Rio de Janeiro, 2021

Rastreabilidade de mensagens instantâneas e vigilância em massa: uma análise crítica do Art. 10 do PL nº 2630/2020

ITS
Instituto
de Tecnologia
& Sociedade
do Rio

Resumo Executivo

O artigo 10 do Projeto de Lei nº 2.630/2020 (‘PL das Fake News”) prevê a obrigatoriedade da rastreabilidade de "mensagens veiculadas em encaminhamentos em massa" por parte dos serviços de mensageria privada como WhatsApp e Telegram. Neste relatório, respondemos perguntas introdutórias acerca do funcionamento técnico da rastreabilidade e analisamos os três problemas principais da medida:

- **Primeiro, a rastreabilidade é tecnicamente ineficiente**: os requerimentos técnicos levam à guarda de todas mensagens, inclusive as privadas, enfraquecendo a criptografia. Ainda assim, é ineficaz para identificar os responsáveis pelas campanhas de desinformação ou pela circulação de mensagens.

- **Segundo, a rastreabilidade implementa a obrigação de monitoramento em massa por parte das plataformas:** a proposta promove um estado de vigilância de todos os usuários de serviços de mensageria privada, sem exceção, pelo Estado e potencialmente pelas plataformas, podendo levar ao arrefecimento do debate público e à autocensura.

- **Terceiro, a rastreabilidade cria um paradoxo em relação ao Marco Civil da Internet e à Lei Geral de Proteção de Dados** ao promover um regime de *maximização* de coleta de dados pessoais, justo quando precisamos nos guiar pelo princípio da *minimização* da coleta e tratamento destes dados. Além disso, cria uma nova vulnerabilidade técnica que poderá levar ao vazamento de dados pessoais e abusos de seus usos

Conclui-se que o PL propõe uma solução jamais testada, podendo levar a violações de direitos fundamentais e contrária à lógica de proteção de dados, além de desconsiderar soluções menos gravosas, como a "escuta de metadados" centrada em indivíduos investigados e a adoção de mecanismos de "follow the money", associando o financiamento da campanha de desinformação aos seus mandantes.

SUMÁRIO:

Introdução

No dia 30 de junho de 2020, o Senado Federal aprovou o projeto de lei nº 2.630 de 2020 que, segundo sua própria ementa, "institui a Lei Brasileira de Liberdade, Responsabilidade e Transparência na Internet". Entre jornalistas e membros da sociedade civil, o projeto ficou conhecido como "PL das fake news", dando destaque ao seu objetivo de combater a desinformação no Brasil a partir de uma série de mudanças no esquema de governança de plataformas digitais. Atualmente, o projeto se encontra em debate na Câmara dos Deputados.

Neste relatório, o ITS Rio apresenta suas considerações sobre um dos pontos mais controversos do projeto, que já atraiu a atenção de diversos pesquisadores e estudiosos do tema: a rastreabilidade de mensagens instantâneas em serviços de mensageria privada como WhatsApp e Telegram.

Antes de se adentrar na análise em si, abaixo serão respondidas algumas perguntas introdutórias que pretendem nivelar o debate a respeito da rastreabilidade de mensagens na Internet, apresentando, assim, um FAQ a respeito do tema.

ENTENDA O QUE O PL PROPÕE E SUAS LIMITAÇÕES

- **O que propõe o PL?**

O PL tem como justificativa a proteção da “Liberdade, Responsabilidade e Transparência” e propõe-se a estabelecer uma série de medidas para aumentar a responsabilidade das plataformas no controle do conteúdo que circula na Internet e, assim, combater a desinformação. Dentre os diferentes dispositivos debatidos, o art. 10 propõe a rastreabilidade das mensagens encaminhadas em aplicativos de mensageria privada.

O projeto sugere que seja possível identificar: (a) quem originou determinado conteúdo que se tornou viral e (b) por quais usuários o conteúdo passou. Isso ocorreria por meio da obrigação de implementação, por parte das plataformas, da rastreabilidade de "mensagens veiculadas em encaminhamentos em massa" em serviços de mensageria privada como WhatsApp e Telegram, conforme texto do substitutivo mais recente:

> Art. 10. Os serviços de mensageria privada devem guardar os registros dos envios de mensagens veiculadas em encaminhamentos em massa, pelo prazo de 3 (três) meses, resguardada a privacidade do conteúdo das mensagens.

> §1º Considera-se encaminhamento em massa o envio de uma mesma mensagem por mais de cinco usuários, em intervalo de até 15 dias, para grupos de conversas, listas de transmissão ou mecanismos similares de agrupamento de múltiplos destinatários.
> §2º Os registros de que trata o caput devem conter a indicação dos usuários que realizaram encaminhamentos em massa da mensagem, com data e horário deste encaminhamento, e o quantitativo total de usuários que receberam a mensagem.
> §3º O acesso aos registros somente poderá ocorrer com o objetivo de responsabilização pelo encaminhamento em massa de conteúdo ilícito, para constituição de prova em investigação criminal e em instrução processual penal, mediante ordem judicial, nos termos da Seção IV do Capítulo III da Lei nº 12.965, de 23 de abril de 2014.
> §4º A obrigatoriedade de guarda prevista neste artigo não se aplica às mensagens que alcançarem quantitativo total inferior a mil usuários, devendo seus registros ser destruídos nos termos da Lei nº 13.709, de 14 de agosto de 2018.

Em outras palavras, as empresas precisariam guardar registros de mensagens que se tornaram "virais" em suas plataformas. Mas e o que é o "viral" para o PL? Segundo a redação do dispositivo, são aquelas mensagens que forem encaminhadas por mais de 5 (cinco) usuários e que, dentro de um intervalo de 15 (quinze) dias, alcançarem um total de 1.000 (mil) usuários devem ter seus registros mantidos pelas empresas por um prazo de até 3 (três) meses. O acesso aos registros ocorreria somente com “o objetivo de responsabilização pelo encaminhamento em massa de conteúdo ilícito”, mediante ordem judicial, visando constituir prova em investigação criminal ou instrução processual penal.

Como se vê, o legislador - sem comprovação de eficiência ou solução técnica possível - parte do pressuposto de que as "mensagens encaminhadas em massa" em aplicações como o WhatsApp e outros são parte significativa do fenômeno de disseminação de desinformação no Brasil. De forma a possibilitar a responsabilização, **o PL exige que sejam incorporadas soluções técnicas necessárias para tornar essas mensagens rastreáveis**, possibilitando, assim, a identificação dos atores envolvidos na cadeia de desinformação.

Vale destacar, desde o início, que propostas como a do PL nº 2.630 de rastreabilidade foram rejeitadas em outros países democráticos, tendo em vista que muitas vezes são apresentadas às pressas e buscam endereçar desafios de uma plataforma específica, violando o princípio da generalidade da lei. Ainda que se proponha um

redesenho de uma dada plataforma no Brasil, por exemplo, atores mal-intencionados podem sempre coordenar esforços para disseminar desinformação usando números estrangeiros, fora da jurisdição brasileira. Assim, na prática, a solução nacional específica é facilmente contornável e não surte os efeitos desejados.

- **Tecnicamente, o que o projeto requer?**

O objetivo é identificar o emissor original de um conteúdo de desinformação (o "viral ilícito"), como também encontrar por quais contas esse conteúdo circulou ("contas investigadas"). A situação concreta é que determinada investigação, ao apreender um celular com a mensagem ilícita, possa retraçar o caminho trilhado pela mensagem até chegar a sua origem.

Embora nenhum dos proponentes ou das pessoas que defenderam publicamente a proposta tenha indicado de que maneira ela poderia ser operacionalizada tecnicamente, deixando a cargo dos serviços de mensageria traduzir em soluções técnicas o que lhes é exigido, anotamos abaixo algumas das obrigações que terão de ser cumpridas com base no que propõe o artigo 10 do PL nº 2.630/2020.

(1) **Primeiro, obriga a coleta de todas as trocas de mensagens, independentemente de sua natureza e do contexto em que ela foi enviada.**

O "viral ilícito" é determinado sempre a posteriori. Sendo assim, **todas** as trocas de mensagens precisam **sempre** ser monitoradas, seja na comunicação entre pares ou nos grupos, pois apenas olhando para trás podemos cumprir com os requisitos técnicos que a lei introduz. Em outras palavras, o mecanismo de vigilância imposto pelo PL não se aplica desde quando algo se torna viral; ele está sempre latente. Princípios como a presunção de inocência e devido processo legal serão flexibilizados para uma quantidade infinita de pessoas, e abusos serão difíceis de serem monitorados.

(2) **Segundo, obriga que toda comunicação tenha o registro do seu criador associado à própria mensagem**.

Para que o emissor original possa ser identificado, é preciso que toda comunicação que se der a partir dele seja guardada, de forma que você se torne vinculado a todos aqueles que fizerem uso de sua mensagem, sem qualquer controle sobre isso. Além disso, é necessário que seja gravado, na mensagem em si, algo que identifique o emissor que enviou a mensagem para o grupo (por mais que não seja ele, necessariamente, quem criou

o conteúdo) Como nestes serviços as mensagens ficam guardadas nos terminais, e não em um intermediário centralizador (um banco de dados central da plataforma), todas as mensagens têm que conter uma espécie de "DNA" de quem as originou, uma espécie de “identidade”, possivelmente o número do chip.

(3) Terceiro, obriga que o intermediário guarde cópia de todas as mensagens encaminhadas pelo usuário.

A plataforma tem acesso a algumas informações quando intermedia a troca de mensagens, como um serviço postal que tem acesso aos dados de endereço de quem envia e recebe uma mensagem. Por princípios técnicos de criptografia ponta-a-ponta, essas informações, assim que usadas para seu fim, são excluídas na sequência. Mesmo os arquivos enviados para grupos, assim que cumprem seu papel, são excluídos assim que possível.

O que o PL obriga é que as informações sejam guardadas preventivamente pelos provedores de aplicação de maneira centralizada em seus servidores (e não de maneira distribuída nos dispositivos terminais dos usuários) e excluídas passados 3 meses. Esse é um grande risco para a criptografia, pois cria um ponto de ataque centralizado para toda rede do provedor de serviços.

Sem as três operações técnicas acima o PL não poderá surtir efeitos. A seguir identificamos porque esses elementos enfraquecem ou, até mesmo, quebram a criptografia ponta-a-ponta.

F.A.Q. DA RASTREABILIDADE

- **O que é a criptografia ponta-a-ponta (sistema E2EE)?**

O sistema E2EE é pensado para gerar um maior grau de privacidade e segurança. É dizer, trata-se de um conjunto de soluções matemáticas com implementações técnicas variadas que garantem a privacidade de seus usuários. A mais comum dessas tecnologias usadas por aplicativos de mensageria é o Protocolo Signal.

O maior perigo associado à quebra ou ao enfraquecimento da criptografia é o acesso ao conteúdo das mensagens trocadas. Como padrão, na E2EE apenas quem envia e recebe as informações têm acesso ao conteúdo. Há a possibilidade de que, uma vez quebrada ou diminuída a proteção oferecida pela criptografia, a própria empresa provedora

do serviço ou mesmo terceiros não autorizados tenham acesso ao conteúdo de mensagens privadas.

O perigo ao E2EE, contudo, não se resume apenas à visualização de mensagens confidenciais, mas diz respeito, também, ao enfraquecimento do conjunto de soluções técnicas que fazem o protocolo funcionar (com os efeitos negativos em cascata que podem ser desencadeados para o nível de confiança nos produtos e serviços da economia digital). Na prática, o risco de vazamento de mensagens é criado quando violamos uma ou mais especificações técnicas do protocolo. É por isso que a introdução da rastreabilidade, por si só, já representa uma potencial vulnerabilização do protocolo (embora em um primeiro momento as mensagens não sejam, necessariamente, reveladas). Dizer que a confidencialidade da mensagem estará segura na mudança proposta pelo PL é uma ingenuidade. É o mesmo que nos anos 1990s imprimir uma agenda telefônica dizendo que a privacidade está garantida. Há na afirmação uma miopia do que a privacidade no futuro próximo será, como uma ingenuidade do que a tecnologia é capaz de produzir com uma base de dados de hash se tornada pública.

- **É possível guardar só metadados, sem registrar o conteúdo da mensagem?**

Da forma como o PL estabelece, não. Toda guarda de dados que o PL ordena deve fazer relação com o conteúdo das mensagens. De forma engenhosa, o PL fala em guarda de metadados (e não de conteúdo). Mas não há alternativa - pelo que o PL propõem - desses metadados não terem uma relação direta e deduzível do conteúdo associado ao metadados.

Metadados são representações codificadas da mensagem em si, que ajudam a circular a mensagem e possibilitam o processo de transmissão. A forma mais clara de implementar o que se pede no PL é criar um metadado único para cada conteúdo diferente. Esse conteúdo (tecnicamente, um *hash*) permite olhar para uma rede e identificar por onde um conteúdo exatamente igual circulou.

Nos protocolos mais fortes de E2EE, contudo, a geração de um metadado único para um conteúdo é tecnicamente negada por um motivo simples: se há relação entre o conteúdo e um dado potencialmente público (como aqueles que circulam no cabeçalho das mensagens), atores que transmitem a mensagem encriptada podem deduzir o que está sendo dito, quebrando a criptografia.

Ou seja: o metadado gerado, pela redação do PL, teria que ser associado diretamente a todos os conteúdos iguais a ele, em contextos ou emissores diferentes, e teria que ser guardado em espaço público e privado para consulta futura. Na prática, todos os conteúdos podem ser acessíveis desde que se tenha um original a se buscar, dando

superpoderes a quem queira saber todos os telefones de quem compartilhou uma mesma mensagem, seja ela grave e ilícita, ou não.

Além disso, há conteúdos que são iguais mas dizem respeito a contextos diferentes. Uma mensagem de "bom dia", por exemplo, pode ser usada em diversas situações. O mesmo pode ser dito a respeito dos emojis, expressões comuns ou imagens. Igualmente, conteúdos distintos de um ponto de vista do código binário falado no nível do hardware (e.g.: bom dia -- b0m d14) pode ser exatamente idêntico de um ponto de vista semântico. Logo, saber todos lugares onde uma mesma mensagem circulou não indica se os contextos são os mesmos. Igualmente, mensagens absolutamente distintas de um ponto de vista estrutural podem circular transmitindo exatamente a mesma informação.

O PL, contudo, segue no caminho contrário ao requerer que haja uma relação 100% confiável entre o conteúdo e seu *hash* correspondente, além, é claro, de obrigar o seu respectivo registro pela plataforma. Essa fragilidade já havia sido excluída do Protocolo Signal e dos demais protocolos de criptografia ponta-a-ponta, mas o PL força o Brasil a reintroduzi-la.

- **É possível guardar dados apenas de mensagens virais, em grupos, sem registrar as mensagens privadas?**

Da forma como o PL exige, não. Todas as mensagens que circulam em grupos de 2 ou mais pessoas serão monitoradas (por no mínimo 15 dias). No cenário mais privado que podemos imaginar, em que há troca de mensagens apenas entre duas contas, sem nunca compartilhar dados vindos de grupos, seria possível desenhar uma solução privada, na qual se geraria o dado de identificação de "conteúdo+autor" para cada troca de mensagem, mas tecnicamente essa informação nunca seria enviada para o intermediário. Esse é um cenário de exceção, e que não representa como aplicativos de mensagem são usados na prática.

A realidade é que a troca e reuso de mensagens de grupo em conversas privadas ou pública acontece a todo momento, fazendo da regra de guarda de dados o padrão. Ou seja: se a cada mensagem enviada para grupos, seja texto ou multimídia, é associada ao "DNA" do usuário que a encaminhou, acompanhada da geração de um hash único, quem envia uma mensagem para um grupo ficaria, daquele ponto em diante, responsável por onde a mensagem circulará na sequência. Você pode criar um conteúdo, que se viralizar independentemente de você, o tornará responsável pela ação de terceiros desconhecidos.

Além disso, uma mensagem recebida em privado e, posteriormente, compartilhada por um segundo usuário para um determinado grupo terá este como seu novo autor, reiniciando toda cadeia de rastreabilidade.

- **Onde ficarão guardados os dados gerados? É possível excluí-los centralmente?**

O PL exige que a plataforma guarde todos os dados necessários para rastrear a mensagem e identificar o usuário inicial. Isso implica que o serviço terá que guardar os dados como faz hoje em relação ao IP para o terminal de acesso.

Contudo, no sistema de mensageria privada, o intermediário é visto como um ponto frágil do sistema, e, por isso, todas informações ficam armazenadas nos próprios terminais. Até porque sistemas descentralizados tendem a ter maior segurança e resiliência, pois a informação disponibilizada em caso de ataques é menor do que ocorreria em um sistema centralizado com banco de dados único. A relação custo-benefício, sendo maior o custo e menor o benefício, favorece a segurança de cadeias descentralizadas.

Na prática, toda informação será guardada de forma duplicada: nos terminais e no intermediário, aumentando a superfície de ataque do sistema e incrementando exponencialmente sua vulnerabilidade. Excluir os dados do intermediário é um procedimento comum. Já apagar informação dos terminais ou quebrar a criptografia é uma solução incomum que pode, também, ser insegura. A criptografia forte impede, no seu design, que o intermediário controle os terminais, e as mensagens geradas são individualizadas, para controle, também por uma opção de design do sistema.

Na prática, dados gerados e guardados nas mensagens nos terminais não podem ser apagados. Logo, a previsão do PL de que dados guardados devem ser armazenados por 3 meses não é implementável: os dados gerados, na prática, ficarão armazenados para sempre nos terminais, mesmo que a cópia deles seja (na melhor das hipóteses) excluída do intermediário central.

- **Os dados guardados não são exatamente os mesmos que são registrados hoje? Há obrigação de coletar novos dados?**

Como regra, o Protocolo Signal não inclui "dentro" das mensagens o dado de quem a gerou. Também não são guardados, como padrão, os dados de "quem enviou a mensagem para quem". Ambos os processos geram rastreabilidade, o que atualmente é combatido pelo Protocolo.

A rastreabilidade enfraquece a criptografia de duas formas. Primeiro, ao criar um banco centralizado de mensagens ou metadados, o que torna possível treinar algoritmos para deduzir o processo criptográfico. A facilidade é similar a ter a correspondência entre algumas senhas e seus hashes, e deduzir o algoritmo que criptografa em si. Logo, a associação entre dois tipos de dados (os dados de quem fala com quem, associado às

mensagens produzidas) cria uma forma nova de armazenamento, produzindo, na prática, um dado novo que o ecossistema antes não tinha.

Nesse sentido, há uma mudança na forma com que hash e mensagem se equivalem, o que de fato produz dados novos sobre o usuário inicial de uma mensagem. A relação entre um usuário e determinado hash torna-se uma informação nova passível de apropriada para tratamento posterior.

Ou seja, é um equívoco dizer que o PL apenas obriga a guardar dados que já existem. O PL obriga a coleta de *novos dados*, e obriga inclusive criar novos usos de dados que já existem, enfraquecendo a criptografia de ponta-a-ponta.

- **Os dados guardados não são parecidos com aqueles previstos pelo Marco Civil da Internet?**

O Marco Civil da Internet (MCI) introduziu a obrigação de guarda de dados, em especial os de registro. Hoje, toda vez que um terminal acessa uma aplicação, guarda-se o IP do terminal que o fez, combinado com a data e a hora de início e fim desse acesso. Em linhas gerais, trata-se da guarda do IP mais provável de identificar de onde veio a ação para se conectar a um terminal, e em qual dia e horário.

O PL introduz um novo tipo de registro: a guarda de dados referentes ao conteúdo. Em vez de registrar quando um usuário acessa o aplicativo (obrigação prevista no MCI), passa-se a guardar dados relacionados ao compartilhamento de um determinado conteúdo. Isto é, introduz-se a guarda não do registro de acesso, mas sim da interação com um determinado conteúdo, guardando a cadeia de encaminhamento de mensagens.

- **Como funcionam as cadeias de encaminhamento de mensagens?**

No caso de encaminhamento direto de conteúdo, quando um usuário de serviço de mensageria privada na Internet recebe uma mensagem e a encaminha diretamente para outros contatos usando uma ferramenta de compartilhamento disponibilizada pela plataforma, os desafios são menores. Neste caso, ao menos em tese, a cadeia de encaminhamentos poderia estar intacta.

No entanto, há que se entender a enorme complexidade da questão. Cadeias de trocas de mensagem não funcionam simplesmente de maneira linear, em que A envia para B, que envia para C e assim por diante. Ou seja, as cadeias de encaminhamento são plúrimas e nem sempre permitem a verificação do autor original da mensagem.

Todavia, nem mesmo a analogia de uma "árvore de encaminhamentos", ainda que elucidativa e comumente evocada pelos aplicativos de mensageria privada, dá conta de

toda a complexidade da questão. Isso porque, no caso de uma árvore, há o pressuposto de que os galhos sempre seguem adiante. No caso de cadeias de mensagens, entretanto, A, B ou C podem estar em diferentes posições e encaminhar e reencaminhar mensagens diversas vezes, incluindo para pessoas que já a receberam anteriormente. Incluem-se, aqui, os grupos, o que apenas potencializa a questão.

Assim, a melhor forma de visualizar a questão pode ser um prato de espaguete - como se vê na imagem abaixo -, no qual os fios (cadeias de encaminhamento de mensagens) se entrelaçam de inúmeras formas. Encontrar a origem de uma mensagem e entender quem foi o maior contribuidor nessa cadeia pode ser extremamente complexo. Mesmo que fosse possível, resta ainda a pergunta se efetivamente saber a origem é relevante para combater a desinformação nos serviços de mensageria.

No relatório "10 mitos e verdades sobre a rastreabilidade", explica-se, por exemplo que a cadeia de encaminhamento de mensagens no aplicativo pode ser vista "como uma árvore com muitos galhos: ao observarmos somente um deles, não sabemos quantos existem no total ou qual a origem dos demais galhos". Igualmente, há uma falsa impressão de que, uma vez implementada a rastreabilidade, as autoridades terão em mãos um único desses “galhos” ou um desses “fios de espaguete” que leva do último a receber uma mensagem até o primeiro a enviá-la.

Na realidade, conforme novas cadeias de encaminhamento são geradas e se valendo da mesma analogia acima, este fio é recortado e embaralhado diversas vezes. Assim, as autoridades terão em mãos, quando muito, apenas uma pequena seção - um “pedaço do espaguete” - sem saber onde estão as demais seções ou mesmo sem saber se

existem outras ou quão relevante é a que se tem ante às outras. Consequentemente, será muito difícil saber, com total certeza, quem deu origem ao fio de espaguete num primeiro momento.

- **Quais são, então, os principais problemas da proposta?**

Não obstante as boas intenções do legislador, no presente relatório serão destacados os três principais motivos pelos quais tal proposta legislativa deve ser considerada temerária e, se aprovada, **poderá causar danos irreparáveis à liberdade de expressão e à privacidade dos usuários de serviços de mensageria privada no Brasil e no mundo.**

Vale ressaltar, desde já, que parte do objetivo deste documento é organizar e sistematizar os argumentos contrários ao art. 10 (rastreabilidade) do PL 2.630/2020 que já foram levantados e articulados por outros atores e instituições desde que o projeto foi apresentado ao Senado Federal em 2020.

Nesse sentido, existem três pontos críticos que devem ser levados em consideração e que serão analisado em maior detalhe neste documento: (i) a eficiência do dispositivo proposto para atacar os problemas de desinformação a que o projeto de lei se propõe resolver; (ii) como a proposta contradiz algumas das promessas dos protocolos de criptografia, que garantem os direitos dos usuários, desde a liberdade de expressão até a privacidade, sem contar inúmeros outros direitos civis e políticos; e (iii) o paradoxo de obrigar aos serviços de mensageria privada de proteger maior aos dados pessoais dos seus usuários, ao mesmo tempo que obriga a criação de mecanismos de coleta massiva e permanente de dados (inclusive pessoais).

1. Ineficiência Técnica da Solução Proposta

A rastreabilidade do Art. 10 do PL 2.630/2020, além de desproporcional, é ineficiente. Ela se vale de instrumentos que não são capazes de atingir tal objetivo, parte de premissas técnicas confusas ou impraticáveis, mira em objetivos que não são capazes de evitar a propagação da desinformação e, ainda, cria novos problemas para a promoção da liberdade de expressão e para a proteção de dados dos usuários de serviços de mensageria privada na Internet.

A rastreabilidade, mesmo antes de qualquer questão de ordem ética ou legal, apresenta um sério problema técnico: é ineficiente para atingir os resultados que almeja. Nas palavras de Canabarro e Rená, o artigo 10 "é ineficaz ao que se propõe e soluciona problemas que não existem, ao preço de criar vários novos problemas graves".

1.1 A rastreabilidade será útil? Poderemos encontrar quem enviou uma mensagem?

Como visto acima, existem, ao menos em tese, duas maneiras complementares de implementar a rastreabilidade exigida pelo PL em uma aplicação protegida pela criptografia ponta-a-ponta:

a) A primeira é através da coleta de metadados relacionados ao encaminhamento da mensagem. Ou seja, o aplicativo poderia coletar dados como "quem mandou a mensagem para quem", "em qual horário a mensagem foi encaminhada", "quando a mensagem foi lida", dentre outros.
b) A segunda, por sua vez, é através da atribuição de um código identificador, conhecido como *hash*, para cada conteúdo específico (ou mesmo para cada usuário, como uma espécie de “assinatura digital” que seguirá todas suas mensagens). Esses códigos poderiam ser monitorados para descobrir de onde a mensagem veio originalmente.

As propostas técnicas devem, concomitantemente, encontrar os responsáveis por mensagens que causem danos e permitir a desestruturação de organizações que criam campanhas de desinformação, ou seja, os atores e as práticas que o projeto de lei se destina a combater. Nesse sentido, a rastreabilidade deveria criar um mecanismo preciso para verificar a origem e a cadeia de mensagens circuladas de maneira massiva.

O problema, no entanto, é que essa precisão é facilmente questionada. Afinal, o modo como os serviços de mensageria funciona permite a realização de três formas de composição de conteúdo: a) a colocação de conteúdos próprios; b) o reencaminhamento de conteúdos de terceiros; e c) a adaptação de conteúdos de terceiros.

1.2 A rastreabilidade pode ser maquiada? É possível alguém incluir pistas falsas ou se esconder no caminho?

Assim, o problema do rastreamento das mensagens é óbvio: a integridade da cadeia pode ser facilmente quebrada. Se uma vírgula é alterada, ou mesmo a mensagem é copiada e colada - uma função que, vale ressaltar, é facilmente automatizada -, a medida já perde seu efeito. Afinal, cria-se uma nova cadeia de encaminhamento. Existem, inclusive, sistemas automatizados que simulam a digitação de textos, podendo, assim, burlar o mecanismo de rastreabilidade proposto pelo artigo 10.

Outra estratégia simples que poderia quebrar a cadeia de encaminhamento seria o envio da mensagem para um número estrangeiro. A questão aqui seria se este número está dentro ou fora da jurisdição brasileira. Ou seja, se a lei é ou não aplicável a esse contato estrangeiro. Entretanto, o projeto de lei não é claro quanto ao seu escopo. Se a lei não fosse aplicável, a inserção de um número estrangeiro poderia cortar a cadeia. Caso fosse aplicável, poderiam existir impactos globais da lei brasileira, gerando potenciais conflitos legais.

Ainda, há outra questão importante: existe a possibilidade de disparo, em massa e em paralelo, por serviços especializados, para espalhar uma mesma mensagem a partir de diversos números no exato mesmo momento. Essa via da desinformação não seria impossibilitada pela rastreabilidade.

Em suma, é muito difícil singularizar com segurança e confiabilidade uma mensagem com base em um código identificador, já que a "identidade" (e integridade) do texto pode ser facilmente alterada. Ainda, novas cadeias de encaminhamento de desinformação são fáceis de serem criadas por quem tem esse objetivo, tornando a rastreabilidade uma medida cuja eficiência será constantemente questionada.

De toda forma, mesmo que fosse possível chegar até a pessoa que realizou o primeiro encaminhamento - o suposto "paciente zero" da desinformação ou "desinformante zero" -, nada garante que ela seja a verdadeira responsável por aquele conteúdo. Primeiro, porque campanhas de desinformação possuem uma natureza multiplataforma: a responsável pode apenas ter copiado aquela mensagem de um site de notícias ou de outra rede social, compartilhando o mesmo conteúdo em um aplicativo de mensageria privada.

Segundo, quem encaminhou a mensagem pela primeira vez não necessariamente é a "mente" por trás de sua elaboração. Ou seja, ainda que se chegue à ponta do espaguete, nada garante que aquele usuário seja de fato o responsável pela confecção da mensagem. Como explica a nota técnica conjunta do IP.rec e do Coding Rights, "mesmo caso esse rastreamento tenha a capacidade de, de fato, chegar no primeiro usuário que enviou o conteúdo investigado (algo bastante questionável), muito provavelmente não será este primeiro usuário o autor daquela mensagem, o que torna a medida ineficaz".

Além disso, há que se considerar a mudança de contexto de determinado encaminhamento e que poderá gerar o enquadramento de pessoas inocentes de maneira involuntária. Exemplo: o jornalista que encaminha uma mensagem com uma denúncia para a apreciação de um grupo que contém o corpo editorial do veículo para o qual trabalha e, por engano, encaminha em paralelo a mesma para um grupo de amigos. Desse último, a mensagem se espalha e ganha viralidade. Independentemente do contexto, o jornalista em questão será enquadrado como o responsável por espalhar desinformação caso a denúncia não seja corroborada (e seja considerada desinformação).

Em suma, a rastreabilidade, além de impor obrigações excessivas às empresas e ser um grande risco para a liberdade e segurança dos usuários - como será reforçado nos tópicos a seguir -, também fracassa no próprio objetivo que diz almejar. Trata-se, na realidade, de um sistema fácil de ser burlado por aqueles que assim desejarem e ineficaz para combater a desinformação. Ainda pior, o artigo 10 do PL 2.630/2020, por sugerir uma solução ineficiente para o desafio da desinformação, pode criar uma falsa sensação de que outras soluções mais estruturais não são necessárias.

Há ainda um complicador a mais. Como bem notaram os participantes de um workshop sobre criptografia organizado pela Internet Society, atribuições digitais pelo uso de metadados ou códigos identificadores não são absolutas.

É possível, assim, que um usuário inocente seja implicado na prática de um ato ilegal com base na modificação de metadados ou de *hashes* por criminosos cibernéticos. Nas palavras do relatório que organizou as conclusões do workshop, "é difícil vincular um usuário a uma determinada mensagem quando a impersonificação online é tão fácil e universal".

Caso o PL seja aprovado, pode ser possível, na melhor das hipóteses, que usuários mal-intencionados escapem de eventual punição e ainda, na pior das hipóteses, incriminem pessoas inocentes.

1.3 A rastreabilidade permite diferenciar claramente o usuário comum da campanha de desinformação?

Vale salientar que o encaminhamento de uma mensagem *popular* - que “viraliza” - não significa, necessariamente, que a pessoa concorda ou defende aquele conteúdo; inclusive, pode tê-lo compartilhado para criticá-lo, para realizar pesquisa, como base de uma inquirição jornalística, para alertar um colega sobre a existência daquela mensagem ou mesmo com fins humorísticos (de sátira). Pessoas nessas circunstâncias também seriam inseridas na cadeia de mensagens da mesma forma. A diferenciação desses casos dependerá de uma compreensão de contexto que pode não estar disponível.

A rastreabilidade não permite a distinção clara entre um ator que potencialmente cometeu um ilícito e um segundo ator que agiu de boa-fé com o intuito de tão somente alertar seus contatos. Como explica Joana Varon, "seria como se tudo que você mandasse nos grupos passasse a ter um selo te identificando - assim, por via das dúvidas, por precaução".

Além disso, não permite diferenciar entre o usuário comum que compartilhou a mensagem - talvez de forma incauta, sem checar os fatos, mas sem intenção de cometer

um crime - e as verdadeiras máquinas de desinformação, que trabalhavam de forma profissional, organizada e financiada.

No entanto, encontrar uma mensagem não necessariamente permite compreender a mensagem que vem logo após a mesma, ou a que vem antes. Ou seja, a compreensão do que se trata a mensagem, indicativo necessário para uma alcunha de legalidade ou ilegalidade não necessariamente estarão disponíveis.

2. Risco da Vigilância em Massa de Interações Privadas

Um segundo ponto que merece destaque neste relatório diz respeito ao risco de monitoramento das interações privadas de usuários de serviços de mensageria privada na Internet. Como se demonstrará, a rastreabilidade prevista pelo artigo 10 do PL nº 2.630/2020 permite - e, em certa medida, obriga - a vigilância em massa de mensagens instantâneas. Ainda que não haja, declaradamente, total quebra da criptografia de ponta a ponta - pois, ao menos em tese, o aplicativo não teria acesso ao conteúdo em si, mas apenas aos metadados -, há um enfraquecimento dos protocolos criptográficos, abrindo espaço para o uso da informação de maneira extensiva.

2.1 Necessidade de rastreamento de todas as mensagens

Um problema sério, que reforça o caráter de vigilância massiva da proposta, reside na necessidade de acompanhamento contínuo de virtualmente todas as mensagens. Dois elementos são cruciais para entender a questão. Primeiro, como visto acima, não há como prever qual mensagem será viral, ou seja, aquelas que, como define o PL, são encaminhadas por mais de cinco pessoas e, dentro de um período de quinze dias, alcançam outras mil.

Seria necessário, então, que os provedores coletassem metadados de absolutamente todas as mensagens de todos os usuários por no mínimo quinze dias. Afinal, uma mensagem enviada no dia um pode viralizar no dia quinze, sendo necessário o registro de toda a cadeia para cumprir com a determinação legal. Na prática, portanto, **cria-se uma gigantesca base de dados com metadados das interações interpessoais de todos aqueles que utilizam um determinado serviço de mensageria privada na Internet.**

Em entrevista para o InternetLab, Riana Pfefferkorn, diretora associada de Cibersegurança e Vigilância no Centro para Internet e Sociedade da Universidade de Stanford, explicou:

> [...] Cada mensagem encaminhada precisaria ser retida por um período de pelo menos 15 dias, só por via das dúvidas, para o caso de que fosse enviada por cinco usuários nesse período. [...] De que forma o serviço sabe, no dia 1, se a mensagem chegará a 1001 ou a 999 usuários até o dia 15? Por isso, me parece que o provedor precisaria manter todos os metadados das mensagens.

Um segundo elemento se torna relevante quando se discute o ciclo de encaminhamento das mensagens. Do modo como se estruturou o artigo, há duas formas de

compreender o prazo de 15 dias que uma mensagem teria para alcançar os critérios estabelecidos na lei: (i) ser enviada por 5 usuários e (ii) atingir pelo menos mil pessoas. A primeira forma parte da ideia de que o início da contagem dos 15 dias seria o dia do "primeiro" envio de uma determinada mensagem. Já a segunda forma demanda que seja contado de trás para frente - sempre que, dentro de 15 dias, uma determinada mensagem for encaminhada por pelo menos cinco pessoas, atingindo no mínimo mil pessoas, ela deve ser rastreada.

Se estivermos falando da primeira interpretação, corre-se o risco de as plataformas não terem obrigação de guardarem os dados de mensagens que somente no dia 16 alcancem os requisitos necessários (compartilhados por 5 usuários atingindo mil pessoas). Isso criaria uma lacuna que poderia ocorrer de maneira orgânica ou ser explorada de maneira intencional (lembrando das possibilidades técnicas, legais ou não, de automação do envio de mensagens).

No caso da segunda interpretação, aí, a questão se relacionaria com a efetividade da medida discutida da parte anterior desse projeto. Se toda mensagem que em 15 dias fosse repassada por 5 usuários até atingir pelo menos mil pessoas, corre-se o risco de que parte da cadeia (tudo o que ocorrer antes dos 15 dias) não tenha sido guardado pela plataforma, haja vista que, de acordo com a lei, não estaria obrigada a guardar. Em tese, toda essa parte da cadeia (antes dos 15 dias) poderia não estar disponível, e a possibilidade de alcançar a origem da mensagem, o "paciente zero", poderia ter se perdido.

Nesse sentido, **tanto os requisitos quantitativos (encaminhamento por pelo menos 5 pessoas, atingindo 1000 indivíduos) e o temporal (15 dias) não limitam efetivamente a quantidade de dados a serem guardados pelos serviços de mensageria**. Para dar vazão à obrigação, há necessidade de criação de uma base de dados massiva e contínua.

2.2 O conteúdo enviado pode vir a ser identificado

Além da existência de um risco real de quebra do protocolo de criptografia ponta-a-ponta, há o risco de que o conteúdo das mensagens - que hoje é protegido pela tecnologia de criptografia - seja exposto. Pfefferkorn alerta:

> [...] A exigência de retenção de metadados poderia levar ao fim da criptografia de ponta a ponta. [...] De que maneira o provedor poderia saber que as mensagens do dia 1 e do dia 14 são a mesma, a não ser que conseguisse ler o conteúdo? Logo, a única forma de saber qual mensagem pode ser considerada "de encaminhamento em

massa" é vendo os conteúdos das mensagens, o que é incompatível com a criptografia de ponta a ponta.

Embora se possa questionar o nível de acesso ao conteúdo (o que a autora chama de “visão”), certo é que, uma vez instaurada a obrigação de rastreabilidade, haverá pelo menos o enfraquecimento da criptografia nestes aplicativos.

Como explica Matthew Green, especialista em criptografia da Johns Hopkins University, "no momento em que você cria um sistema que pode voltar no tempo para desmascarar algumas pessoas que encaminharam um conteúdo, você criou um sistema que pode desmascarar qualquer pessoa que encaminha qualquer conteúdo". O que o autor está dizendo é que, se é criado um sistema que permita a partir de um determinado conteúdo encontrar quem o encaminhou, qualquer conteúdo fica a mercê desse mesmo sistema.. Ou seja, a lógica criptográfica de proteção dos conteúdos é fragilizada.

Adicionalmente, o relatório da Internet Society explicita que os riscos de exposição existem não só frente ao governo, uma vez que garantir o acesso para um é abrir espaço para a mesma fragilidade ser explorada por todos. Isso significa que, ao estipular que autoridades investigativas e judiciais podem ter acesso ao mecanismo que permite a rastreabilidade, outros podem ter acesso ao mesmo sistema. Ou seja, "uma vez encontrados por atores mal-intencionados, os mesmos métodos usados para garantir o acesso pelas autoridades ou pelas próprias plataformas poderiam ser explorados para práticas nefastas". É dizer, não existe meio de garantir que as vulnerabilidades criadas para acessos "excepcionais" não sejam exploradas também por criminosos.

2.3 Noção ilusória de que coleta apenas de metadados não viola a intimidade e a privacidade

Há que se entender que a criação desses bancos de dados de informações relativas a encaminhamentos de mensagens que permitam a rastreabilidade podem servir para mais do que meramente identificar as pessoas na cadeia de encaminhamentos. Um banco de dados tão vasto pode servir para inúmeras funções - particularmente porque os serviços de mensageria não teriam acesso apenas a uma cadeia de mensagens, mas sim a um todo de mensagens.

Para atingir o objetivo de rastreabilidade, cria-se um vínculo direto do usuário aos conteúdos de suas mensagens instantâneas, mesmo que apenas por meio de metadados. Como explicitado acima, é diferente do que ocorre nas obrigações legais já existentes de manter "registros telefônicos" ou "registros de conexão" e "de acesso a aplicações de

internet" - como determinado, inclusive, pelo Marco Civil da Internet -, que não têm o condão de vincular a pessoa ao conteúdo trafegado.

Determinar a coleta destes metadados tende a impactar também na privacidade em sentido amplo, pois eles carregam consigo, mesmo que através de inferências, informações essenciais da vida das pessoas. É preciso desmistificar a ideia de que o metadado (ou seja, o dado sobre o dado) não diz nada a respeito das pessoas e, consequentemente, que a sua coleta não viola a privacidade.

Informações acerca de locais frequentados, quais pessoas são contatadas e com que frequência ou mesmo padrões de consumo podem "dizer basicamente tudo sobre a vida de uma pessoa" - nas palavras de um ex-procurador geral da NSA, agência de inteligência estadunidense envolvida no escândalo global de espionagem revelado por Edward Snowden.

Conforme afirmado por Canabarro e Renâ, "coletar, armazenar e tratar metadados é tão perigoso quanto acessar o conteúdo das comunicações privadas". No mesmo sentido, como ilustrou o jornalista Glenn Greenwald, responsável pelas reportagens sobre documentos vazados por Snowden, em entrevista para o InternetLab:

> [...] O metadado mostra mais sobre a sua vida do que o conteúdo. Por exemplo, se uma mulher quer abortar e liga para uma clínica de aborto, você não precisa saber o que ela disse para obter uma informação sobre ela. [...] O metadado mostra mais do que o conteúdo da sua comunicação, porque mostra quem são seus amigos, as pessoas com quem está trabalhando, seu ativismo, com quem está falando. Pode servir para criar uma ideia muito forte e profunda sobre quem você é, o que faz, por qual motivo.

A revelação de cadeias completas de mensagens pode ser invasiva não apenas por, potencialmente, expor relacionamentos individuais, mas também porque, conforme afirmado em documento elaborado pela EFF, "a história completa de certas mensagens pode revelar a estrutura e os membros de uma comunidade inteira, por exemplo, de pessoas que compartilham determinada crença ou interesse, ou que falam certo idioma minoritário, mesmo que nenhuma delas esteja realmente envolvida em atividades ilegais".

Isto é temerário pois afeta de forma desproporcional parcelas vulneráveis da população, em especial ativistas dos direitos humanos e jornalistas investigativos que dependem da criptografia de aplicativos como WhatsApp e Signal para se comunicar com segurança com sua rede de apoio e fontes. É dizer, a imposição da rastreabilidade não vem a custo apenas da criptografia e da privacidade nestas plataformas digitais, mas da própria democracia.

Como se sabe, a criptografia é particularmente vital para a proteção da comunidade LGBTQIA+ e outras pessoas marginalizadas, que são beneficiárias de comunicações criptografadas por questões de segurança. Instituir a obrigação de uma vigilância privada massiva, portanto, é um risco para a privacidade de todos, mas que será desproporcionalmente carregado por parcelas historicamente excluídas da população.

2.4 Risco de autocensura e arrefecimento do debate público

O fato de se compilar dados relacionados às interações entre as pessoas em serviços de mensageria, por si só, já pode gerar autocensura e causar o arrefecimento do debate público (do inglês *chilling effect*). Exemplos claros aparecem quando se foca em temas mais delicados, como expressões de foro íntimo e manifestações políticas em geral.

Em um quadro no qual a percepção genérica será de que tudo que qualquer pessoa escreve pode ser rastreado até elas, há grande probabilidade de que muitos, por vergonha ou medo de retaliação, deixem de se expressar ou de se comunicar acerca de diversos temas importantes para elas e para a sociedade.

Igualmente em vez de incentivar a existência de fontes jornalísticas, de "*whistleblowers*", faz justamente o contrário, cria barreiras, dificulta

Na índia, onde há um debate semelhante sobre rastreabilidade, alguns ativistas políticos declararam que já não se sentem mais seguros em usar aplicações como o WhatsApp e, quando as usam, evitam tratar de temas mais sensíveis por receio de serem monitorados e perseguidos pelo governo.

A necessidade de ordem judicial no Brasil é uma salvaguarda fraca ante a percepção social de que existe esse "banco de dados" geral e que, como visto acima, tende a ser permanente. Igualmente, como já explicitado, o modo como as cadeias de encaminhamento são complexas faz com que interações de todos os tipos e de inúmeras pessoas sejam incluídas em uma potencial requisição de dados.

Nesse sentido, a percepção de monitoramento já deve causar um impacto nas relações entre as pessoas, particularmente em casos sensíveis. Pode-se esperar maior reticência, por exemplo, em situações de abuso, de necessidade de acesso a informação sobre doenças e em contextos de oposição a abuso de autoridade.

2.5 Importância multifacetada da proteção da confidencialidade das mensagens

A proteção gerada pela criptografia é essencial em inúmeros aspectos, mesmo para além da fundamental garantia de liberdade de expressão e comunicação. Há evidências de

que o enfraquecimento da criptografia afeta negativamente o desenvolvimento econômico dos países.

Uma pesquisa realizada pela equipe da Law & Economics Consulting Associates (LECA) na Austrália, após a aprovação em 2018 de uma lei que expande os poderes do governo para contornar sistemas de proteção de dados, demonstrou que a nova legislação aumentou a incerteza dos negócios, reduziu as oportunidades de crescimento e trouxe danos às empresas.

Na mesma toada, Jacqueline Abreu, especialista em proteção de dados e segurança pública, pondera que o artigo 10 cria "falsas expectativas de que sempre vai ser possível rastrear a origem e, se não for, é porque o provedor está dificultando". Consequentemente, isso alimentará "um cenário de insegurança jurídica pouco atraente a operações no Brasil". Ou seja, além de violar direitos fundamentais e a própria ordem democrática por promover um cenário de vigilância em massa, a medida também cria sérios problemas econômicos.

Em adição, a criptografia é importante para proteger a segurança de crianças e adolescentes, pois é uma das formas de garantir que comunicações (incluindo imagens, vídeos e áudios) trocados entre menores continuarão privadas e não cairão nas mãos de eventuais exploradores. Medidas que criam brechas em sistemas criptográficos, como a rastreabilidade proposta pelo PL, facilitam a violação da confidencialidade dessas comunicações e colocam esses conteúdos privados em risco.

Em suma, ao criar um sistema para supostamente combater a desinformação na era digital, o legislador brasileiro está criando um sistema de vigilância massiva que coloca direitos fundamentais, a ordem democrática e a própria economia digital em risco.

3. O Paradoxo da Proteção de Dados Pessoais

Nesta terceira e última seção, ressaltamos que o artigo 10 do PL nº 2.630/2020 vai em sentido oposto ao da Lei Geral de Proteção de Dados e do Marco Civil da Internet, legislações basilares que têm como um de seus pontos centrais o princípio da minimização da coleta e do tratamento de dados pessoais. É paradoxal que seja imposto às empresas por lei, de um lado, uma lógica de proteção de dados e, por outro, seja criada uma nova obrigação legal de uma coleta massiva de uma série de dados que antes não seriam objeto de tratamento específico.

3.1 Substituição do *privacy-by-design* pelo *surveillance-by-design*

Em função da ascensão das leis de proteção de dados em todo o mundo, se popularizaram os sistemas de privacidade desde a concepção (conhecidos pelo nome *privacy-by-design*). Com isso, empresas passaram a estruturar seu modelo de negócio a partir da retenção mínima de dados, deixando de rastrear informações que não são necessárias para o aprimoramento de suas funcionalidades e para a persecução de seus objetivos empresariais. Nessa linha, alguns dos sistemas contemporâneos de mensageria privada são estruturados de forma a impedir que o provedor de serviços e o desenvolvedor tenham acesso aos dados ou conteúdos de mensagens trocadas.

Muitas dessas empresas não necessariamente dispõe de tecnologia ou outra forma para começar a rastreá-los sem alterar, de forma intensa, suas estruturas, que valorizam a proteção da privacidade. Dessa forma, a "adequação" ao artigo 10 do PL significaria um verdadeiro retrocesso. Como explica Katitza Rodriguez no blog da EFF, "o WhatsApp usa uma implementação de *privacy-by-design* que protege as comunicações dos usuários ao fazer com que um encaminhamento seja indistinguível de uma nova mensagem do ponto de vista do servidor". Entretanto, **a obrigação de rastreabilidade das mensagens faz com que a empresa tenha que reestruturar sua aplicação para dar visibilidade a uma informação que antes era invisível.**

Em tempos de *privacy-by-design*, o artigo 10 parece querer impor um *surveillance-by-design*, ou seja, uma vigilância desde a concepção. É dizer, cria-se um estado de vigilância em massa obrigatório por parte das plataformas de mensageria privada a mando do próprio Estado, contrariando os princípios fundantes da legislação nacional sobre o tema, em especial do Marco Civil e da LGPD. Como bem pontuou o WhatsApp em seu relatório "10 mitos e verdades sobre a rastreabilidade":

> [...] A exigência de que as plataformas coletem mais dados do que o necessário, apenas para fornecê-los às autoridades de aplicação da

lei, contraria normas internacionais, contradiz diretamente o princípio da minimização de dados do Marco Civil da Internet e da Lei Geral de Proteção de Dados, e aumenta o escopo da vigilância generalizada.

3.2 Riscos econômicos da afronta ao princípio da minimização do tratamento dos dados pessoais

O ataque ao princípio de minimização do tratamento de dados pessoais também apresenta riscos econômicos. Esse princípio, previsto no Marco Civil e na LGPD, além de proteger os titulares, também reduz riscos e custos operacionais de serviços e aplicações de Internet. Os aplicativos de mensageria teriam que complexificar sobremaneira sua estrutura técnica, pois precisariam criar e manter enormes bases de dados e capacidade de processamento para fazer a correlação entre dados de usuários e mensagens.

Aplicativos de mensageria privada geralmente consomem poucos dados e são considerados, assim, aplicações "leves", o que torna seu uso mais acessível e democrático. Nada obstante, a rastreabilidade acaba exigindo um aumento considerável do fluxo de dados, sobrecarregando servidores e tornando o serviço mais lento e oneroso.

Para piorar, quanto mais complexo o sistema, mais vulnerável e suscetível ele é a falhas de funcionamento e ataques a bancos de dados pessoais. Em outras palavras, a medida, além de ir contra elementos basilares da Lei Geral de Proteção de Dados, também poderia tornar os titulares mais vulneráveis a ataques cibernéticos.

Em um país como o Brasil que enfrenta uma epidemia de mega vazamentos de dados pessoais, muitas vezes financiados por uma indústria hospedada na chamada *dark web*, criar vulnerabilidades estruturais como esta é uma temeridade. O caminho deve ser o da minimização de dados e do fortalecimento das tecnologias que garantem a privacidade dos usuários, como a criptografia ponta-a-ponta, e não o contrário. A solução para a desinformação deve vir em harmonia com a proteção de dados e não com o seu enfraquecimento.

Conclusão

Novas soluções para combater a desinformação vêm sendo testadas e desenvolvidas pela comunidade nacional e internacional, com participação séria e atenta de organizações da sociedade civil, como o ITS Rio. No debate público, realizado de forma aberta e sem precipitações, é possível identificar fraquezas, pontos fortes e boas ideias dentre as novas sugestões que são ventiladas. O PL nº 2.630/2020, ainda que possa ter a boa intenção de "combater as *fake news*", atropela o debate e cria mais problemas do que soluções.

Existem países que já ameaçaram banir os aplicativos de mensageria privada que não compartilham com o governo dados de localização e identificação legal dos usuários. É o caso da China, Rússia e Turquia. São governos que, de fato, buscam impedir que seus cidadãos possam realizar comunicações privadas e seguras com base em tecnologias avançadas de criptografia. É, no mínimo, temerário que o Brasil queira se aproximar desses exemplos.

Mas mais do que isso, como destacou o WhatsApp, "a exigência de rastreabilidade tornará o Brasil uma verdadeira exceção no cenário internacional, pois nenhuma democracia no mundo exige o rastreamento de mensagens privadas" - com a exceção recente da Índia.

No país asiático também há uma discussão, com fortes críticas da sociedade civil, acerca da criação de “IDs digitais” para garantir a rastreabilidade. Ainda que não seja um tema muito discutido no cenário nacional, seus problemas são inúmeros e merecem ser destacados. Trata-se de um identificador que permitiria vincular a pessoa às mensagens que enviou na plataforma. O identificador pode ser visível a todos da cadeia de mensagens (como uma assinatura de e-mail inapagável) ou criptografado de forma a ser acessível apenas pelo serviço de mensageria em casos de ordens judiciais.

Três críticas principais podem ser feitas a essa proposta. Primeiro, a criação de elemento de identificação que permitisse rastrear um conteúdo a alguma pessoa pode ser uma prova muito fraca de que a mensagem efetivamente foi encaminhada por aquela pessoa. Isso porque é relativamente fácil falsear identidades na Internet. Ademais, outra pessoa poderia enviar as mensagens pelo dispositivo que foi identificado em um primeiro momento. Provas que identificam responsabilidade criminal devem ser robustas, e as provas que dependem desse arranjo no contexto da rastreabilidade seriam necessariamente tecnicamente fracas ou tênues.

Segundo, esses elementos de identificação no contexto da rastreabilidade acrescentariam vulnerabilidades ao ecossistema digital. O mesmo sistema poderia ser usado por pessoas mal-intencionadas para monitorar o envio de mensagens por

determinados usuários. Trata-se de uma grande ameaça para a liberdade de expressão de todos os indivíduos, mas especialmente de populações marginalizadas ou envolvidas em manifestações políticas.

Terceiro, seria impraticável a existência de um sistema multiplataforma no contexto da rastreabilidade em função das especificidades técnicas de cada serviço. Para o Internet Society, a rastreabilidade inter-plataforma demandaria, na prática, um registro central de cada aparelho e cada cliente de aplicativo no mundo inteiro, que seria, então, responsável por identificar cada ID digital e possibilitar o rastreamento das cadeias de encaminhamento. Isso afetaria a inovação de uma forma inimaginável, além de criar novos riscos para a privacidade e segurança da informação.

Cumpre ressaltar, ainda, que a rastreabilidade no Brasil, assim como na Índia, é uma ameaça para a criptografia e a privacidade no mundo inteiro. Avaliando o contexto indiano, a Revista Wired alertou que exigir o enfraquecimento da criptografia ponta-a-ponta para revelar a identidade dos usuários que encaminharam determinadas mensagens vai impactar a privacidade dos mais de 400 milhões de usuários na Índia e potencialmente outros bilhões de usuários ao redor do mundo. Afinal, as interações na Internet são essencialmente globais e aplicativos de mensageria privada possibilitam que usuários de diversas nacionalidades interajam entre si.

O combate à desinformação é um problema complexo e não há, ao menos até hoje, solução unívoca. Todavia, algumas das respostas interessantes que se delineiam nos debates mais recentes passam pela análise da arquitetura da desinformação. Uma delas é a implementação de soluções como a limitação do número de encaminhamentos que um usuário pode fazer quando a mensagem se torna viral, diminuindo, assim, a distribuição de conteúdos desinformativos na plataforma sem enfraquecer a criptografia. Como avalia a EFF, essas medidas são caracterizadas pela inclusão de barreiras e limites técnicos que não demandam gestão ou a moderação do conteúdo em si.

Além disso, a implementação de medidas focadas em potenciais desinformadores – como a implementação de "escutas" (temporárias, justificadas e permitidas por ordem judicial), que monitoram metadados – é uma alternativa menos danosa do que tratar todos os usuários de todos os serviços de mensageria privada como suspeitos em potencial, invertendo, assim, o princípio constitucional da presunção de inocência.

Para enfrentar sistematicamente a desinformação na Internet é preciso, antes, pensar em estratégias para desestruturar os sistemas de pessoas e tecnologias responsáveis por produzir, segmentar e distribuir mensagens falsas para determinadas parcelas da população, sistemas estes chamados por Philip Howard de "máquinas da mentira" (do inglês *lie machines*). Uma forma de assim proceder seria "seguir o dinheiro" (“*follow the money*”) para descobrir quem é o ator ou a instituição responsável pelo

financiamento de campanhas de desinformação. Infelizmente, o artigo 10 não contribui com essa empreitada e sequer possibilita a identificação segura e confiável dos verdadeiros responsáveis pela desinformação.

O debate acerca do combate à desinformação está longe do fim. De toda forma, a rastreabilidade de mensagens e discussões legislativas a toque de caixa não são a solução. Como se argumentou neste relatório, são três os principais problemas. Primeiro, a solução proposta pelo artigo 10 do PL nº 2.630/2020 é tecnicamente falha e ineficiente, sendo relativamente fácil de burlar ou, até mesmo, transferir a culpa para inocentes. Segundo, a exigência da rastreabilidade criará um estado de vigilância em massa, impactando desproporcionalmente os mais vulneráveis. Por fim, em terceiro lugar, o PL segue na contramão do Marco Civil e da LGPD, abandonando o paradigma do *privacy-by-design* e inaugurando uma era de *surveillance-by-design*.