritor.

O que fica em aberto é mais estreito do que a pergunta inicial. O decreto não fixa os prazos de contestação; autoriza a autoridade competente a regulamentá-los (art. 16-D, parágrafo único), e a contestação em tempo útil depende, entre outras coisas, de que tais prazos existam e sejam curtos. As garantias do art. 16-E são desencadeadas pela notificação; o decreto não diz se alcançam a moderação que o provedor exerce por suas condições de uso, nem trata da redução de alcance, que permanece a medida menos visível e menos controlada. O controle administrativo da diligência é sistêmico, exercido pela autoridade de proteção de dados (art. 19-A). O controle da decisão particular é judicial e está previsto na própria tese: quem teve o conteúdo removido pode pedir seu restabelecimento demonstrando a ausência de ilicitude, sem que o provedor seja condenado a indenizar, e o provedor ou o autor da publicação podem requerer tutela provisória para impedir a retirada (STF, 2026, itens 5.5 e 5.6). A questão deixa de ser a existência de controle individual e passa a ser sua suficiência, sua acessibilidade e sua tempestividade. O restabelecimento sem indenização põe, além disso, um problema que toca diretamente o critério: restaurar a publicação não restaura a oportunidade de participação perdida enquanto ela esteve fora, e a mensagem devolvida depois do debate a que se destinava volta a um campo que já se reorganizou sem ela. Por fim, anúncios, impulsionamentos pagos e redes artificiais recebem presunção de responsabilidade afastável pela atuação diligente (art. 16-L), o que concentra a responsabilidade em quem amplifica; mas a solidariedade fixada na tese, que o voto vencido do ministro Mendonça apontou como fonte de efeito inibitório sobre as manifestações dos usuários (MIGALHAS, 2026), reorganiza a condição de todo falante. Se as garantias dos itens 5.5 e 5.6 bastam para justificar essa reorganização perante os falantes, ou se a ausência de reparação pela participação perdida deixa parte dela sem justificação, é a pergunta que a tese deixa ao critério; respondê-la exigiria reconstruir os fundamentos dos votos, o que aqui não se faz.

Este é um artigo publicado em acesso aberto (Open Access) sob a licença CreativeCommons Attribution, que permite uso, distribuição e reprodução em qualquer meio, sem restrições desde que o trabalho original seja corretamente citado.

Ano VII, v.1 2026 | submissão: 29/10/2026 | aceito: 05/11/2026 | publicação: 10/11/2026

8. Três regimes perguntados pelo mesmo critério

O regime brasileiro não é o único a ter redesenhado, na última década, a responsabilidade das plataformas. A comparação com dois outros — o Regulamento dos Serviços Digitais da União Europeia e o Online Safety Act britânico — serve a um propósito determinado: verificar se as perguntas do critério distinguem regimes, e se as lacunas encontradas no brasileiro são inevitáveis ou apenas não resolvidas.

No diagnóstico, os três convergem na separação entre o sistema e a decisão particular. O Regulamento europeu combina um mecanismo de notificação e ação, aplicável a cada conteúdo (art. 16), com o dever das plataformas de muito grande dimensão de avaliar e atenuar riscos sistêmicos, entre os quais os efeitos negativos sobre a liberdade de expressão, o discurso cívico e os processos eleitorais (arts. 34 e 35) (UNIÃO EUROPEIA, 2022). O Online Safety Act impõe deveres de avaliação de risco e de cuidado quanto ao conteúdo ilegal (ss. 9 e 10), supervisionados pela autoridade reguladora das comunicações por meio de códigos de prática (REINO UNIDO, 2023). O decreto brasileiro responsabiliza o provedor por falha sistêmica e veda a punição por decisões isoladas. Nos três, o regime é examinado como regime, e a decisão particular, como aplicação.

Na justificação dirigida ao atingido, a diferença é nítida. O Regulamento europeu exige exposição de motivos clara e específica a todo destinatário afetado por uma restrição — remoção, bloqueio de acesso, despromoção e outras restrições de visibilidade, suspensão de pagamentos, do serviço ou da conta —, tanto quando a decisão se funda na ilegalidade do conteúdo quanto quando se funda em sua incompatibilidade com as condições de uso (art. 17). Exige ainda que as plataformas apliquem suas condições de uso de modo diligente, objetivo e proporcionado, com a devida consideração pelos direitos fundamentais, entre eles a liberdade de expressão (art. 14, n. 4). O regime brasileiro exige razões na remoção por notificação e silencia sobre a redução de alcance e sobre a moderação por condições de uso. O que no Brasil aparece como ponto em aberto tem, na União Europeia, regra expressa.

No controle, a diferença está no intervalo entre o provedor e o juiz. O Regulamento europeu impõe um sistema interno de tratamento de reclamações, gratuito e acessível por pelo menos seis meses após a decisão (art. 20), e abre a via de órgãos extrajudiciais de resolução de litígios, certificados, sem prejuízo do recurso ao Judiciário (art. 21); as duas obrigações não alcançam, com as ressalvas ali previstas, as micro e pequenas empresas (art. 19). O prazo do art. 21 mostra, porém, que a via extrajudicial não resolve por si a perda de participação num

Este é um artigo publicado em acesso aberto (Open Access) sob a licença CreativeCommons Attribution, que permite uso, distribuição e reprodução em qualquer meio, sem restrições desde que o trabalho original seja corretamente citado.

Ano VII, v.1 2026 | submissão: 29/10/2026 | aceito: 05/11/2026 | publicação: 10/11/2026

debate datado: a decisão pode levar até noventa dias, prorrogáveis por outros noventa nos litígios altamente complexos. O critério alcança, assim, também o modelo europeu. O Online Safety Act exige procedimentos de reclamação que alcançam, entre outros, o usuário cujo conteúdo foi removido ou restringido (s. 21). O regime brasileiro oferece a contestação perante o próprio provedor e os remédios judiciais dos itens 5.5 e 5.6 da tese, mas nenhuma instância intermediária. É nesse intervalo que a contestação em tempo útil se decide: a via interna é decidida por quem restringiu, e a demora de qualquer via externa, judicial ou extrajudicial, conta contra a participação que se pretendia proteger. Se uma instância intermediária seria mais rápida no Brasil é questão empírica que a comparação normativa não resolve.

O Online Safety Act acrescenta um traço que os outros dois não têm na mesma forma. Os serviços classificados na Categoria 1, categoria jurídica definida pelo próprio regime, devem operar sistemas que levem em conta, nas decisões de moderação, a importância da livre expressão do conteúdo de importância democrática e do conteúdo jornalístico (ss. 17 e 19), e todos os provedores devem ter particular consideração pela liberdade de expressão ao cumprir seus deveres (s. 22). Lido por aproximação — interpretativa, não uma influência demonstrada —, o dispositivo trata a regra que organiza o discurso público como mais exigente do que a que organiza um espaço temático, o que se aproxima da distinção de Post. O decreto brasileiro aproxima-se disso ao mandar considerar a finalidade informativa, crítica, satírica ou paródica da publicação (art. 16-G, § 2º), mas o faz como critério de exame do conteúdo, não como dever de proteção de uma categoria.

A comparação não mostra que um regime seja justo e os outros não; o critério não julga regimes inteiros por uma nota só. Mostra três coisas mais precisas. As perguntas do critério distinguem os regimes: os três convergem no diagnóstico e divergem na justificação e no controle. Os pontos em aberto do regime brasileiro — razões para a redução de alcance e para a moderação por condições de uso, e uma instância de revisão independente do provedor — têm correspondentes normativos em outros ordenamentos; os mecanismos identificados ampliam as alternativas institucionais disponíveis para discussão, sem demonstrar que sua ausência no Brasil tenha sido escolha deliberada nem que bastariam aqui. E o preenchimento no texto não garante o funcionamento: se a exposição de motivos é lida, se a reclamação interna reverte decisões, se os órgãos extrajudiciais decidem a tempo, são perguntas empíricas que a comparação normativa não responde e que o critério apenas formula.

Este é um artigo publicado em acesso aberto (Open Access) sob a licença CreativeCommons Attribution, que permite uso, distribuição e reprodução em qualquer meio, sem restrições desde que o trabalho original seja corretamente citado.

Ano VII, v.1 2026 | submissão: 29/10/2026 | aceito: 05/11/2026 | publicação: 10/11/2026

9. Conclusão

A praça fica aberta, e o que a mantém aberta não é a ausência de resposta. A fala que reorganiza condições de que outros vivem — a garantia de que um grupo não será atacado pelo que é, a confiabilidade de um procedimento, o acesso ao exercício de uma posição — deve resposta a essas pessoas, pela posição que elas ocupam e não pela intenção de quem falou. Mas a resposta devida não se converte, por si, em autorização para restringir. Entre o atalho que passa do dano à contenção e o que passa da proteção à dispensa de justificação há um caminho mais longo, que ambos evitam.

O caminho tem três passos, e cada um pede argumento próprio. O diagnóstico identifica a conduta, a condição e as pessoas, e distingue a prova do que foi dito da prova do dano ou do risco. A justificação diz por que a alteração foi injustificada, pela negação de um status ou pela obstrução de um exercício, sem que o emprego de um rótulo, a falsidade ou a ofensa decidam a questão por si. A resposta é escolhida por comparação concreta com as alternativas — extensão, duração, reversibilidade, urgência, efeitos sobre terceiros — e exige de quem a adota um título e um controle. Os três trabalhos anteriores ofereceram elementos para os dois primeiros passos; o terceiro pedia a construção aqui tentada.

A intervenção que restringe a expressão reorganiza a praça tanto quanto a fala — mais, quando quem intervém pode impor decisões — e responde às mesmas perguntas: que condição alterou, perante quem se justifica, quem a controla. A simetria não está nas medidas, desiguais porque as posições o são; está nas perguntas. Retirar uma publicação com a prova preservada, as razões ditas e o recurso em tempo útil não fecha a praça; suprimir os meios de contestar a própria supressão a fecha. Lido por essas perguntas, o direito brasileiro aparece mais próximo do critério do que o debate que o cerca sugere — responde a atos, prevê a resposta antes da contenção, exige nome de quem fala, pune quem embaraça o voto — e, no regulamento do Marco Civil, já exige razões dirigidas ao usuário atingido, meios de contestação e escrutínio proporcional à gravidade do caso. O que permanece aberto é mais preciso: os prazos de contestação, que o decreto não fixa, a suficiência dos remédios judiciais diante da participação perdida, as garantias da moderação exercida por condições de uso, o controle da redução de alcance e o alcance da proteção eleitoral contra o engano praticado em massa. A comparação com os regimes europeu e britânico mostra que parte desses pontos já recebeu resposta normativa em outros ordenamentos, o que amplia as alternativas institucionais em discussão, sem demonstrar que bastariam no Brasil.

Este é um artigo publicado em acesso aberto (Open Access) sob a licença CreativeCommons Attribution, que permite uso, distribuição e reprodução em qualquer meio, sem restrições desde que o trabalho original seja corretamente citado.

Ano VII, v.1 2026 | submissão: 29/10/2026 | aceito: 05/11/2026 | publicação: 10/11/2026

O resultado é argumentativo e tem três partes. A proteção da expressão não dispensa a justificação de quem intervém, falando, nas condições de participação de outras pessoas: nesse sentido, a liberdade de expressão não é irrestrita, porque não é imunidade às razões que se devem a quem a fala alcança. Esse dever de justificação não autoriza, por si, nenhuma restrição: dever razões a alguém e estar autorizado a restringir alguém são relações diferentes. E a restrição, quando vem, exige fundamento material suficiente — seja a proteção contra um dano ou um risco injustificado, seja a organização justificável de um espaço com finalidade determinada —, comparação com as alternativas menos onerosas, competência de quem a impõe e controle efetivo em tempo útil; nenhuma dessas condições basta sem as outras. Ficaram de fora, por exigirem outro método e outra evidência, a demonstração de que determinada fala tenha causado determinado dano e a avaliação, decisão por decisão, da resposta institucional ao 8 de janeiro.

Referências

ALEXY, Robert. Teoria dos direitos fundamentais. Tradução de Virgílio Afonso da Silva. São Paulo: Malheiros, 2008.

AMAYA, Amalia. Legal Justification by Optimal Coherence. Ratio Juris, v. 24, n. 3, p. 304-329, 2011.

BENESCH, Susan. Dangerous Speech: A Proposal to Prevent Group Violence. Dangerous Speech Project, 2013.

BRASIL. Constituição da República Federativa do Brasil de 1988.

BRASIL. Decreto-Lei nº 2.848, de 7 de dezembro de 1940 (Código Penal), arts. 138 a 140, 286, 359-L a 359-N, 359-P e 359-R.

BRASIL. Lei nº 4.737, de 15 de julho de 1965 (Código Eleitoral), arts. 297 e 323.

BRASIL. Lei nº 7.716, de 5 de janeiro de 1989, arts. 2º-A e 20.

BRASIL. Decreto nº 8.771, de 11 de maio de 2016, na redação do Decreto nº 12.975, de 20 de maio de 2026.

BRASIL. Lei nº 12.965, de 23 de abril de 2014 (Marco Civil da Internet), arts. 19 e 21.

BRASIL. Lei nº 14.197, de 1º de setembro de 2021.

BRASIL. Lei nº 14.532, de 11 de janeiro de 2023.

Este é um artigo publicado em acesso aberto (Open Access) sob a licença CreativeCommons Attribution, que permite uso, distribuição e reprodução em qualquer meio, sem restrições desde que o trabalho original seja corretamente citado.

Ano VII, v.1 2026 | submissão: 29/10/2026 | aceito: 05/11/2026 | publicação: 10/11/2026

BRASIL. Decreto nº 12.975, de 20 de maio de 2026. Altera o Decreto nº 8.771, de 11 de maio de 2016, que regulamenta a Lei nº 12.965, de 23 de abril de 2014. BRASIL. Decreto nº 12.976, de 20 de maio de 2026. Dispõe sobre a proteção das mulheres no ambiente digital.

BRASIL. Supremo Tribunal Federal. ADPF 130. Rel. Min. Ayres Britto. Julgamento em 30 abr. 2009.

BRASIL. Supremo Tribunal Federal. ADPF 187. Rel. Min. Celso de Mello. Julgamento em 15 jun. 2011.

BRASIL. Supremo Tribunal Federal. ADI 7261. Rel. Min. Edson Fachin. Referendo da medida cautelar julgado em 26 out. 2022.

BRASIL. Supremo Tribunal Federal. HC 82.424. Rel. p/ acórdão Min. Maurício Corrêa. Julgamento em 17 set. 2003.

BRASIL. Supremo Tribunal Federal. RE 1.037.396 (Tema 987) e RE 1.057.258 (Tema 533). Rel. Min. Dias Toffoli e Min. Luiz Fux. Julgamento em 26 jun. 2025; embargos de declaração julgados em 17 jun. 2026. Informativo STF, n. 1222, 30 jun. 2026.

BRASIL. Supremo Tribunal Federal. ADI 7981. Em tramitação.

BRASIL. Tribunal Superior Eleitoral. Resolução nº 23.714, de 20 de outubro de 2022.

CONNOLLY v. R [2025] EWCA Crim 657. Court of Appeal (Criminal Division), Inglaterra e País de Gales.

DEUTSCHLAND. Bundesverfassungsgericht. BVerfGE 93, 266 (“Soldaten sind Mörder”). Decisão de 10 out. 1995.

DIETRICH, Pablo. O amanhã é possível? Rio de Janeiro: Autografia, 2026a.

DIETRICH, Pablo. Da visibilidade à competência presumida: desempenho digital e atribuição de aptidão política. Revista Tópicos, 2026b. DOI 10.70773/revistatopicos/792268195.

DIETRICH, Pablo. Quando a praça vira trincheira: mitos políticos e o enfraquecimento do debate democrático. 2026c. Inédito.

DIETRICH, Pablo. Doing Wrong in the Name of Good. Zenodo, 2026d. Preprint. DOI: 10.5281/zenodo.23744884.

DIETRICH, Pablo. O dano recíproco e a condição comum: Coase diante da pergunta sobre quem reorganiza. 2026e. Inédito.

DOUEK, Evelyn. Content Moderation as Systems Thinking. Harvard Law Review, v. 136, n. 2, p. 526-607, 2022.

FORST, Rainer. The Right to Justification: elements of a constructivist theory of justice. Translated by Jeffrey Flynn. New York: Columbia University Press, 2012.

Este é um artigo publicado em acesso aberto (Open Access) sob a licença CreativeCommons Attribution, que permite uso, distribuição e reprodução em qualquer meio, sem restrições desde que o trabalho original seja corretamente citado.

Ano VII, v.1 2026 | submissão: 29/10/2026 | aceito: 05/11/2026 | publicação: 10/11/2026

HABERMAS, Jürgen. Direito e democracia: entre facticidade e validade. Tradução de Flávio Beno Siebeneichler. Rio de Janeiro: Tempo Brasileiro, 1997. v. 1.

HEGEL, G. W. F. Princípios da filosofia do direito. Tradução de Orlando Vitorino. São Paulo: Martins Fontes, 1997.

HIS MAJESTY'S INSPECTORATE OF CONSTABULARY AND FIRE & RESCUE SERVICES (HMICFRS). An inspection of the police response to the public disorder in July and August 2024: Tranche 2. London, 7 maio 2025.

HOUSE OF COMMONS. Science, Innovation and Technology Committee. Social media, misinformation and harmful algorithms. Second Report of Session 2024–25, HC 441. London, 2025a.

HOUSE OF COMMONS. Home Affairs Committee. Police response to the 2024 summer disorder. Second Report of Session 2024–25, HC 381. London, 2025b.

KAUFMANN, Arthur. Filosofia do direito. Prefácio e tradução de António Ulisses Cortês. Lisboa: Fundação Calouste Gulbenkian, 2004.

MIGALHAS. Responsabilidade de redes: veja como votaram ministros do STF em embargos. Migalhas, 11 jun. 2026.

MILL, John Stuart. Sobre a liberdade. Tradução de Pedro Madeira. Lisboa: Edições 70, 2006.

POST, Robert. Participatory Democracy and Free Speech. Virginia Law Review, v. 97, n. 3, p. 477-489, 2011.

RAWLS, John. Two Concepts of Rules. The Philosophical Review, v. 64, n. 1, p. 3-32, 1955.

REILLY, Paul. Southport riots: why social media's role in unrest is overblown. The Conversation, University of Glasgow, 2024.

REINO UNIDO. Online Safety Act 2023, c. 50, ss. 9, 10, 17, 19, 21 e 22.

SHKLAR, Judith N. The Faces of Injustice. New Haven: Yale University Press, 1990.

SILVA, Virgílio Afonso da. O proporcional e o razoável. Revista dos Tribunais, São Paulo, v. 798, p. 23-50, 2002.

SOLUM, Lawrence B. Procedural Justice. Southern California Law Review, v. 78, p. 181-321, 2004.

TAVRIS, Carol; ARONSON, Elliot. Mistakes Were Made (but Not by Me): why we justify foolish beliefs, bad decisions, and hurtful acts. Orlando: Harcourt, 2007.

Este é um artigo publicado em acesso aberto (Open Access) sob a licença CreativeCommons Attribution, que permite uso, distribuição e reprodução em qualquer meio, sem restrições desde que o trabalho original seja corretamente citado.

Ano VII, v.1 2026 | submissão: 29/10/2026 | aceito: 05/11/2026 | publicação: 10/11/2026

UNIÃO EUROPEIA. Regulamento (UE) 2022/2065 do Parlamento Europeu e do Conselho, de 19 de outubro de 2022, relativo a um mercado único para os serviços digitais (Regulamento dos Serviços Digitais), arts. 14, 16, 17, 20, 21, 34 e 35.

WALDRON, Jeremy. The Harm in Hate Speech. Cambridge, MA: Harvard University Press, 2012.

Este é um artigo publicado em acesso aberto (Open Access) sob a licença CreativeCommons Attribution, que permite uso, distribuição e reprodução em qualquer meio, sem restrições desde que o trabalho original seja corretamente citado.