Editorial – poliTICs 33

O Instituto Nupef tem procurado acompanhar o panorama de riscos de segurança da Internet, particularmente em função de manter um projeto de serviços Web sem fins de lucro orientado a organizações da sociedade civil (o Tiwa) e atuar na implantação de redes comunitárias, além de manter uma presença histórica no debate internacional de governança da Internet.

O Brasil é hoje bem servido em estruturas de atendimento a incidentes. Uma lista atualizada em abril de 2021 inclui 34 grupos de resposta a incidentes de segurança dedicados ao monitoramento e proteção a setores ou redes específicas, ou de caráter mais amplo, como o CERT.br[1], mantido pelo NIC.br[2], ou o CAIS/RNP[3] (em parceria com 15 centros acadêmicos de atendimento a incidentes), dedicado a responder a incidentes no âmbito da Rede Nacional de Ensino e Pesquisa (RNP)[4]. Mundialmente a rede de centros de resposta é muito grande e diversificada. Só na União Europeia há mais de 640 entidades-membros da rede de CSIRTs[5].

Este número da poliTICs inclui uma resenha produzida originalmente pelo CERT.br para o Encontro NETmundial[6], adaptada e atualizada para esta publicação com autorização dos/as autores/as originais.

Publicamos também uma resenha produzida por Access Now[7] sobre o avanço das técnicas da chamada "realidade estendida" (RE), que motivam uma corrida das principais operadoras de plataformas sociais em busca de novas ofertas de serviços imersivos. Ao lado de possíveis impactos positivos, a RE apresenta riscos ampliados para os direitos humanos, especialmente pelo potencial de captura de dados em tempo real em novo patamar, aprofundando os riscos de violação da privacidade. Para agravar ainda mais o cenário, as regulações de proteção existentes estão longe de alcançar as possibilidades dessas novas técnicas.

Finalmente, a professora Meredith Whittaker alerta para a proliferação da inteligência artificial (AI) mais estimulada pela profusão de dados disponíveis do que por avanços nos algoritmos. O controle pelas grandes empresas de tecnologia sobre o uso da AI torna os principais centros de desenvolvimento dependentes dessas empresas, orientando o avanço tecnológico segundo regras impostas por elas -- uma reprodução da influência militar sobre a pesquisa científica durante a Guerra Fria.

Boa leitura!

---

[1] *https://cert.br/*
[2] *https://nic.br*
[3] *https://www.rnp.br/sistema-rnp/cais*
[4] *https://rnp.br*
[5] *https://www.enisa.europa.eu/topics/csirts-in-europe/csirt-inventory/certs-by-country-interactive-map*
[6] *https://netmundial.br/pt/*
[7] *https://www.accessnow.org/*

A Importância de uma abordagem multissetorial para a segurança cibernética efetiva

Texto baseado em artigo de Cristine Hoepers, Klaus Steding-Jessen e Henrique Faulhaber para o Encontro NETmundial de 2014, adaptado e atualizado para esta edição da poliTICs.

1. Introdução

Muitas das ameaças de segurança da Internet são cada vez mais complexas, afetando vários setores ao mesmo tempo e exigindo esforços coordenados para ser detectadas e efetivamente mitigadas. Isto é especialmente verdadeiro para os incidentes envolvendo botnets, spam, malware e DDoS (Distributed Denial of Service).

Nos últimos 30 anos foram criados vários fóruns com várias partes interessadas e iniciativas que tratam de ameaças à segurança da Internet - a maioria deles têm sido muito bem sucedidos em juntar diferentes setores para em conjunto mitigar os incidentes de segurança e o cibercrime. Todos estes esforços revelaram que a eficácia depende da cooperação entre os diferentes setores, e que a segurança cibernética não pode ser alcançada através de uma única organização ou estrutura. Além disso, os governos precisam participar mais em fóruns de segurança e melhorar a cooperação com os outros setores. Novos fóruns e iniciativas não devem substituir as estruturas existentes, que devem visar a alavancagem e melhorar as estruturas já existentes dos vários grupos de interesse.

O cenário fica mais complicado quando as infraestruturas críticas nacionais estão ligados à Internet, tornando-se expostas às mesmas vulnerabilidades que outros sistemas, e podem ser atacadas pelas mesmas ferramentas ou técnicas usadas para ataques em outros contextos.

A proteção das infraestruturas críticas e redes governamentais ligadas à Internet envolve os aspectos de defesa e de segurança da Internet - a proteção destas infraestruturas é feita na maioria das vezes por organizações governamentais. O preocupante é que estamos vendo cada vez mais as questões que são puramente de segurança da Internet sendo percebidas pelos governos como questões puramente de defesa. Isso está levando a um cenário em que, por exemplo, a cooperação vital já existente entre as CERT (Computer Emergency Response Teams) que têm responsabilidade nacional está sendo prejudicada por uma tendência a transferir todos os recursos de segurança da Internet existentes para organizações governamentais ou de inteligência.

O ecossistema de segurança, estabilidade e resiliência da Internet deve permanecer multissetorial. A cooperação entre os diferentes setores e partes interessadas, hoje já existente, é a chave para mitigar muitas das ameaças atuais.

No restante desta proposta, vamos discutir brevemente vários fóruns multissetoriais e iniciativas em andamento, identificando os seus pontos fortes, e destacar temas que precisam ser considerados quando se discute uma estrutura para melhorar a abordagem multissetorial que viabilize uma segurança cibernética mais eficaz.

2. Fóruns existentes envolvendo múltiplos agentes sociais

Há vários fóruns internacionais já existentes que congregam diferentes agentes, cooperando para lidar com incidentes de segurança e mitigar ameaças específicas. A maioria destes fóruns foram criados para mitigar categorias específicas de ataques ou ameaças. Como hoje em dia o cenário de ameaças mudou e há uma prevalência do que é tecnicamente conhecido como ameaças combinadas, a maioria dessas organizações está lidando com questões de segurança semelhantes. O que se segue é uma descrição de cada um destes organismos.

2.1. FIRST

O FIRST é o Fórum de Equipes de Resposta a Incidentes e Segurança[1]. A CSIRT (Equipe de Resposta a Incidentes de Segurança)[2], às vezes também referida como CERT, é uma organização de serviços que é responsável por receber, analisar e responder a relatos e atividades de incidentes de segurança informática. Seus serviços são geralmente realizados para um público definido, como uma entidade ou organização empresarial, governamental ou educacional, uma região ou país, uma rede de pesquisa, ou um cliente pago[3].

O primeiro CSIRT, o Centro de Coordenação CERT, foi criado em Novembro de 1988, após o incidente de segurança conhecido como "verme da Internet" ou "verme de Morris", derrubou grande parte da Internet, e deixou clara a necessidade de esforços mais coordenados para responder a incidentes de segurança[4]. Após este incidente, foram criadas várias outras equipes. A primeira foi formada em 1990, em resposta a um segundo verme, o "verme WANK"[5], e este incidente destacou a necessidade de uma melhor comunicação e coordenação entre equipes de diferentes organizações.

FIRST é uma confederação internacional de equipes confiáveis de resposta a incidentes de computador[6] que cooperativamente lida com incidentes de segurança informática e promove programas de prevenção de incidentes. O FIRST reúne uma grande variedade de CSIRTs de todo o mundo, incluindo entidades educacionais e comerciais, distribuidores, entidades nacionais, governo e militares. Os membros do FIRST desenvolvem e compartilham informações técnicas, ferramentas, metodologias, processos e melhores práticas, e usam seus conhecimentos combinados, habilidades e experiência para promover um ambiente mais seguro na Internet.

2.2 CiviCERT

Esta é uma rede de CERTs e provedores de conteúdo e serviços de Internet que colaboram com entidades civis para prevenir e lidar com temas de segurança digital[7]. A rede é uma iniciativa de um grupo de organizações sem fins de lucro, provedores e indivíduos conhecida como RaReNet (Rapid Response Network)[8] que buscam contribuir tempo e recursos para a conscientização sobre segurança digital da sociedade civil. Entre os membros estão entidades como a Access Now, Anistia Internacional, Colnodo, Electronic Frontier Foundation e Human Rights Watch.

2.3. CSIRTs com responsabilidade nacional

Desde 2006, o Centro de Coordenação CERT (CERT/CC)[9] tem promovido uma reunião técnica anual para CSIRTs com responsabilidade nacional. Esta reunião é uma oportunidade para as organizações responsáveis por proteger a segurança das nações, economias e infraestruturas críticas de discutir os desafios que eles enfrentam ao cumprir este papel. Como resultado destas reuniões, um fórum online é mantido durante todo o ano, bem como uma lista de CSIRTs com responsabilidade nacional[10].

Vale ressaltar que existem diferentes modelos de CSIRT nacionais, que vão desde estruturas sem fins de lucro, acadêmicas e até equipes de governo. Além disso, vários países têm mais de uma equipe, o que demonstra a complexidade de aumentar a segurança cibernética e realizar o tratamento de incidentes a nível nacional.

2.4. APWG

O APWG (Grupo de Trabalho Antiphishing)[11] foi fundado em 2003, momento em que a sua missão era combater ataques de *phishing*.[12] Mas, como a tecnologia evoluiu, APWG não é mais focado apenas em *phishing*, mas na mitigação de outros ataques que são usados para cometer crimes cibernéticos. APWG tem mais de dois mil membros e parceiros de pesquisa em todo o mundo, desde instituições financeiras, varejistas, provedores de soluções, provedores de Internet, empresas de telecomunicações, CSIRTs,

universidades, empreiteiros da defesa, agências de aplicação da lei, grupos de comércio, agências multilaterais e governamentais.

2.5. M3AAWG

O M3AAWG (Grupo de Trabalho Móvel de Mensagens, Malware e Anti-Abuso)[13] aglutina a indústria de mensagens para trabalhar de forma colaborativa no tratamento exitoso de várias formas de abuso de mensagens, tais como spam, vírus, ataques de negação de serviço e outras explorações de mensagens. Para isso, desenvolve iniciativas M3AAWG nas três áreas necessárias para resolver o problema de abuso de mensagens: colaboração da indústria, tecnologia e políticas públicas.

2.6. ISOC

A ISOC (Internet Society)[14] é uma organização dedicada a garantir que a Internet permaneça aberta e transparente. Ela tem iniciativas em políticas de Internet, padrões de tecnologia e desenvolvimento futuro. A ISOC tem um projeto especial chamado "Projeto Combatendo o Spam", em parceria com MAAWG, dedicado a mostrar a formadores de políticas, de forma clara e eficaz, as ferramentas e parcerias industriais que estão disponíveis para combater spam.

3. Exemplos de iniciativas multissetoriais exitosas nacionais e internacionais

Nos últimos anos, CSIRTs, operadores de rede e os membros dos fóruns acima mencionados envolveram-se em alguns projetos e grupos de trabalho destinados a mitigar grandes ameaças específicas, implementando as melhores práticas ou compreendendo melhor o ambiente de ameaças na Internet. Nesta seção, vamos descrever algumas dessas iniciativas multissetoriais exitosas.

3.1. Grupo de Trabalho Conficker

Começando no final de 2008, e continuando até junho de 2010, uma coalizão de pesquisadores de segurança trabalhou para resistir a um ataque na Internet por software malicioso conhecido como Conficker. Esta coligação ficou conhecido como "Grupo de Trabalho Conficker", e pareceu ser bem sucedida de várias maneiras, entre as quais a cooperação sem precedentes entre organizações e indivíduos em todo o mundo, em ambos os setores público e privado[15].

O trabalho deste grupo envolveu membros dos órgãos de governança da Internet, software e fornecedores de hardware, provedores de conteúdo, universidades e centros de pesquisa, e foi vital para mitigar

transportadores maliciosos do verme e para ajudar a limpar sistemas em toda a Internet. Um documento de lições aprendidas pode ser encontrado na página inicial do sítio mencionado.

3.2. DCWG

O DCWG (Grupo de Trabalho DNS Changer)[16] foi um grupo ad hoc de especialistas no assunto, e incluiu membros de organizações como Georgia Tech, Internet Systems Consortium, Mandiant, National Cyber-Forensics e Training Alliance, Neustar, Spamhaus, Equipe Cymru, a Trend Micro, e Universidade do Alabama em Birmingham. O trabalho do DCWG foi coordenado com as investigações do FBI, e recebeu ajuda de vários CERTs nacionais e provedores de acesso.

Este grupo de trabalho foi criado para ajudar a sanar os servidores DNS maliciosos da Rove Digital. O botnet operado pela Rove Digital alterava as configurações de DNS do usuário, direcionando as vítimas para servidores de DNS maliciosos em centros de dados na Estônia, Nova York e Chicago. Os servidores de DNS maliciosos davam respostas falsas e maliciosas, alteravam buscas do usuário e promoviam produtos falsificados e perigosos. Como cada pesquisa na Web começa no DNS, o malware mostrava aos usuários uma versão alterada da Internet.

A cooperação entre todos esses setores permitiu alertar gradualmente e ajudar a desinfetar dispositivos dos usuários finais, sem interromper o seu acesso à Internet.

3.3. Iniciativas multissetoriais a nivel nacional

Existem várias iniciativas multissetoriais nacionais. Nesta seção, vamos descrever brevemente algumas dessas iniciativas.

3.3.1. Conselho de Ciberssegurança Holandês

O Conselho de Ciberssegurança Holandês tem 15 membros do governo, da indústria e da comunidade científica, para um total de três cientistas, seis do setor público e seis representantes do setor privado[17]. O Conselho, apoiado por uma secretaria independente, supervisiona a estratégia de segurança cibernética nacional holandesa e oferece recomendações ao governo holandês e à sociedade. O papel que o Conselho desempenhou durante o incidente DigiNotar, por exemplo, demonstrou a eficácia deste tipo de parceria público-privada no domínio digital[18].

Em julho de 2013, o Conselho emitiu um parecer sobre a nova Estratégia Nacional de Segurança Cibernética, publicado em Outubro de 2013. As recomendações eram voltadas especificamente para a necessidade de uma estreita cooperação e coordenação em matéria de detecção e resposta a

incidentes. Somente através compartilhamento ativo de informação, resposta oportuna e colaboração contínua pode um ambiente digital seguro ser estabelecido.

3.3.2. CCC - Centro Japonês de Ciberlimpeza

O CCC é um organismo central que promove a limpeza de bots e prevenção de reinfecção de computadores infectados dos usuários, baseado na cooperação entre o governo, os fornecedores de software e os provedores de acesso[19]. O CCC tem um Comitê Gestor que supervisiona três grupos de trabalho: o grupo de operação do sistema de contramedidas a bots, o grupo de análise dos programas bots, e o grupo promotor da prevenção da infecção por bots.

3.3.3. Iniciativa de Gestão da Porta 25 - CGI.br

Durante muito tempo, o Brasil esteve presente na maioria dos rankings como país topo de reenvio de spam. Determinado a reverter essa situação, o Comitê Gestor da Internet no Brasil (CGI.br) realizou, desde 2005, uma série de atividades, tais como estudos acadêmicos e análises técnicas, que levaram à adoção Gestão da Porta 25 como a medida mais eficaz para evitar o abuso da infraestrutura de banda larga no Brasil por spammers. Esta iniciativa foi liderada pelo Grupo de Trabalho Antispam do CGI.br (CT- Spam), que manteve um fórum onde os diferentes atores puderam reunir-se[20].

Por quase 20 anos, o Brasil desenvolveu um modelo de governança da Internet multissetorial. Portanto, uma medida de tamanha importância como o bloqueio de tráfego de saída da porta 25 em redes residenciais não poderia ser adotada sem a devida consulta a todos os setores afetados, que foram solicitados a contribuir para este processo de tomada de decisão.

Reunindo a experiência de empresas de telecomunicações, mais de uma dezena de milhares de provedores de serviços de Internet, representantes da sociedade civil e da comunidade acadêmica, bem como a equipe técnica do CGI.br, o processo de adoção da gestão da porta 25 foi amplamente discutido. Isto foi especialmente importante porque a implementação exigiu um esforço concertado, assegurando que os provedores de serviços de e-mail passassem a oferecer a submissão de mensagens através de uma porta diferente (587), e migraram pelo menos 90% da base de seus usuários antes que os provedores de banda larga bloqueassem a porta de saída 25.

Também é importante ressaltar que tanto a Agência Nacional de Telecomunicações (Anatel) e o Ministério da Justiça têm desempenhado um papel fundamental no fornecimento de suporte para as empresas de telecomunicações e as entidades de defesa do consumidor,

respectivamente. Anatel assinou um Acordo de Cooperação com CGI.br, que deu às empresas de telecomunicações fundamentos jurídicos para prosseguir com a adoção. O Ministério da Justiça, por outro lado, publicou uma nota técnica explicando os benefícios de tais medidas para os consumidores.

Como resultado desta iniciativa, o Brasil já não é listado como um dos principais países no reenvio de spam no mundo, de acordo com vários rankings públicos.

3.3.4. CERT.br - Computer Emergency Response Team Brasil

O CERT.br é mantido pelo NIC.br, uma organização sem fins lucrativos criada para implementar as decisões e projetos elaborados pelo Comitê Gestor da Internet no Brasil - CGI.br. Todas as atividades do CERT.br levam em conta a necessidade de envolver todas as partes interessadas para aumentar com êxito o nível de segurança a capacidade de lidar com incidentes nas redes conectadas à Internet no Brasil[21].

Além de realizar atividades de tratamento de incidentes, o CERT.br também trabalha para aumentar a percepção sobre segurança na comunidade brasileira, mantendo um projeto de alerta prévio, com o objetivo de identificar novas tendências e correlacionar eventos de segurança, bem como alertando redes brasileiras envolvidas em atividades maliciosas. CERT.br também ajuda os novos CSIRTs a estabelecer suas atividades no país.

Um exemplo claro do sucesso dessa abordagem é o Projeto Brasileiro de Honeypots Distribuidos, que, através de uma rede de honeypots distribuídos no espaço Internet brasileiro, aumenta a capacidade de detecção de incidentes, correlação de eventos e análise de tendências no país[22]. Estes honeypots são sensores passivos que fornecem percepção situacional valiosa, sem coletar o tráfego normal de dados nem realizar qualquer tipo de vigilância. Este projeto tem sensores em mais de 40 organizações brasileiras parceiras, que vão desde os setores do governo e da energia, a academia, provedores de serviços Internet e provedores de telecomunicações.

3.3.5 Iniciativas do NIC.br de apoio à cibersségurança

O NIC.br, através de seus programas permanentes, mantém várias iniciativas de monitoramento e apoio sobre a segurança das redes e dispositivos da Internet. Entre estas estão:

Campanha Fique Esperto[23] -- É uma iniciativa que une governo e entidades privadas com o objetivo de informar às pessoas sobre como evitar golpes usuais do nosso novo mundo digital. Durante o período da campanha, são

apresentadas, mensalmente, novas dicas relacionadas aos problemas mais comuns e às medidas de prevenção que podem ser tomadas para evitá-los.

Cidadão na Rede[24] – conduzido pela equipe do Ceptro.br[25], incentiva boas práticas relacionadas à cidadania digital e ao bom uso da Internet, alcançando o maior número possível de seus usuários. Com animações curtas, que explicam de maneira simples como usar a rede de forma correta e responsável, abrangendo questões técnicas e comportamentais, dicas importantes podem ser transmitidas e compartilhadas pela rede. Empresas e organizações interessadas podem se tornar parceiras dessa iniciativa, fazendo o download gratuito dos vídeos ou solicitando a inclusão do seu logo em uma versão customizada dos vídeos, para divulgação em seu site, ou outros canais.

Internet Segura[26] -- Idealizado pelo CGI.br, o portal Internet Segura reúne iniciativas de conscientização sobre segurança e uso responsável da Internet no Brasil, auxiliando os internautas a localizar as informações de interesse e incentivando o uso seguro da Internet. O portal também traz iniciativas de outras entidades e instituições sobre o uso seguro da Internet.

3.3.6. CAIS/RNP

Este centro de atendimento a incidentes de segurança é mantido há mais de 20 anos pela Rede Nacional de Ensino e Pesquisa (RNP)[27], e oferece apoio à adoção de boas práticas de segurança nas instituições acadêmicas e outras conectadas à extensa e diversificada rede da RNP. CAIS é um dos primeiros centros de resposta a incidentes a operar nacionalmente, e mantém parceria com 15 CSIRTs de entidades acadêmicas no Brasil. Oferece também apoio a órgãos públicos.

4. A necessidade de melhoria da colaboração multissetorial em cibersegurança

Atingir um nível satisfatório de segurança na Internet não é uma tarefa fácil, mas a experiência acumulada por várias iniciativas de sucesso demonstra que, para ser eficaz, qualquer iniciativa de segurança cibernética precisa envolver vários intervenientes. Mais do que isso, a realidade é que na maioria das vezes, as medidas de segurança devem ser tomadas por administradores de sistemas, operadores de rede ou profissionais de segurança em suas próprias redes. No entanto, a cooperação com os outros é a chave para ser capaz de entender as ameaças e melhor avaliar a eficácia de suas ações.

No documento "Grupo de Trabalho Conficker: Lições Aprendidas"[28], publicado em Janeiro de 2011, embora não apareça o conceito de

"multissetorial", alguns dos fatores de sucesso listados indicam a importância da cooperação e do envolvimento de diferentes grupos de interesse. Aqui estão alguns exemplos:

- utilizar um modelo de confiança; o grupo de trabalho deve ter um tamanho gerenciável para ser eficaz e incluir aqueles diretamente afetados, mas grande o suficiente para incluir um universo mais amplo de pessoas afetadas;
- incorporar um modelo de consenso, sem hierarquia, para permitir que o grupo se adpte e responda à rápida mudança das condições;
- ganhar a participação e apoio de órgãos reguladores e de governo relevantes;
- formalizar a comunicação com grupos de interesse em vez de contar com redes sociais.

Estes quatro pontos trazem à luz questões como a rápida mudança do cenário de ameaças, a necessidade de uma comunicação rápida, o envolvimento e apoio dos governos e o fato de que os vários grupos de interesse precisam cooperar.

Embora o Grupo de Trabalho Conficker tenha sido muito bem sucedido, bem como outras iniciativas listadas na seção anterior, ainda existem alguns setores que poderiam melhorar a sua cooperação. Por exemplo:

- Os Grupos de Operadores de Rede (NOGs) e Registros Regionais de Internet (RIRs) deveriam estar mais envolvido com as questões de segurança. Há algumas áreas, como a segurança de roteamento (e protocolos recentemente propostos como RPKI ou SBGP) ou DNSSEC que precisam de adoção em todo o mundo para serem eficazes. Os RIRs também podem trabalhar mais próximos à comunidade CSIRT para melhorar o sistema WHOIS e com isso ajudar o processo de tratamento de incidentes.
- Os fornecedores de software precisam envolver-se e serem mais proativos – afinal, a maior parte dos problemas de segurança que enfrentamos hoje são problemas relacionados com software. O verdadeiro desafio é melhorar a segurança do software e elevar a indústria de software a um nível mais maduro.
- Os governos, incluindo os setores militares e de inteligência, além de estratégias de segurança e defesa tradicionais, precisam melhorar a sua consciência da natureza multissetorial da Internet e a importância vital da cooperação para enfrentar as ameaças à segurança. Eles precisam participar mais nos fóruns nacionais e

internacionais de segurança e melhorar a cooperação com outros setores.

Considerando-se as estratégias de segurança cibernética de governos, é de salientar que cerca de 130 representantes de vários setores, incluindo entidades públicas e privadas, instituições de conhecimento e organizações sociais, estiveram envolvidos na elaboração da "Estratégia Nacional de Segurança Cibernética 2 - Da consciência à capacidade" da Holanda (NCSS2)[29]. A estratégia começa com a seguinte declaração: "Estamos caminhando de estruturas a coalizões em que todas as partes – nacionais e internacionais – são representadas, a fim de atingir os padrões suportados."

E acrescenta: "A correlação entre segurança, liberdade e benefícios sócio-econômicos propostos no NCSS2 é um equilíbrio dinâmico, que se destina a ser realizado em um diálogo constantemente aberto e pragmático entre todos os intervenientes, tanto nacionais como internacionais. (...) A fim de trazer o diálogo sobre segurança cibernética entre as várias partes interessadas para um novo nível de maturidade, as três seguintes áreas de gestão são de extrema importância : (auto-) regulação, transparência e desenvolvimento do conhecimento."

Este é um bom exemplo do reconhecimento da importância de uma abordagem multissetorial para a segurança, estabilidade e resiliência do ecossistema da Internet.

5. Recomendações

Como já mencionado, alcançar um nível satisfatório de segurança na Internet não é uma tarefa fácil, e as iniciativas já discutidas envolvendo vários grupos de interesse são bons exemplos de estruturas que podem efetivamente lidar com questões correntes e emergentes de segurança cibernética. Portanto, recomenda-se que todas as organizações nacionais e internacionais envolvidas com a governança da Internet, por exemplo, os governos, os registradores regionais de endereços IP (RIRs), as Nações Unidas, a União Europeia, os capítulos da Internet Society, entre outros, devam levar em consideração o seguinte :

1. A experiência acumulada pelas diversas iniciativas de sucesso descritas nesta contribuição demonstra que, para ser eficaz, qualquer iniciativa de segurança cibernética depende da cooperação entre os diferentes atores, e isso não pode ser alcançado através de uma única organização ou estrutura.

2. Há atores que ainda precisam envolver-se mais, como operadores de rede e desenvolvedores de software.

3. Governos, incluindo os setores militares e de inteligência, além das estratégias de segurança e defesa tradicionais, precisam melhorar a sua consciência da natureza multissetorial da Internet e da importância vital da cooperação para enfrentar as ameaças à segurança. Eles precisam participar mais nos fóruns nacionais e internacionais de segurança e melhorar a cooperação com outras partes interessadas.

4. Há espaço e necessidade de novos fóruns e iniciativas, mas não devem substituir as estruturas existentes. Qualquer nova iniciativa deve ter como objetivo alavancar e melhorar as estruturas multissetoriais já em vigor hoje.

1 *https://www.first.org/*

2 *https://www.csirt.org/*

3 *https://resources.sei.cmu.edu/library/asset-view.cfm?assetID=485652*

4 *https://en.wikipedia.org/wiki/Morris_worm*

5 *https://en.wikipedia.org/wiki/WANK_(computer_worm)*

6 "Computador" hoje é entendido como o universo gigantesco de dispositivos que podem ser conectados à Internet – tabuletas, celulares, computadores portáteis e de mesa, dispositivos embarcados fixos ou móveis, veículos etc.

7 *https://www.civicert.org/*

8 *https://www.rarenet.org/*

9 *https://www.kb.cert.org/vuls/*

10 *https://www.sei.cmu.edu/our-work/cybersecurity-center-development/index.cfm*. A lista de CSIRTs europeus pode ser vista aqui: *https://csirtsnetwork.eu/*

11 *https://apwg.org/*

12 *https://en.wikipedia.org/wiki/Phishing*

13 *https://www.m3aawg.org/*. O grupo hoje é chamado de M3AAWG – Message, Malware, Mobile Anti-Abuse Working Group.

14 *https://www.internetsociety.org/*

15 *https://www.senki.org/operators-security-toolkit/security-organizations/conficker-working-group-archive-of-materials/*

16 *https://www.crunchbase.com/organization/the-dns-changer-working-group---dcwg*

17 *https://www.ncsc.nl/english/current-topics/news/best-practices-in-computer-network-defense.html*

18 *https://en.wikipedia.org/wiki/DigiNotar*

19 *https://www.ccc.go.jp/en_ccc/*

20 *https://www.cert.br/docs/palestras/certbr-ct-spam-cgibr2008.pdf*

21 *https://www.cert.br/*

22 *https://honeytarg.cert.br/honeypots/index-po.html*

23 *https://fe.seg.br/*

24 *https://cidadaonarede.nic.br*

25 *https://ceptro.br/*

26 *https://internetsegura.br/*

27 *https://www.rnp.br/sistema-rnp/cais*

28 *http://docs.media.bitpipe.com/io_10x/io_102267/item_465972/whitepaper_76813745321.pdf*

29 *https://www.enisa.europa.eu/topics/national-cyber-security-strategies/ncss-map/NCSS2Engelseversie.pdf*

Mundos virtuais, pessoas reais: direitos humanos no metaverso

Access Now[1] — 09-12-2021

No dia 10 de dezembro é celebrado o Dia Internacional dos Direitos Humanos. Nesse dia de 1948, a Assembleia Geral da ONU adotou a Declaração Universal dos Direitos Humanos[2], o documento que estabelece os princípios e as bases dos instrumentos de direitos humanos atuais e futuros. Em homenagem a este aniversário, a Access Now e a Electronic Frontier Foundation (EFF)[3] conclamam governos e empresas a abordar os direitos humanos no contexto da realidade virtual e aumentada[4] (RV e RA[5]) e garantir que esses direitos sejam respeitados e aplicados.

As tecnologias de realidade estendida[6] (RE), incluindo realidade virtual e aumentada, são a base de ambientes digitais emergentes, incluindo o chamado metaverso. Elas ainda estão em um estágio inicial de desenvolvimento e adoção, mas as empresas que conformam a "Big Tech" estão investindo pesadamente nessas tecnologias[7], e há uma corrida para conquistar o domínio e consolidar os monopólios do que os investidores e executivos de tecnologia afirmam ser a próxima geração de computação e de mídia social.

Como qualquer outra tecnologia, a RE pode ter muitos impactos positivos em nossas vidas diárias. Pode ser uma ferramenta útil em áreas como medicina[8], ciência[9] e educação[10]. Os artistas estão usando RE de maneira criativa para transformar os mundos virtuais em sua tela e criar novas formas de expressão. Protestos e movimentos sociais também têm usado essas tecnologias para aumentar a consciência sobre questões coletivas ou para fazer ouvir sua voz em situações em que essa prática é fisicamente impossível ou perigosa[11].

No entanto, a RE também representa riscos substanciais para os direitos humanos. Fones de ouvido RV e óculos RA, juntamente com outros “wearables”, podem agravar o cenário de coleta de dados cada vez mais invasiva e de vigilância onipresente[12]. Esta coleta de dados, às vezes feita por empresas que priorizam o lucro em detrimento das proteções[13], prepara o terreno para invasões sem precedentes em nossas vidas, nossas casas e até mesmo em nossos pensamentos, já que os dados coletados por dispositivos RE são usados para publicidade direcionada e para permitir novos formas de “psicografia biométrica”[14] — inferir nossos desejos e inclinações mais profundos. Uma vez coletados, há pouco que os usuários podem fazer para mitigar os danos causados por vazamentos de dados ou

dados sendo monetizados por terceiros. Esses dispositivos também coletarão grandes quantidades de dados sobre nossos lares e espaços privados e podem permitir que governos, empresas e agentes da lei tenham acesso ilegítimo a nossas vidas[15], exacerbando as já severas invasões em nossa privacidade.

Essas novas tecnologias também criam novos caminhos para o assédio e o abuso online[16]. Os óculos RA podem minar drasticamente as expectativas de privacidade em espaços públicos e privados. Uma pessoa que usa os óculos pode gravar facilmente o que está ao seu redor em segredo[17], o que pode tornar-se ainda mais perigoso se tecnologias de vigilância, como reconhecimento facial[18], forem incorporadas.

Aprendemos muitas lições com tudo que deu errado e certo com a geração atual de dispositivos inteligentes e mídia social, e precisamos aplicar essas lições agora para garantir que todos possam tirar proveito das tecnologias RE e do metaverso sem sacrificar direitos humanos fundamentais que temos que preservar.

Aqui estão os tópicos que merecem cuidadosa e urgente consideração:

- Sabemos que a autorregulação sobre proteção de dados e diretrizes éticas não são suficientes para conter os danos causados pela tecnologia.
- Sabemos que precisamos que os padrões de direitos humanos sejam centrais no desenvolvimentos da RE para garantir que nossos direitos não sejam apenas respeitados, mas também estendidos ao metaverso.
- Precisamos de regulamentação e aplicação adequadas para proteger a privacidade e outros direitos humanos das pessoas no metaverso.
- Também precisamos nutrir a tecnologia de base respeitadora dos direitos que vem sendo desenvolvida hoje. Os legisladores precisam estar atentos para que as grandes empresas de tecnologia não engulam todos os seus concorrentes antes que eles tenham a chance de desenvolver alternativas que respeitem os direitos das plataformas dominantes orientadas para a vigilância.

Para este fim, pedimos aos governos que garantam que as proteções contra o alcance amplo e a intrusão estatal e corporativa se apliquem ao RE, como segue:

- Os governos devem promulgar ou atualizar a legislação de proteção de dados que limite a coleta e o processamento de dados para incluir dados gerados e coletados por sistemas de RE, incluindo inferências médicas ou psicográficas[19]. Os governos devem definir claramente esses dados como dados pessoais sensíveis e fortemente protegidos de acordo com a lei, mesmo quando não qualifiquem para serem classificados como dados biométricos, pessoais ou informações de identificação pessoal de acordo com a legislação atual. A legislação deve reconhecer que os sistemas de RE podem ser usados para fazer inferências problemáticas e invasivas[20] sobre nossos pensamentos, emoções, inclinações e vida mental privada[21].
- Autoridades independentes responsáveis devem agir para fazer cumprir as leis de proteção de dados e proteger os direitos das pessoas. Pesquisas mostram que as "escolhas" de privacidade das pessoas para permitir que as empresas processem seus dados são normalmente involuntárias[22], sujeitas a vieses cognitivos e/ou contornáveis devido a limitações humanas, padrões obscuros[23], lacunas legais e as complexidades do processamento de dados moderno. As autoridades devem exigir transparência e controle não apenas sobre os dados coletados, mas também sobre o uso ou divulgação das inferências que a plataforma fará sobre os usuários (seu comportamento, emoções, personalidade etc), incluindo o processamento de dados pessoais em segundo plano. Assim, o paradigma legal de notificação e escolha, tal como é praticado hoje, precisa ser desafiado.
- O metaverso não deve pertencer a nenhuma empresa. Os reguladores da concorrência devem tomar medidas para salvaguardar a diversidade de plataformas de metaverso e evitar monopólios sobre infraestrutura e hardware, para que os usuários não se sintam presos a uma determinada plataforma para desfrutar de plena participação na vida cívica, pessoal, educacional, social ou comercial, ou sentir que eles têm que tolerar essas falhas para permanecer conectados aos espaços vitais da existência humana. Essas intervenções pró-competitivas devem incluir escrutínio de fusões, separação estrutural de empresas dominantes de elementos adjacentes de suas cadeias de abastecimento, proibições de conduta anticoncorrencial, como preços predatórios, interoperabilidade obrigatória de protocolos-chave e estruturas de dados e salvaguardas legais para operadores que usam engenharia reversa e outra “interoperabilidade adversária” para melhorar a segurança,

acessibilidade e privacidade de um serviço. Esta não é uma lista exaustiva e, se as tecnologias do metaverso se consolidarem, quase certamente darão origem a novos direitos humanos, preocupações e soluções de concorrência específicos das novas tecnologias.

- Os governos devem garantir que os 13 Princípios Internacionais sobre a Aplicação dos Direitos Humanos à Vigilância das Comunicações[24] sejam aplicados, os privilégios existentes contra a intrusão do governo sejam reafirmados e as proteções legais sejam estendidas a outros tipos de dados, como dados psicográficos e comportamentais e inferências extraídas deles.
- Os governos devem aumentar a transparência em torno do uso de RE. À medida que os governos começam a usar o RE para treinamento e simulações, deliberação e tomada de decisões e reuniões públicas, novos tipos de informações serão produzidos que constituirão registros públicos que devem ser ofertados ao público de acordo com as leis de liberdade de informação.
- À medida que as tecnologias RE se tornam onipresentes, as empresas devem respeitar e os governos devem proteger o direito das pessoas de reparar, alterar ou investigar a funcionalidade de seus próprios dispositivos.
- Como na vida real, os governos devem abster-se de censurar a liberdade de expressão e inibir as liberdades jornalísticas e, em vez disso, encorajar trocas participativas no mercado de ideias. Com o surgimento de iniciativas regulatórias em todo o mundo que ameaçam prejudicar a liberdade de expressão, é fundamental aderir a medidas proporcionais, consistentes com os Princípios de Santa Clara[25], equilibrando objetivos legítimos com a liberdade de receber e divulgar informações.

As empresas têm a responsabilidade de defender os direitos humanos de acordo com os Princípios Orientadores sobre Negócios e Direitos Humanos da ONU[26], um padrão global de "conduta esperada para todas as empresas onde quer que operem", aplicável em todas as situações.

A responsabilidade corporativa de respeitar os direitos humanos também significa abordar os impactos adversos que podem ocorrer, como segue:

- As empresas devem se comprometer publicamente a requerer que os governos sigam o processo legal necessário para acessar os dados de RE[27], notificar os usuários quando permitido por lei, publicar regularmente relatórios de transparência, utilizar criptografia (sem *backdoors*) e lutar para limitar os dados que

podem ser acessados ao que é necessário, adequado e proporcional.

- As empresas, incluindo fabricantes e fornecedores, não devem apenas proteger o direito de seus usuários à privacidade contra a vigilância do governo[28], mas também o direito de seus usuários à proteção de dados. Eles devem resistir ao impulso, muito comum no Vale do Silício, de "coletar tudo", caso possa ser útil mais tarde. Em vez disso, as empresas devem aplicar princípios estritos de minimização de dados (*privacy-by-design*), coletando apenas o que é necessário para a funcionalidade de base ou para fornecer serviços específicos que os usuários solicitaram e concordaram, e retendo-os apenas pelo tempo necessário. Quanto menos dados as empresas coletam e armazenam agora, menos problemas inesperados surgirão posteriormente se os dados forem roubados, violados, reaproveitados ou apreendidos por governos. Qualquer processamento de dados também deve ser justo e proporcional.
- As empresas devem ser claras com os usuários sobre quem tem acesso aos seus dados, incluindo dados compartilhados como parte dos contratos de emprego ou matrícula escolar, e adotar políticas de transparência fortes, declarando explicitamente as finalidades e os meios de processamento de dados, e permitindo que os usuários acessar e transferir seus dados com segurança.
- O desenvolvimento e a implantação da tecnologia RE devem ser examinados para identificar e abordar os riscos potenciais aos direitos humanos e garantir transparência, proporcionalidade, justiça e equidade.

Para investidores:

- Os investidores devem avaliar seus portfólios para determinar onde podem investir em tecnologias de RE e usar sua influência para garantir que as empresas do portfólio sigam os padrões de direitos humanos no desenvolvimento e implantação de tecnologias de RE.

Ativistas de direitos digitais e a comunidade de RE em geral têm um papel significativo a desempenhar na proteção dos direitos humanos, como segue:

- Os entusiastas e analistas de RE devem priorizar dispositivos abertos e que levem em conta a privacidade, mesmo que sejam apenas acessórios de entretenimento. Ativistas e pesquisadores

devem concentrar-se na criação de um futuro em que as tecnologias de RE tenham em conta os melhores interesses dos usuários e da sociedade em geral.

- Defensores e ativistas dos direitos digitais devem começar a investigar tecnologias de RE agora e fazer suas demandas serem ouvidas por empresas e reguladores, para que sua experiência possam informar desenvolvimentos e proteções governamentais neste estágio inicial.
- As comunidades de RE devem educar-se sobre as implicações sociais e de direitos humanos das tecnologias que estão desenvolvendo e comprometerem-se com práticas responsáveis.

Nossos dados em RE devem ser usados em nossos próprios interesses, não para prejudicar-nos ou manipular-nos. Não vamos deixar a promessa da próxima geração de computação falhar da mesma forma que a geração anterior. O futuro é amanhã, então vamos torná-lo um futuro em que gostaríamos de viver.

1 Tradução do original **"**Virtual worlds, real people: human rights in the metaverse**",** *https://www.accessnow.org/human-rights-metaverse-virtual-augmented-reality/*

2 *https://www.un.org/en/observances/human-rights-day/know-your-rights*

3 *https://www.eff.org/*

4 *https://www.accessnow.org/what-is-augmented-reality-risks/*

5 *https://www.eff.org/deeplinks/2020/10/augmented-reality-must-have-augmented-privacy*

6 *https://www.eff.org/issues/xr*

7 *https://www.wsj.com/articles/big-tech-seeks-its-next-fortune-in-the-metaverse-11636459200*

8 *https://www.theverge.com/2021/10/20/22736894/fda-vr-tv-movies-treatment-lazy-eye-amblyopia*

9 *https://www.caltech.edu/about/news/virtual-reality-scientists*

10 *https://gupea.ub.gu.se/bitstream/2077/39977/1/gupea_2077_39977_1.pdf*

11 *https://docubase.mit.edu/project/ar-occupy-wall-street/*

12 *https://rd.springer.com/chapter/10.1007/978-3-030-42504-3_16*

13 *https://www.theguardian.com/technology/2021/oct/03/former-facebook-employee-frances-haugen-identifies-herself-as-whistleblower*

14 *https://scholarship.law.vanderbilt.edu/cgi/viewcontent.cgi?article=1000&context=jetlaw*

15 *https://www.eff.org/deeplinks/2020/10/come-back-warrant-my-virtual-house*

16 *https://jessica-outlaw.medium.com/harassment-in-social-vr-stories-from-survey-respondents-59c9cde7ac02*

17 *https://www.accessnow.org/facebook-ray-ban-stories-smart-glasses-privacy-review/*

18 *https://www.eff.org/deeplinks/2020/10/augmented-reality-must-have-augmented-privacy*

19 *https://scholarship.law.vanderbilt.edu/jetlaw/vol23/iss1/1/*

20 *https://scholar.google.com/citations?view_op=view_citation&hl=en&user=DAG_O0EAAAAJ&citation_for_view=DAG_O0EAAAAJ:UeHWp8X0CEIC*

21 *https://www.youtube.com/watch?v=pIpD4-gYImU*

22 *https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3881776*

23 *https://www.eff.org/deeplinks/2021/05/help-bring-dark-patterns-light*

24 *https://necessaryandproportionate.org/principles/*

25 *https://santaclaraprinciples.org/*

26 *https://www.ohchr.org/documents/publications/guidingprinciplesbusinesshr_en.pdf*

27 *https://www.eff.org/deeplinks/2020/10/come-back-warrant-my-virtual-house*

28 *https://www.eff.org/who-has-your-back-2017*

Inteligência Artificial -- O alto custo da captura

Meredith Whittaker -- Professora de Pesquisa Minderoo na Universidade de Nova York (NYU) e diretora do AI Now Institute[1]

Este é um momento perigoso. Sistemas computacionais privados comercializados como inteligência artificial (IA) estão se espalhando por nossa vida e instituições públicas, concentrando o poder industrial, aumentando a marginalização e modelando silenciosamente o acesso a recursos e informações.

Ao considerar como enfrentar essa investida da IA industrial, devemos primeiro reconhecer que os "avanços" em IA celebrados na última década não foram devidos a avanços científicos fundamentais em técnicas de IA. Eles foram e são principalmente o produto de dados significativamente concentrados e recursos de computação que residem nas mãos de algumas grandes corporações de tecnologia. A IA moderna é fundamentalmente dependente de recursos corporativos e práticas de negócios, e nossa crescente confiança nesse tipo de IA concede um poder desordenado sobre nossas vidas e instituições a um punhado de empresas de tecnologia — a "Big Tech". Também dá a essas empresas uma influência significativa tanto na direção do desenvolvimento da IA quanto nas instituições acadêmicas que desejam pesquisá-la. Isso significa que as empresas de tecnologia estão surpreendentemente bem posicionadas para moldar o que sabemos — e não sabemos — sobre IA e os negócios por trás disso, ao mesmo tempo que seus produtos de IA estão trabalhando para moldar nossas vidas e instituições.

Percepções

- O controle da "Big Tech" sobre os recursos de IA tornou as universidades e outras instituições dependentes dessas empresas, criando uma teia de relacionamentos conflitantes que ameaçam a liberdade acadêmica e nossa capacidade de compreender e regular essas tecnologias corporativas.
- Para garantir pesquisa independente e rigorosa, bem como meios de defesa capazes de entender e verificar essas tecnologias e as empresas por trás delas, precisamos organizar-nos, dentro da tecnologia e dentro da universidade.

Examinando a história da influência das forças armadas dos EUA sobre a pesquisa científica durante a Guerra Fria, vemos paralelos com a influência atual da indústria de tecnologia sobre a IA. Essa história também oferece exemplos alarmantes da maneira como o domínio militar dos EUA trabalhou para moldar a produção de conhecimento acadêmico e para punir aqueles que discordaram.

Hoje, a indústria de tecnologia está enfrentando uma pressão regulatória crescente e está aumentando seus esforços para criar narrativas tecnológicas positivas e para silenciar e afastar os críticos da mesma forma que os militares dos EUA e seus aliados fizeram no passado. Como um todo, vemos que o domínio da indústria de tecnologia na pesquisa de IA e na produção de conhecimento coloca pesquisadores e defensores críticos dentro e fora da academia em uma posição vulnerável. Isso ameaça privar as comunidades da linha de frente, os legisladores e o público de conhecimentos vitais sobre os custos e consequências da IA e da indústria responsável por ela — exatamente no momento em que esse trabalho é mais necessário.

Revendo a extensão da influência atual das grandes empresas de tecnologia sobre a IA e as pesquisas de IA, é útil começar com uma breve história da atual virada para a IA. Dado que o campo de IA tem quase 70 anos e passou por vários "invernos de IA", por que a IA cresceu na última década? E do que estamos falando quando falamos sobre IA? Responder a essas perguntas destaca a mutabilidade do termo IA. Também focaliza nossa atenção na centralidade dos recursos corporativos concentrados para o atual boom da IA, e como o controle monopolístico desses recursos deu a um punhado de empresas de tecnologia a autoridade para [re]definir o campo da IA, ao mesmo tempo que encerram o conhecimento sobre os sistemas de IA por trás do segredo das corporações.

Em 2012, uma equipe de pesquisa baseada em Toronto criou um algoritmo chamado AlexNet que venceu o Desafio de Reconhecimento Visual em Grande Escala da ImageNet. Isso marcou um momento importante na história recente da IA e foi um grande negócio na indústria de tecnologia. Ele demonstrou que o aprendizado de máquina supervisionado foi surpreendentemente eficaz no reconhecimento de padrões preditivos quando treinado usando poder computacional significativo e grandes quantidades de dados rotulados[2]. O algoritmo AlexNet dependia de técnicas de aprendizado de máquina que tinham quase duas décadas. Mas não foi o algoritmo que foi um avanço: foi o que o algoritmo poderia fazer

quando combinado com dados e recursos computacionais em grande escala.

AlexNet mapeou um caminho a seguir para grandes empresas de tecnologia que buscam cimentar e expandir seu poder. Os recursos dos quais o sucesso do AlexNet dependia eram os que as grandes empresas de tecnologia já controlavam: vasta infraestrutura computacional, enormes quantidades de dados (e sistemas em funcionamento para processá-los e armazená-los), alcance de mercado consolidado que garantiu a coleta persistente de dados, bem como o capital para contratar e reter talentos escassos. Yoshua Bengio, um dos precursores da pesquisa de IA, colocou de forma simples: "O poder [da computação], a experiência e os dados estão todos concentrados nas mãos de algumas empresas"[3].

O ano de 2012 mostrou o potencial comercial do aprendizado de máquina supervisionado e o poder do termo IA como um gancho de marketing. As empresas de tecnologia rapidamente [re]classificaram o aprendizado de máquina e outras abordagens dependentes de dados como IA, enquadrando-as como o produto de uma inovação científica revolucionária. As empresas adquiriram laboratórios e startups e trabalharam para lançar a IA como uma ferramenta multifuncional de eficiência e precisão, adequada para quase todos os fins em inúmeros campos. Quando dizemos que a IA está em toda parte, é por isso.

A retórica e o capital fluindo dessas empresas serviram para redefinir o campo de pesquisa de IA, inundando-o com financiamento e focando a atenção do campo em técnicas intensivas de computação e dados, além de questões de pesquisa. Laboratórios universitários e startups que queriam desenvolver e estudar IA foram levados a depender de acesso a caros ambientes de computação em nuvem operados por grandes empresas de tecnologia e lutando para acessar dados, uma dinâmica que só se intensificou a partir de 2012. J. Nathan Matias, professor da Cornell e líder do Citizens and Technology Lab, aponta a extensão dessa dependência quando observa que "alguns campos não poderiam existir sem estreitos laços com a indústria"[4].

Isso não significa que os pesquisadores nesses domínios estejam comprometidos. Tampouco significa que não haja direções de pesquisa que possam mitigar tais dependências. Significa, no entanto, que as questões e incentivos que animam o campo nem sempre são os pesquisadores individuais que decidem. E que os termos do campo — incluindo quais perguntas valem a pena responder e quais respostas

resultarão em subsídios, prêmios e estabilidade — são desordenadamente moldados pela virada corporativa para IA intensiva em recursos e pelos incentivos da indústria de tecnologia que a impulsionam.

Um movimento recente da Universidade de Stanford ilustra essa dinâmica. Em agosto de 2021, a universidade anunciou o novo Centro de Pesquisa em Modelos de Fundação (CRFM), cujo lançamento foi acompanhado por um relatório de mais de 100 autores que caracteriza esses modelos como uma "mudança de paradigma" em IA significativa o suficiente para justificar um novo e caro centro de pesquisa[5]. Ao longo do relatório, os modelos de base são considerados inevitáveis, de ponta e resultado do progresso científico.

O que são modelos de fundação? Não é surpresa se você não souber. O nome foi cunhado por Stanford em seu relatório e materiais de lançamento do CRFM, renomeando o que antes se conhecia como modelos de linguagem grandes (LLMs). LLMs — pense em GPT-3 e BERT, entre outros — são algumas das técnicas mais intensivas de dados e computação em IA e, portanto, estão entre as mais capturadas pela indústria. Eles também ganharam muita atenção da mídia recente e foram sujeitos a fortes críticas sobre o viés das técnicas, os custos ambientais e o poder concentrado[6].

Além de simplesmente valorizar as técnicas capturadas pela indústria como de ponta, a *rebranding* de Stanford trabalha para distanciar os LLMs desse legado de crítica. E embora o relatório reconheça que "a pesquisa sobre a construção de modelos de fundação em si ocorreu quase exclusivamente na indústria", ele enquadra as questões de poder concentrado não como questões que deveriam nos fazer reconsiderar a confiança nessas tecnologias, mas sim como problemas que podem ser resolvidos facilitando restrições de controle para que instituições como Stanford também obtenham um pedaço: "A indústria em última análise toma decisões concretas sobre como os modelos de fundação serão implantados, mas também devemos nos apoiar na academia, com sua diversidade disciplinar e incentivos não comerciais"[7].

Os esforços para expandir o acesso à pesquisa de IA também seguem esse padrão, considerando formas de IA de uso intensivo de dados e computação e concentrando-se exclusivamente em como fazer com que mais pessoas tenham acesso a esses recursos concentrados. Ao examinar uma "solução" proposta para esse problema de estrutura restrita, ficamos cara a cara com a extensão da captura da indústria.

Em março de 2020, a Comissão de Segurança Nacional de Inteligência Artificial (NSCAI), presidida pelo ex-CEO do Google Eric Schmidt e dirigida por outros executivos de tecnologia, recomendou que o governo dos EUA financiasse o que chamou de infraestrutura nacional de pesquisa de IA, em nome da "democratização" do acesso à pesquisa de IA. Essa recomendação foi adotada na Lei de Autorização de Defesa Nacional (NDAA) de 2021, que determina a criação de "um sistema que forneça a pesquisadores e estudantes de campos e disciplinas científicas acesso a recursos de computação, co-localizados com conjuntos de dados governamentais e não governamentais publicamente disponíveis e prontos para a inteligência artificial"[8]. Seguindo a diretiva da NDAA, o Escritório de Política Científica e Tecnológica da Casa Branca e a Fundação Nacional de Ciência lançaram recentemente o National AI Research Resource (NAIRR), nomeando uma força-tarefa para arquitetar suas políticas e implementação.

Por que um corpo governamental conflituoso, habitado por executivos de tecnologia, recomendaria "democratizar" o acesso às infraestruturas que estão no centro de seu poder concentrado? Porque essa proposta não reduziria esse poder. Na verdade, se implementada, quase certamente consolidaria e expandiria o poder e o alcance das grandes empresas de tecnologia. O domínio das "Big Techs" sobre a infraestrutura de pesquisa e desenvolvimento de IA vai além de fornecer "plataformas neutras". Essas empresas controlam as ferramentas, os ambientes de desenvolvimento, as linguagens e o software que definem o processo de pesquisa de IA — elas fazem a água em que a pesquisa de IA nada. Mesmo que fosse desejável (o que, dados os danos e falhas da IA, deve ser questionável), não há cenário plausível no qual uma infraestrutura nacional de pesquisa pudesse ser construída de forma significativa fora do atual ecossistema da indústria de tecnologia. Fazer isso exigiria lançar uma nova plataforma, desenvolver software e habituar dezenas de milhares de pesquisadores a novas ferramentas e interfaces, enquanto contrata milhares de engenheiros, desenvolvedores de software, certificadores de garantia de qualidade e pessoal de suporte necessário para manter permanentemente um sistema grande e caro.

Na prática, então, essas propostas para "democratizar" o acesso às infraestruturas de pesquisa de IA equivalem a apelos para subsidiar ainda mais os gigantes da tecnologia, licenciando infraestrutura conhecida dessas empresas de forma que lhes permita continuar a definir os termos e condições da IA e da pesquisa de IA. Ao mesmo tempo, centros como o novo CRFM de Stanford estão preparados para consolidar ainda mais esse

domínio, apresentando técnicas de IA dependentes da indústria como a vanguarda da pesquisa em IA.

De programas de doutorado patrocinados pela indústria a iniciativas que colocam escritórios de empresas de tecnologia literalmente no meio das universidades, ou a parcerias entre a National Science Foundation e a Amazon para definir os parâmetros de "justiça" em IA e conceder bolsas para aqueles que atendem aos seus critérios positivistas[9], vemos uma infinidade esquemas para aproximar a academia das empresas de tecnologia. Isso se estende a acordos de dupla afiliação, que são comuns no campo de IA e equivalem a empresas que contratam professores de IA, permitindo-lhes manter seus títulos acadêmicos e cargos. Acadêmicos com dupla afiliação recebem o salário de uma empresa de tecnologia, trabalham em estreita colaboração com funcionários de tecnologia e se valem da infraestruturas de pesquisa corporativa, ao mesmo tempo que publicam pesquisas sob o aval da universidade. Esses acordos ajudam a proteger as empresas de acusações de que estão contribuindo para a fuga de cérebros ao contratar pesquisadores de universidades. Eles também permitem que as empresas recrutem profissionais para responder a perguntas interessantes para empresas de tecnologia, ao mesmo tempo que criam a aparência de disciplinas acadêmicas que investem orgânica e independentemente nessas mesmas questões.

O fato de esses arranjos conflitantes serem tratados como prática padrão provavelmente está relacionado à clareza com que os pesquisadores de IA e as universidades reconhecem sua dependência de grandes empresas e dos recursos que controlam. Maja Pantic, professora de aprendizado de máquina que trabalha para a Samsung e tem dois cargos no Imperial College London, disse ao Financial Times que "simplesmente não poderia continuar trabalhando apenas na academia, não temos os recursos de computação, eu não poderia pagar pessoas para trabalharem para mim e eu não tinha dinheiro para criar poder de processamento "[10]. Ela e muitos outros enfrentam a escolha de se aliar a uma empresa, com todas as condições tácitas que essa dependência exige, ou de ser incapazes de fazer o tipo de trabalho que iguala prestígio e sucesso acadêmico.

A extensão da influência da indústria de tecnologia sobre o domínio de pesquisa de IA tem paralelos com o domínio das forças armadas dos EUA sobre a pesquisa científica durante a Guerra Fria. As empresas de tecnologia estão se inspirando em um manual semelhante.

Escrevendo em 1946, logo após o término da Segunda Guerra Mundial, o General Dwight D. Eisenhower redigiu um memorando intitulado "Recursos Científicos e Técnicos como Patrimônios Militares" que propunha trazer cientistas e pesquisadores mais diretamente para o planejamento militar dos EUA, argumentando que isso permitiria aos militares construir confiança com os cientistas, para ter um assento na primeira fila para novos desenvolvimentos científicos e — por meio de financiamento e proximidade colegial — deixar os militares dos EUA conduzirem as questões de pesquisa de forma a garantir que os cientistas estejam "familiarizados com nossos problemas fundamentais"[11]. Três anos após o memorando, em 1949, os EUA obtiveram evidências de que a União Soviética estava testando armas nucleares. Isso ajudou a colocar o plano de Eisenhower em ação, catalisando a criação de escritórios e agências de pesquisa em ramos militares dedicados a financiar e moldar a pesquisa[12].

Relevante para o nosso caso é o poder que isso deu aos militares dos EUA na orientação da pesquisa científica e das instituições que a abrigavam. Essa influência foi aplicada não apenas para garantir que a pesquisa acadêmica fosse animada por questões e preocupações militares dos EUA, mas também para punir denunciantes, controlar a dissidência e incentivar a complacência em face de afirmações exageradas mascaradas pela autoridade científica. É aqui, nessas histórias mais sombrias, que confrontamos o alto custo da captura — seja militar ou industrial — e suas implicações perigosas para a liberdade acadêmica e a produção de conhecimento capaz de responsabilizar o poder.

Aldric Saucier era um cientista que trabalhava para o Exército dos EUA na controversa Iniciativa de Defesa Estratégica (IDE/SDI) do então presidente Ronald Reagan. A IDE foi uma iniciativa militar maciça que recrutou cientistas de todo o país na tentativa de construir um escudo de mísseis balísticos. A proposta era fantástica, e muitos na comunidade de pesquisa a consideraram cientificamente infundada e com probabilidade de aumentar as chances de uma guerra nuclear. Quando Saucier relatou desperdício, fraude e hipérbole dentro do programa, o então secretário de Defesa Dick Cheney supervisionou sua demissão, junto com uma campanha para desacreditar publicamente sua experiência científica[13]. Fora dos laboratórios de pesquisa comandados por militares, os dissidentes também foram ameaçados. Cientistas de universidades organizaram um boicote à pesquisa e ao financiamento da IDE. Em resposta, o congressista de Indiana, Dan Burton, ameaçou cortar fundos para universidades onde os professores recusavam bolsas relacionadas à

IDE. Enquanto isso, a liderança da universidade no Laboratório Nacional Lawrence Livermore, há muito aliada às visões nucleares dos militares dos EUA, trabalhou para expulsar o físico Hugh DeWitt, que se manifestou contra o papel do laboratório em exacerbar a corrida armamentista. Enquanto DeWitt conseguiu manter sua posição, a ele foram negado aumentos e promoções e foi excluído de pesquisas relevantes[14]. O subsecretário de Defesa para Pesquisa e Engenharia, Donald Hicks — na época encarregado dos contratos de pesquisa do Pentágono — chegou a intimidar publicamente os pesquisadores. Em uma entrevista, Hicks afirmou que, embora os professores pudessem falar em um "país livre", eles também eram "livres para manter a boca fechada... Eu também sou livre para não lhes dar dinheiro"[15]. O Wall Street Journal publicou um editorial aplaudindo Hicks.

Com o benefício de uma retrospectiva, sabemos que os críticos da IDE estavam amplamente corretos sobre as falhas e a lógica perigosa que impulsionava o programa. Mas seus argumentos e análises apoiados em evidências não os salvaram de retaliação, coerção financeira e difamação, mesmo dentro de instituições supostamente dedicadas à liberdade acadêmica.

O fato de um punhado de grandes empresas de tecnologia atualmente ter influência semelhante em relação à pesquisa de IA deve nos alarmar, especialmente diante das evidências crescentes do desejo da tecnologia de moldar uma narrativa positiva em resposta à crescente pressão regulatória e pública, ao lado da clara disposição da indústria de silenciar e punir críticos. Os exemplos abundam, desde a revogação do acesso de dados do Facebook a pesquisadores da NYU que examinaram o papel da empresa na insurreição de 6 de janeiro; ao Google, instruindo os pesquisadores internos a "atingir um tom positivo" em suas descobertas[16], enquanto direciona "aliados acadêmicos" externos para levantar questões sobre a intervenção regulatória[17]; aos ataques capciosos da Amazon contra jovens pesquisadores negros que revelaram lógicas racistas em seus produtos, enquanto retaliam contra os trabalhadores que se organizaram contra os danos climáticos da empresa. O Google também demitiu Timnit Gebru, depois de exigir que ela e seus co-autores removessem seus nomes de um artigo crítico dos LLMs que são essenciais para o roteiro de produtos do Google, e que Stanford recentemente reformulou e revalorizou. A lista continua, fornecendo um bom barômetro de onde essas empresas traçam os limites — pesquisas e divergências que ameaçam o crescimento e a receita.

Além de punir os dissidentes e desmoralizar as pesquisas que consideram ameaçadoras, as empresas de tecnologia estão trabalhando para cooptar e neutralizar as críticas. Elas fazem isso em parte financiando e elevando seus críticos mais fracos, muitas vezes instituições e coalizões que se concentram na chamada ética da IA, e enquadram questões de poder e domínio da tecnologia como questões de governança abstratas que tomam a forma atual da indústria de tecnologia como um dado e a proliferação da IA como inevitável. Paralelamente, as empresas de tecnologia também defendem remédios tecnocráticos como "recompensas contra os preconceitos sobre a IA" e correções de justiça que tratam a discriminação habilitada pela tecnologia como um problema de código ruim e engenharia "bugada"[18]. Essas abordagens resultam em ótimas relações públicas. Elas também servem para definir os engenheiros de elite como árbitros do "preconceito", enquanto excluem estruturalmente acadêmicos e defensores que não têm treinamento em ciência da computação, mas cujo foco nas assimetrias de poder racializadas e na economia política da IA são essenciais para compreender e abordar os malefícios da IA.

Tudo isso está acontecendo em um cenário em que as instituições acadêmicas, cada vez mais operando como empresas em busca de grandes investidores, têm dificuldade em ignorar as vantagens financeiras e de reputação que as parcerias de tecnologia e financiamento trazem. Essa dinâmica é agravada pela crescente precarização dos empregos acadêmicos, nos quais cada vez menos acadêmicos têm a segurança no emprego ou a solidariedade sindical necessária para contestar com segurança políticas que possam comprometer a liberdade acadêmica. Isso dá às empresas de tecnologia uma alavancagem crescente não apenas sobre as pesquisas que financiam diretamente, mas também sobre as decisões sobre quais trabalhos serão incluídos e excluídos na universidade em geral.

Também não podemos ignorar o ataque em andamento contra o trabalho que revela o racismo e a desigualdade estruturais. Os think-tanks de extrema direita e os apparatchiks republicanos estão pressionando as instituições educacionais a eliminarem a pedagogia e a pesquisa com foco na justiça racial, que eles descartam levianamente sob o termo *teoria crítica da raça*. Esse ataque à liberdade intelectual é importante por muitas razões. A prática acadêmica e de movimentos atenta ao capitalismo racial e ao racismo estrutural forneceu muitos dos métodos e estruturas essenciais para o trabalho crítico envolvendo as implicações sociais da tecnologia. Ajudou a enfocar a crítica tecnológica além de noções

superficiais de preconceito para exames das maneiras pelas quais essas tecnologias replicam padrões de marginalização racial e concentram o poder nas mãos daqueles com recursos escassos e caros para desenvolver e implantar IA. Essa linha de crítica já influenciou fortemente o discurso público e a agenda regulatória global de uma forma que as empresas de tecnologia estão resistindo ativamente.

Então, qual é o caminho a seguir? Para começar, acadêmicos, defensores e formuladores de políticas que produzem e contam com trabalho crítico de tecnologia devem confrontar e caracterizar a dinâmica da captura, cooptação e compromisso da tecnologia, e logo. Isso significa incorporar críticas reflexivas às condições e à criação do conhecimento, e aos compromissos e compensações enfrentados pelos trabalhadores do conhecimento sobre os quais as instituições interessadas têm poder. Dada a política de proximidade cordial que informa as redes de prestígio acadêmico enquanto trabalha para confundir os limites entre os trabalhadores acadêmicos e da indústria, isso certamente será desconfortável. Mas caracterizar essas dinâmicas é a única maneira de abordá-las e de pleitear questões que nos permitam imaginar e exigir futuros alternativos.

Este é exatamente o tipo de intervenção que está ameaçada pela captura da indústria de pesquisas em IA. Então, como apoiar esse trabalho crítico e proteger aqueles que o fazem dentro e fora da academia?

Aqui nos voltamos para o papel central dos trabalhadores de tecnologia organizados, aqueles que fizeram incursões em toda a indústria nos últimos cinco anos e trabalhadores acadêmicos organizando-se em um ambiente onde o mito do gênio individual serve para sustentar a desigualdade, mesmo quando o mercado de trabalho desmorona. A luta dos trabalhadores acadêmicos contra a precariedade da profissão é também uma luta pela liberdade acadêmica. Oportunidades de carreira estáveis e controle mais democrático sobre a universidade ajudariam a desviar o equilíbrio da influência de doadores ricos e patrocinadores da grande indústria. Os trabalhadores de tecnologia organizados, por sua vez, têm um papel a desempenhar na verificação por dentro do poder de sua indústria, lutando por mais controle sobre o trabalho que fazem e trabalhando para conter a influência de seus empregadores na academia e além. Nessa capacidade, poderíamos imaginar pesquisadores e cientistas organizados exigindo um redirecionamento das generosas dotações do Congresso dos EUA que atualmente subscrevem a infraestrutura de

pesquisa de IA nacional, usando sua experiência e posição para exigir doações em apoio a universidades, escolas de trabalhadores, verdadeiramente públicas e acessíveis, e programas que integram comunidades com experiência vivida no panteão de alunos e especialistas interrogando tecnologia[19]. Claro, dado que a indústria de tecnologia não apenas redige políticas por meio de conselhos nomeados pelo Congresso, como o NSCAI, mas também gasta mais do que as indústrias do petróleo e do tabaco em lobby, está claro que qualquer intervenção desse tipo exigirá uma séria luta organizada.

Um futuro em que o Congresso dos Estados Unidos apoie de modo decidido o trabalho crítico verdadeiramente democrático e independente não parece próximo no horizonte. Mas a organização dentro da academia e dos locais de trabalho de tecnologia também pode nos ajudar a proteger a nós mesmos e ao interesse público no curto prazo, preparando-nos para defendermos uns aos outros diante da pressão institucional e desenvolvendo fortalezas de cuidado e responsabilidade mútua que nos permitem caracterizar as dinâmicas de coerção e captura com mais segurança. Isso não será fácil; exigirá o confronto de culturas de competição e território que marcam tanto a tecnologia quanto os locais de trabalho acadêmicos. Mas as apostas são muito altas, e aqueles que pretendem moldar o que sabemos (e não sabemos) sobre IA e a indústria responsável são bem organizados e com muitos recursos. Em suma, esta é uma batalha de poder, não apenas uma disputa de ideias, e estar certos sem a estratégia e a solidariedade para defender nossa posição não nos protegerá.

Referências

1 Meredith Whittaker é Professora de Pesquisa Minderoo na Universidade de Nova York (NYU) e diretora do AI Now Institute. Sua pesquisa se concentra nas implicações sociais da IA e da indústria responsável por ela. Antes de ingressar na NYU, ela fundou e liderou o Grupo de Pesquisa Aberta do Google. mw3900@nyu.edu. Copyright detido pela autora. Direitos de publicação licenciados para a Association for Computing Machinery. Versão em português autorizada pela autora. A autora agradece a Theodora Dryer, Lilly Irani, Nantina Vgontzas, Sarah Myers West e Michael Decker por sua leitura cuidadosa e generosa e sugestões, e a Nicole Weber pela excelente assistência em pesquisa. A edição rigorosa e cuidadosa de J. Khadijah Abdurahman também foi fundamental para a versão final deste texto.

2 Krizhevsky, A., Sutskever, I., and Hinton, G.E. ImageNet classification with deep convolutional neural networks. *Advances in Neural Information Processing Systems* 25 (NIPS 2012); *https://proceedings.neurips.cc/paper/2012/file/c399862d3b9d6b76c8436e924a68c45b-Paper.pdf*

3 Murgia, M. AI academics under pressure to do commercial research. *Financial Times*. 13-03-2019; *https://www.ft.com/content/94e86cd0-44b6-11e9-a965-23d669740bfb*

4 Matias, J.N. Why we need industry-independent research on tech & society. *Citizens and Tech Lab*. Janeiro 2020; *https://citizensandtech.org/2020/01/industry-independent-research/*

5 Bommasani, R. et al. On the opportunities and risks of foundation models. 2021; arXiv preprint arXiv:2108.07258

6 Bender, E.M., Gebru, T., McMillan-Major, A., and Shmitchell, S. On the dangers of stochastic parrots: Can language models be too big? *Proc. of the 2021 ACM Conference on Fairness, Accountability, and Transparency*. ACM, Nova York, 2021, 610-623; *https://dl.acm.org/doi/10.1145/3442188.3445922*

7 Bommasani, R. et al. *op.cit.*

8 *https://www.congress.gov/116/bills/hr6395/BILLS-116hr6395enr.pdf*

9 *https://www.nsf.gov/funding/pgm_summ.jsp?pims_id=505651*

10 Murgia, M. *op.cit.*

11 Melman, S. *Pentagon Capitalism: The Political Economy of War*. McGraw-Hill, New York, 1970.

12 Krinsky, R. *Swords and sheepskins: Militarization of higher education in the United States and prospects of its conversion*. Bulletin of Peace Proposals 19, 1 (1988), 33-51; *http://www.jstor.org/stable/44481371*

13 Lardner, Jr., G. Army accuses SDI critic of falsifying credentials. *Washington Post*. 14-04-1992; *https://www.washingtonpost.com/archive/politics/1992/04/14/army-accuses-sdi-critic-of-falsifying-credentials/13ffe75f-50f8-4654-9027-536c30880c13/*

14 Martin, B. Science: contemporary censorship. In *Censorship: A World Encyclopedia*, Vol 4. D. Jones, ed. Fitzroy Dearborn, Londres, 2001, 2167–2170; *https://documents.uow.edu.au/~bmartin/pubs/01cescience.html*

15 Hiatt, F. Official seeks like minds in 'Star Wars'. *Washington Post*. 13-05-1986; *https://www.washingtonpost.com/archive/politics/1986/05/13/official-seeks-like-minds-in-star-wars/28ddbdc0-d55f-4cac-a742-b17176e6adcf/*

16 Dave, P. and Dastin, J. Google told its scientists to 'strike a positive tone' in AI research – documents. *Reuters*. 23-12-2020; *https://www.reuters.com/article/us-alphabet-google-research-focus/google-told-its-scientists-to-strike-a-positive-tone-in-ai-research-documents-idUSKBN28X1CB*

17 Satariano, A., and Stevis-Gridneff, M. Big tech turns its lobbyists loose on Europe, alarming regulators. *New York Times*. 14-12-2020; *https://www.nytimes.com/2020/12/14/technology/big-tech-lobbying-europe.html*

18 Vanian, J. Why Microsoft and Twitter are turning to bug bounties to fix their A.I. *Fortune*. 10-08-2021; *https://fortune.com/2021/08/10/why-microsoft-and-twitter-are-turning-to-bug-bounties-to-fix-their-a-i/*

19 Quero agradecer especialmente a Lilly Irani e Nandina Vgontzas por seu cuidado e generosidade intelectual, que me ajudaram a mapear esse conjunto de pontos.