

Roteiro esquemático dos relatórios dos workshops

a) Informações sobre a atividade:

O risco do algoritmo: o uso de relatórios de impacto à proteção de dados pessoais no combate à discriminação algorítmica

Proponentes e co-proponentes: Laboratório de Políticas Públicas e Internet (LAPIN), terceiro setor.

Palestrantes:

Academia: Edson Prestes e Silva Junior (Universidade Federal do Rio Grande do Sul)

Professor da Universidade Federal do Rio Grande do Sul; Líder do Phi Robotics Research Group e do CNPq Research Fellow. Membro Sênior da IEEE Robotics and Automation Society (IEEE RAS) e da IEEE Standards Association (IEEE SA).

Sociedade Civil: Tarcízio Silva (Fundação Mozilla)

Tarcízio Silva é Tech + Society Fellow na Fundação Mozilla, alocado na organização Ação Educativa. Mestre em Comunicação (UFBA); doutorando em Ciências Humanas e Sociais (UFABC). Autor do livro "Racismo Algorítmico" (Edições Sesc, 2022).

Setor Público: Thiago Moraes (Autoridade Nacional de Proteção de Dados - ANPD)

LLM Law & Technology (Tilburg University), Mestre em Ciências da Informação, Bacharel em Direito e em Engenharia de Redes (UnB). Encarregado de Proteção de Dados Pessoais e Coordenador de Tecnologia e Pesquisa na Autoridade Nacional de Proteção de Dados (ANPD). CIPP/E, CIPM.

Setor Privado: Maria Cecília Oliveira Gomes (idwall)

Advogada especializada em Privacidade e Proteção de Dados. Encarregada e Coordenadora Sênior de Privacidade e Proteção de Dados na Idwall. Doutoranda em Filosofia e Teoria Geral do Direito na USP. Professora do Data Privacy Brasil. Professora convidada da USP, PUCRS e de outras instituições de ensino.

Moderador(a): Victor Mulin - Lapin (mini bio) Mestre em Direito e Informática (Universidade do Minho) e Bacharel em Direito (UCAM). Servidor Público do Tribunal de Justiça de Minas Gerais e Pesquisador em Proteção de Dados e Inteligência Artificial pelo LAPIN.

Relator(a): Mariana Freitas - Lapin - Graduada em Ciência Política, atua em relações governamentais na área de Tecnologia e Finanças. Pesquisadora no Lapin no grupo de Governança e apoio legislativo aos times do laboratório.

b) Estruturação do workshop

Objetivos e resultados:

O objetivo do painel foi abordar o fenômeno da discriminação algorítmica no cenário brasileiro. Assim, o painel possuía os seguintes objetivos:

Como ocorre a discriminação algorítmica; qual o impacto real da discriminação algorítmica na vida das pessoas; e qual a importância de relatórios de impacto de proteção de dados na avaliação e mitigação de riscos aos indivíduos, em particular nos contextos de discriminação algorítmica.

Assim, foi priorizada a cooperação e a interação de pontos de vista multidisciplinares, o que ocorreu com a integração da fala dos painelistas, que foram se complementando e permitindo o entendimento do processo que ocorre para a discriminação algorítmica, que era o terceiro objetivo do painel.

Os objetivos foram alcançados paulatinamente, primeiramente com a explicação técnica do que é a Inteligência Artificial e como ocorre o processo para a formação da discriminação algorítmica, com a primeira fala de Edson Prestes. Em seguida, houve a exposição de como a discriminação algorítmica se apresenta para o mundo real e qual o real impacto na indivíduos, com a fala de Tarcízio Silva. Houve ainda, uma compreensão de como o problema tem sido abordado globalmente e como o relatório de impacto tem se encaixado nesse contexto, com a fala de Thiago Moraes. E a finalização, que se deu com a fala de Maria Cecília, abordou o contexto histórico da ferramenta do RIPD, e as abordagens e funções pretendidas em relação a esse contexto.

Justificativa em relação à governança da Internet:

O painel abordou de maneira multidisciplinar o tema da discriminação algorítmica, passando por explicações que foram desde como funcionam os sistemas de IA até as abordagens jurídicas que diferentes nações tem feito para conter efeitos negativos que a IA pode vir a provocar.

A discussão foi necessária para demonstrar para o público, desde o que mais possui conhecimento até os que ainda não haviam tido contato com a temática, como os sistemas de IA têm expandido seu uso e assumido um protagonismo nas relações diárias.

O painel ainda foi realizado em uma janela de oportunidade, uma vez que o Brasil se encontra em um momento de construção regulatória com votação de um projeto de lei que regula a IA no Brasil. Assim, a discussão sobre a discriminação algorítmica se demonstrou necessária e que gerou frutos, uma vez que auxiliou na construção da contribuição à Comissão de Juristas de IA no Senado.

Por fim, levando em conta os princípios e as boas práticas para a boa governança da Internet, o painel deu foco ao RIPD, demonstrando como tal documento pode se tornar aparato importante

na construção de um ambiente jurídico seguro para todos os agentes envolvidos, uma vez que pode vir a proporcionar uma ponte entre questões de tecnologia e do ordenamento jurídico.

Metodologia:

Foi realizado um painel de cerca de 1h30 de duração, que permitiu com que cada painalista realizasse uma fala de cerca de 10 minutos, dando ao público conhecimento a ensejar.

Foi então aberto espaço para perguntas, que foi realizada de maneira híbrida, permitindo que tanto o público presente como o público virtual conseguissem enriquecer o debate. Devido ao número de inscritos, foi realizado um bloco de perguntas e assim os painelistas puderam respondê-las de forma a relacionar e sumarizar as questões abordadas. As perguntas que conseguiram abordar todos os participantes e permitiram o aprofundamento desde as questões técnicas até as que abordavam a necessidade de se pensar em questões sociais dentro de áreas técnicas da tecnologia.

Assim, cada palestrante realizou suas conclusões finais e foi levantado o ponto sobre a participação social na consulta pública do Senado Federal, mais especificamente na comissão de juristas para Inteligência Artificial, ponto importante para a equipe do Lapin, uma vez que pretendemos permitir que o debate se estende-se para além do tempo do debate.

Houve a divulgação do painel nas redes sociais e, após o evento, foram colhidos feedbacks e houve um retorno extremamente positivo do ponto de vista da integração de diferentes grupos de pessoas impactadas em buscar conhecimento sobre o tópico do painel.

c) Síntese dos debates

1) Edson Prestes

Edson Prestes iniciou questionando o que seria a discriminação algorítmica, pontuando que existe uma certa confusão no que seria a discriminação em termos de algoritmos, ou seja, se seria o código o agente discriminador ou não.

Prestes pontuou que é preciso partir do ponto que estamos lidando com computadores, que compreendem apenas símbolos, ou seja, nas palavras do professor, máquinas que não compreendem racionalmente o que estão fazendo. Assim, defendeu que não se pode atribuir uma característica inerentemente humana, de discriminar outro indivíduo, a um computador.

O professor seguiu explicando que para combater a discriminação é preciso olhar para os dados que são inseridos em um sistema de IA. Pontuou que existem os dados de entrada e os de saída, e a partir disso, se descobre uma função que faz o caminho da entrada para a saída. Explicou que a resposta de um sistema de IA é referente aos dados que alimentam o mesmo, assim, dados viesados, que são criados e colhidos pelas próprias pessoas, que devem ser o grande foco ao se pensar discriminação por algoritmos.

Edson Prestes chamou a atenção, entretanto, para a opacidade dos modelos que não permitem a compreensão de como os processos internos do sistema ocorrem. Isso permite que sistemas deem 99% de acurácia, o que em um primeiro momento pode parecer satisfatório, porém o que se encontra no 1% restante pode ser extremamente danoso a depender do tipo de aplicação que está sendo desenvolvido. A exemplo seriam sistemas que realizam a contratação de pessoas ou que identifiquem terremotos.

O painelista ainda ressalta que, com tais dados de entrada, os sistemas podem reproduzir padrões discriminatórios já existentes na sociedade por meio de associações e inferências. E realiza o questionamento se a preocupação com dados pessoais sensíveis é suficiente.

2) Tarcízio Silva

Tarcízio Silva fez sua apresentação em torno de uma percepção da justiça epistêmica e da categoria de racismo algorítmico, a qual rejeita a ideia de um algoritmo racista. Silva ressaltou que é preciso observar esse posicionamento de um ponto em que relações de poder estão incorporadas nos algoritmos e suas aplicações, e que são pouco vistas devido a opacidade dos sistemas.

O pesquisador abordou sobre os impactos discriminatórios da inteligência artificial, como o aumento de casos de racismo, misoginia e classismo. Assim, ressaltou que a sociedade precisa começar a discutir sobre a redistribuição de métodos e do aumento da capacidade de investigação frente a esses casos mencionados. Silva abordou a diferença entre o número de estudos que propõem novos modelos de Inteligência Artificial e estudos que propõem a análise e escrutínio destes modelos. Há também uma desproporcionalidade entre a capacidade de investimento privado em pesquisa em relação à capacidade de investimento de instituições acadêmicas, principalmente do setor público. Outro ponto necessário de espaço para conversa seria a promoção de um alinhamento entre o setor privado e os pesquisadores quanto à questão da ética.

Tarcízio defendeu a necessidade de que os reguladores repensem os limites da aplicação da autorregulação, já que, em suas palavras, há um alto índice de má fé por parte das *big techs* e *start-ups*.

Em relação à opacidade, outra questão a ser observada é que a opacidade algorítmica como crença de que a tecnologia é neutra se vincula à negação do racismo, caracterizando uma dupla-opacidade. E assim, Tarcízio defendeu que migremos da ideia de ética dos dados, para justiça dos dados, de forma a referenciar pesquisadores e eticistas de origem feminista e decolonial de forma posicionar o estudo da história, cultura e contexto como questões centrais para o debate.

Sugeriu que as avaliações de risco sejam implementadas no planejamento e no uso da IA, permitindo a classificação entre os diferentes sistemas. Já os relatórios de impacto seriam voltados para os sistemas já implementados. Tarcízio ressaltou a importância da participação de desenvolvedores, comunidades impactadas, *experts* de domínio, especialistas legais, autoridades legais, dentre outros *stakeholders*.

Tarcizio Silva fez a observação de que a opacidade gera uma injustiça epistêmica que seria o “dano a alguém especificamente em sua capacidade de conhecedor” conforme proposto por Miranda Fricker.

Por fim, diante da justiça epistêmica, as análises de riscos e impactos devem levar em consideração a regulação e os compromissos das empresas devem estar voltados aos direitos. Com isso, as avaliações devem englobar o sistema sociogênico como um todo. Sugeriu a existência dos chamados “Experts de Domínio” como atores essenciais no planejamento e na análise dos relatórios de risco e de impacto, bem como a necessidade de representação de titulares de dados e grupos impactados.

3) Thiago Moraes

Representante do setor público, Thiago Moraes inicia seu posicionamento afirmando que não pode haver distinção entre uma abordagem baseada em risco e uma abordagem baseada em direitos. De modo que o uso dos relatórios de impacto de proteção de dados não devem ser usados apenas como um meio de mitigar risco inerentes aos ativos de negócio, mas sim visando a proteção dos direitos das pessoas afetadas.

Ultrapassado esse ponto, Thiago Moraes apresenta um panorama internacional relativo a propostas feitas por governos e organizações de relatórios/avaliações de impacto a serem utilizados em sistemas de inteligência artificial, sendo elas:

Canadá - Avaliação de Impacto Algorítmico: destinada a identificar riscos e avaliar impactos em uma ampla gama de áreas, incluindo: os direitos de indivíduos ou comunidades; a saúde ou o bem-estar de indivíduos ou comunidades; e a sustentabilidade contínua de um ecossistema.

NIST - Moldura para gestão de riscos de IA (ainda em elaboração): destinado a identificar riscos às pessoas (liberdades e direitos fundamentais, direitos coletivos e direitos sociais), às instituições (sistemas técnicos e operações de negócio); e ao sistema (sistema financeira ou cadeia de suprimento global, por exemplo).

UNESCO - Avaliação de Impacto Ética: destinada a identificar e avaliar benefícios, preocupações e riscos da IA, bem como a prevenção, mitigação e monitoramento de riscos sobre os direitos humanos e as liberdades fundamentais, em particular, mas não limitado à/ao(s): direitos de grupos marginalizados; pessoas vulneráveis e pessoas em situação de vulnerabilidade; direitos trabalhistas; meio ambiente e ecossistema; implicações éticas e sociais; e participação democrática.

OCDE - Moldura de Avaliação de Risco (risk assessment framework): onde, baseado no princípio 1.4 (c) da OCDE, os atores de IA devem, com base em seus papéis, contexto e capacidade de agir, aplicar uma abordagem de

gerenciamento de risco para cada fase do ciclo de vida do sistema de IA de forma contínua para abordar riscos relacionados a sistemas de IA, incluindo privacidade, segurança digital, segurança física e discriminação. A proposta da OCDE leva em consideração 05 vetores, em que deve-se avaliar os riscos que podem afetar o vetor "Pessoas e Planeta" a partir dos demais vetores, sendo eles o "Contexto Econômico", os "Dados entradas", as "Tarefas e Saídas" e o "Modelo de IA".

Por fim, Thiago Moraes apresenta o Relatório de Impacto de Proteção de Dados, previsto no artigo 5, XVII da Lei Geral de Proteção de Dados (LGPD), que é um documento que deve gerir os riscos às liberdades civis e aos direitos fundamentais, bem como definir medidas, salvaguardas e mecanismos de mitigação de risco.

Thiago aproveita para mencionar que, apesar de não ter sido desenvolvido especificamente para análise de sistemas de IA, o RIPD está em destaque atualmente por conta da LGPD e do crescente número de aplicações em IA. A partir disso, Thiago Moraes inicia uma comparação entre o nosso RIPD e a proposta da OCDE de Avaliação de Riscos, buscando entender eventual compatibilidade entre os dois instrumentos.

Através de uma comparação entre as exigências do RIPD e os vetores da proposta da OCDE, Thiago destaca que o RIPD possui certas similaridades, como a proteção das "Pessoas" e a análise do "Contexto Econômico", dos "Dados e Entradas" e das "Tarefas e Saídas". Contudo, ao não levar em consideração os "Modelos de IA" e a própria proteção do "Planeta", Thiago concluiu que o RIPD não é maduro o suficiente para liderar as análises dos riscos provocados por sistemas de IA. Dessa forma, o RIPD poderia ser um ensaio para aprofundar maiores discussões e buscar, durante as discussões de regulação de IA no Brasil, desenvolver um instrumento válido e apto para proteger liberdades e direitos fundamentais que podem estar em risco mediante o uso crescente de aplicações de IA.

Por fim, foi abordada a agenda regulatória da ANPD para o tema do RIPD, mencionou que já foi realizada uma tomada de subsídios sobre o tema e que o RIPD é um tema bastante desafiador porque se conecta com vários outros temas.

4) Maria Cecília Oliveira Gomes

Representante do setor privado, Maria Cecília inicia sua apresentação a partir da seguinte provocação: o RIPD pode ou não ser um instrumento de resolução de problemas relacionando à discriminação algorítmica e aos demais riscos da IA?

Maria Cecília iniciou a sua fala com o panorama histórico das avaliações de impacto, que nasceram nas décadas de 60 e 70 com uma proposta de trabalho numa perspectiva ex-ante em relação à prevenção de riscos associados a um projeto, uma política, alguma tomada de decisão necessária. Neste sentido, as avaliações de impacto foram tomando forma em diferentes áreas ao longo das décadas, como os relatórios de impacto ambiental. E a partir da

década de 90, passamos a ver esse instrumento dialogar com regulações de proteção de dados.

Com isso, Maria Cecília retoma que a proposta destes instrumentos se desenvolveu em torno da sua função de prevenção, a qual se diferencia da mitigação. Maria Cecília afirma que o RIPD pode ser uma das ferramentas para solução de problemas, mas não a única. Na visão da Maria Cecília, o RIPD é útil na medida em que ele possibilita ter uma 'fotografia' geral dos riscos envolvidos no uso de uma tecnologia ou aplicação específica, podendo assim analisar os riscos que podem ser jurídicos, éticos, técnicos, regulatórios e os riscos residuais.

Levando o assunto para o ponto de vista regulatório, pois o Brasil está discutindo um Marco Legal para a IA, Maria Cecília converge com Thiago Moraes afirmando que este é o momento de definirmos a estrutura e a metodologia de uma avaliação de impacto para a inteligência artificial que vá além dos riscos do negócios e proteja os direitos das pessoas afetadas. Para a painelistas, na questão metodológica não há uma dicotomia entre uma abordagem baseada em riscos e uma abordagem baseada em direitos, porque a abordagem baseada em riscos busca prevenir riscos a direitos.

Maria Cecília aproveita para sensibilizar sobre a transparência necessária ao lidar com riscos aos direitos, ressaltando a importância (e possibilidade) de publicizar as avaliações de impacto, para que haja um diálogo em relação à prevenção e para que haja um escrutínio público na forma como a tecnologia está sendo avaliada.

Perguntas:

- 1) Quais medidas empresarias de combate à discriminação algorítmica poderiam ser adotadas? Um conselho dentro da empresa? Uma ouvidoria especializada??**

Respostas diretas para essa pergunta:

Maria Cecília alerta, pela visão do setor privado, que além da pluralidade necessária para lidar com tais riscos, também é importante iniciar um processo de conscientização dentro das empresas e times de produto envolvidos.

Já **Tarcízio Silva** se posicionou cético quanto às questões éticas sendo praticadas por parte das empresas, mas mencionou que códigos de condutas e princípios corporativos para diversas situações têm sido um caminho, além do envolvimento de experts éticos e de domínio independentes.

- 2) Como a ANPD atualmente vê a questão da publicização dos RIPD? Existem conteúdos que não devem ser publicizados? Pensando na acessibilidade dos titulares de dados há esse material, existe alguma preocupação em relação ao design?**

Resposta direta para essa pergunta:

Thiago Moraes lembrou que esse é um dos pontos que foi perguntado na tomada de subsídios, e que assim que a minuta foi liberada, esse ponto voltará ao debate. **Maria Cecília** relembrou que em ordenamentos jurídicos internacionais sobre proteção de dados, a publicização não é obrigatória por lei, mas é considerada uma boa prática. Nos EUA, no âmbito federal, a publicação dos relatórios é obrigatória para a administração pública. No panorama atual das discussões no Brasil percebe-se que, para o setor público seria obrigatória uma publicização, enquanto para o setor privado não se tem consenso. Neste sentido, ainda não temos um panorama uniforme e harmônico em relação à quando a publicização deve ser realizada e de que forma ela deve ser realizada.

Tarcizio Silva pontuou sobre a complexidade no tópico da publicização do relatório, porém defendeu que o documento pode vir a público levando em conta os níveis de risco e o setor em questão, a exemplo deste último pode ser mencionado como dados relativos a saúde das pessoas ou a segurança pública são pontos mais delicados e necessários de certas salvaguardas.

3) Como foi o processo de análise da política de privacidade do Whatsapp?

Resposta direta para essa pergunta:

Thiago Moraes afirmou que a chegou-se a um momento em que entendeu-se que há um limite de até onde se pode analisar uma política de privacidade sem adentrar todo o sistema de compliance de uma organização. Ao longo da análise, eles analisaram Relatórios de Impacto e sugeriram recomendações, o que permitiu entender que outros processos fiscalizatórios deveriam começar a partir desse. Todo esse processo está público no site da ANPD.

4) A ANPD vai desenvolver um modelo de RIPD?

Resposta direta para essa pergunta:

Thiago Moraes afirmou que essa era uma resposta que ele não era capaz de dar naquele momento, porque a ANPD ainda não deliberou sobre isso, embora as práticas internacionais levam a indicar que não é válido fechar um modelo muito específico, pois leva a entender que basta usar aquele modelo que estará em conformidade. Contudo, a ANPD tem o propósito de desenvolver algo que possa ser levado para o debate público.

5) Qual a importância da revisão de decisões automatizadas levando em consideração que as comunidades **negras, mulheres, analfabetos, pessoas com deficiência e LGBTQIA+ são as mais afetadas por tais decisões? O princípio da não discriminação é suficiente?**

Tarcizio Silva defendeu que as avaliações de impacto são ferramentas muito importantes para compreender a sociedade e permitir que discriminações sejam mitigadas.

- 6) Deveríamos constituir uma Agência autônoma composta por representantes da sociedade, notadamente por aquelas comunidades que são diretamente afetadas, para conduzir investigação independente e descentralizada dos riscos e danos ocasionados pela tecnologia?**

Tarcízio Silva relatou concordar plenamente com a existência de uma agência autônoma, uma vez que já existem exemplos de como a autorregulação realizada por grandes empresas que dominam infraestruturas essenciais a países como o Brasil já se demonstrou falha.

- 7) Levando em consideração o racismo algoritmo, quais seriam as características principais que o Brasil deveria adotar em sua avaliação de impacto de inteligência artificial?**

Sem resposta direta para essa pergunta

Respostas gerais

Edson Prestes defendeu que não necessariamente um sistema será enviesado por conta dos dados que foram fornecidos ao sistema, algo que exige uma necessidade de prevalência da transparência. Citou que o posicionamento das empresas em colocar a transparência como inimiga dos segredos de negócio seria uma grande falácia, já que a transparência possui inúmeros níveis, ou seja, um sistema transparente não é sinônimo de um código aberto. Por fim, Edson lembrou do desafio de se estabelecer uma formação para profissionais de tecnologia que englobe questões éticas e de direitos humanos.