

Descolonizando os Dados Estruturados da Internet



RELATÓRIO

EM OUTUBRO DE 2021, MAIS DE 40 PARTICIPANTES DE TODO O MUNDO SE REUNIRAM PARA UMA CONVERSA SOBRE A DESCOLONIZAÇÃO DOS DADOS ESTRUTURADOS DA INTERNET, ÀS VÉSPERAS DA WIKIDATACON. O EVENTO FOI ORGANIZADO CONJUNTAMENTE PELA WHOSE KNOWLEDGE?, WIKI MOVIMENTO BRASIL E WIKIMEDIA DEUTSCHLAND. ESTE É O RELATÓRIO SOBRE E AS REFLEXÕES QUE SURGIRAM A PARTIR DELE.

Imagem por Whose Knowledge?, CC BY-SA 4.0, via [Wikimedia Commons](#)

Imagem por Whose Knowledge?, CC BY-SA 4.0, via [Wikimedia Commons](#)



Conteúdos

Por que organizamos essa conversa?

O que fizemos, e como o fizemos?

O que surgiu

 Perspectivas e provocações

 Imaginações e implementações

 Escuta e aprendizado

Quais os próximos passos?

Gratidão

Por que organizamos essa conversa?

Os dados estruturados estão no centro de como a internet, como a conhecemos atualmente, funciona. São informações organizadas de tal forma que podem ser facilmente lidas, compreendidas e processadas por máquinas. Inúmeros aplicativos e plataformas estão construídos sobre tais sistemas de dados estruturados, incluindo ferramentas criadas por provedores de conteúdo como Google e iniciativas colaborativas como Wikidata. Como resultado, tais dados permeiam todos os cantos de nossa vida diária online, desde mapas e sistemas de navegação até avaliações e classificações de filmes.

Através desses sistemas, grandes quantidades de dados são selecionadas, organizadas e classificadas em relação a outros dados. Desta forma, são informados e estruturados por e em torno de regulamentações, tradições e epistemologias específicas. Em uma tentativa de categorizar o mundo, eles prescrevem certos enquadramentos e visões sobre ele. Em outras palavras,

estão longe da neutralidade de dados frequentemente presumida. Analisando a dinâmica de poder dos dados estruturados, podemos examinar quais visões, qual agenda, quais ontologias (categorias de classificação) e quais decisões constroem e sustentam essas classificações e sistemas. Podemos, então, trabalhar em conjunto para construir sistemas de dados estruturados mais justos e equitativos. Essa transformação dos sistemas de dados estruturados — e do próprio conhecimento — requer uma discussão e um compromisso contínuos que não se limitem a empresas de tecnologia ou setores específicos. Ela precisa ser um esforço coletivo de pessoas, organizações e comunidades com diferentes formas de conhecimento e expertise, um processo que chamamos de aliança (com base no neologismo do inglês allyship) ou "solidariedade em ação". Mais importante ainda, porém, é que esse processo compartilhado coloque ao centro a expertise, as experiências vividas e a liderança da maioria minorizada do mundo.

[Como já compartilhamos antes](#), apenas uma fração do conhecimento público

online é produzida sobre ou por mulheres, pessoas não-brancas, pessoas LGBTQIA+, comunidades indígenas e povos do Sul Global, em línguas que não são o inglês ou outras línguas coloniais. No entanto, somos a maioria da população mundial, a maioria das pessoas online, e as mais afetadas pela forma como os dados estruturados são usados ou abusados. Somos as comunidades mais impactadas pelas estruturas históricas e contínuas de poder e privilégio e sistemas interligados de opressão, como colonização, patriarcado, racismo, homofobia, classismo, castismo, capacitismo e mais além. É por isso que precisamos começar a descolonizar a internet como uma prática, não simplesmente uma metáfora: precisamos transformar as diferentes formas de poder e privilégio que moldam nossas tecnologias digitais.

Saiba mais sobre por que a justiça do conhecimento é importante para dados estruturados a partir [desta apresentação](#) de nossa co-diretora Anasuya Sengupta.

A *Descolonizando os Dados Estruturados da Internet* nasceu da necessidade urgente de tal conversa

Que apresentação poderosa e sincera sobre a descolonização do conhecimento e sua importância para os dados estruturados de @Anasuyashh @WhoseKnowledge na #WikidataCon2021. Obrigada . Não estamos aqui para preencher as lacunas, mas para reestruturar e repensar as formas de fazer as coisas.

FLAVIA DORIA

@AFLAVIADORIA NO TWITTER, 29 DE OUTUBRO DE 2021

coletiva, trazendo ao centro pessoas que muitas vezes são marginalizadas. Em 13 de outubro de 2021, convidamos mais de 40 pessoas ativistas do conhecimento, organizadoras comunitárias, criadoras de tecnologia e outras "aliadas incomuns" para juntar-se a um espaço seguro, multilíngue e colaborativo. Esse grupo de pessoas brilhantes era composto, em sua maioria, por pessoas que se identificam como mulheres (71%), do Sul Global (66%) e com origem indígena/negra/não-branca (82%). A conversa serviu como pré-conferência para a [WikidataCon 2021](#), e foi organizada conjuntamente por [Whose Knowledge?](#), [Wikimedia Deutschland](#) e [Wiki Movimento Brasil](#). Como resultado, priorizamos também a participação de comunidades da América Latina e do Caribe.

5

O que fizemos, e como o fizemos?

A *Descolonizando os Dados Estruturados da Internet* foi organizada em três painéis diferentes. No Painel 1, palestrantes convidadas ofereceram *Perspectivas* e *Provocações* para delinear as conversas gerais. Em seguida, participantes se dividiram em sete grupos para discussões denominadas *Imaginações* e *Implementações*. Finalmente, ouvintes relataram os "insights" de cada grupo em uma sessão de *Escuta e Aprendizado*, na qual compartilhamos reflexões e ideias para os próximos passos.

Iniciamos a conversa expondo claramente três valores e compromissos fundamentais: amor, respeito e solidariedade. Com base nesse conjunto de princípios orientadores, queríamos que participantes estivessem conscientes de suas posicionalidades e privilégios, e que pudessem ser seus eus plenos e múltiplos durante toda a sessão. Nosso objetivo era promover interações centradas na escuta e no aprendizado, e no respeito (de cada pessoa pela outra e

pelo nosso tempo em conjunto). Ao nos envolvermos em conversas significativas, queríamos construir as bases para áreas-chave e ações concretas futuras.

Realizamos a conversa virtualmente, e oferecemos interpretação simultânea em inglês, português e espanhol. Queríamos que todas as pessoas pudessem se expressar plenamente em um idioma com o qual se sentissem confortáveis e que pudessem compartilhar suas experiências e conhecimentos umas com as outras, independentemente das barreiras linguísticas.

Foi realmente uma conversa extremamente esclarecedora, e ouvir as vozes de diferentes disciplinas acadêmicas, o diálogo de saberes – que sempre enriquece tanto – e não só isso, mas também de tantos países, e com tantos idiomas.

Uma pequena Babel.

MONICA CECILIA CALDERON

SESSÃO DE ENCERRAMENTO, 13 DE OUTUBRO DE 2021

Além disso, estabelecemos práticas claras de privacidade a fim de garantir a segurança e o bem-estar das pessoas durante toda a chamada. Por exemplo, participantes podiam manter suas câmeras desligadas se preferissem e precisavam obter consentimento explícito para publicar e atribuir qualquer citação identificável de outra pessoa participante antes de compartilhá-la nas mídias sociais durante e após o evento.

Eis como as conversas se desenrolaram.

O que surgiu

Perspectivas e provocações

A conversa começou com um painel chamado "Perspectivas e Provoações", no qual enquadrámos as questões urgentes em torno da descolonização de dados estruturados, e as pessoas palestrantes explicaram o que isso significava para elas. Essa primeira parte contou com os insights de Lydia Pintscher, que é gerente de produtos da Wikidata; Matariki Williams, escritora, curadora e historiadora Māori, que está

atualmente no Ministério da Cultura e Patrimônio em Aotearoa, Nova Zelândia; Paz Peña, ativista e techie feminista do Chile; e Stacy Allison-Cassin, acadêmica da Universidade de Toronto, bibliotecária, entusiasta da Wikidata e cidadã da Nação Métis de Ontário, Canadá. Esse painel inicial forneceu as bases sobre o que está em jogo quando falamos de dados estruturados.

Começamos situando as questões e os enquadramentos que cobriríamos durante o evento: o que são dados estruturados e o que significa construir sistemas em torno deles que centralizem múltiplas epistemologias. A fim de avançarmos, colocamos ao centro a análise de poder: questionando como o controle sobre os recursos — desde os projetos de plataformas até o acesso ao conteúdo em seus idiomas — se manifesta em diferentes níveis online. Perguntas como "de quem é a história?", "quem é a pessoa contadora de histórias e a pessoa curadora?", e "de quem são os sistemas de categorização?" permearam muitas intervenções. À medida que o painel avançava, nos aprofundamos na importância de dados estruturados e

construímos as bases para imaginar outras possibilidades.

Nessa primeira parte do evento, painelistas contestaram a suposta neutralidade dos dados estruturados. Nós enquadrámos os dados estruturados como pedaços de ideologia, não como categorias neutras que classificam o mundo de uma certa maneira. Participantes navegaram pelas nuances que surgem quando questionamos como os dados são classificados e com base em que epistemologias, ao mesmo tempo em que trazemos ao centro as necessidades e o direito de recusa e determinação de nossas comunidades. Como construímos alternativas baseadas na emancipação e na libertação? Como foi muito bem questionado por painelistas, essa conversa precisa ir além de um item não-essencial e ser reconhecida como um assunto urgente que requer recursos para ser mudada e melhorada.

Aqui está uma prévia desta discussão inicial. As respostas foram editadas e condensadas por uma questão de clareza e concisão.

O que significam dados estruturados, e por que é importante que falemos sobre eles?

Temos todos esses diferentes sistemas que interagem, que estruturam nosso movimento pelo espaço de maneiras particulares, e às vezes essa forma de pensar realmente nos guia [...]. Em cada um desses momentos em que encontramos um desses sistemas de estruturação, há um pequeno pedaço de ideologia que conceitua o mundo de uma determinada maneira. É uma interação entre esses pedaços de informação ou dados e o sistema de informação que os rodeia. [...] Portanto, os dados estruturados para mim começam naquele lugar fundamental de como estruturamos o mundo ao nosso redor, de como interagimos com diferentes sistemas de informação. Então, chegamos a um sistema como a Wikidata, que eu acho realmente maravilhosa de muitas maneiras, como sua capacidade de ser uma colaboração entre comunidades e de ter certa agência para mudar algumas dessas estruturas de conhecimento.

Imagem por Whose Knowledge?, CC BY-SA 4.0, via [Wikimedia Commons](#)



— STACY ALLISON-CASIN

O que significa ter múltiplos conhecimentos, ou epistemologias, no cerne dos dados estruturados?

Uma das coisas em que tenho pensado é a forma como o poder desempenha um papel nos dados estruturados. E como alguém que é uma pessoa usuária e não uma pessoa que projeta esses dados estruturados, especialmente como uma indígena em um país colonizado, a maneira como isso se mostra nas áreas em que trabalho é bastante grande. [...] Não nos encaixamos no sistema, pois o sistema não foi projetado por nós. Isso tem muitas implicações na forma como o próprio conhecimento foi compartilhado, ou coletado, e reunido e incluído ao longo do tempo. [...] Quando falamos de processos internos, do ponto de vista do Ministério, se trata de fazer e estruturar dados de uma forma que torne nosso trabalho interno mais eficiente. Mas o que isso significa para nossos parceiros? Estamos esperando que se encaixem em um modelo que facilite nosso trabalho, sem colocar ao centro suas aspirações sobre como elas lidam com seu próprio conhecimento? [...] Então, quando se trata de pensar em abordar dados estruturados de e com múltiplas epistemologias, isso já acontece, e acontece naturalmente como uma pessoa indígena.

Imagem por Whose Knowledge?, CC BY-SA 4.0, via [Wikimedia Commons](#)



— MATARIKI WILLIAMS —

Como podemos reimaginar dados estruturados, especialmente a partir de uma perspectiva feminista e anti-colonial?

A estruturação dos dados nesses sistemas é uma continuação histórica da fantasia branca, eurocêntrica, colonial e patriarcal que pressupõe que os dados flutuam. Eles flutuam livres de suas origens, são despojados de nossos corpos, de nossa história e, portanto, são livres para navegar pelo mundo como uma moeda neutra, uma moeda universal. [...] Também acredito que os dados estruturados muitas vezes atuam como uma força hegemônica, porque eles literalmente trazem consigo formas de ver o mundo, como fazem as forças hegemônicas. Esta força hegemônica se apresenta como inevitável, procura ser a única forma de ordenar o mundo. [...] Se queremos ver os dados estruturados a partir de uma perspectiva feminista, temos que lutar pela democratização dessa discussão.

Imagem por Whose Knowledge?, CC BY-SA 4.0, via [Wikimedia Commons](#)



— PAZ PEÑA

O que você gostaria de ver feito de forma diferente em dados estruturados hoje, que nos ajudaria a chegar a esse espaço de emancipação e libertação?

Na Wikidata, estamos tentando fornecer dados estruturados de uma forma mais justa, mais equitativa, mais representativa das diferentes experiências e do mundo em que vivemos. [...] Estamos tentando encontrar um equilíbrio entre fornecer dados úteis e utilizáveis sobre as quais as pessoas possam construir aplicativos e, ao mesmo tempo, fazer justiça às pessoas e ajudá-las a se sentirem representadas e vistas nesses dados, e tê-los como uma representação de seu mundo. E isso às vezes torna os dados mais complexos no Wikidata. Frequentemente, nos dizem que nossos dados são difíceis de usar devido a isso. E, se eu pudesse mudar uma coisa, seria encontrar maneiras de preencher essa lacuna.

Imagem por Whose Knowledge?, CC BY-SA 4.0, via [Wikimedia Commons](#)



LYDIA PINTSCHER —

Imaginações e implementações

Durante essa parte do evento, participantes se dividiram em sete grupos menores, que combinavam membros de diferentes origens e especialidades. Cada grupo tinha uma pessoa facilitadora, assim como uma pessoa encarregada de ouvir e trazer os destaques da conversa ao plenário geral. Dessa vez, fizemos perguntas voltadas para o futuro: como seriam os dados estruturados com múltiplas epistemologias? Pedimos às pessoas que imaginassem uma internet em 2040 que tivesse ao centro os dados estruturados com múltiplas epistemologias.

Os grupos detalharam diferentes elementos relacionados ao tema: a necessidade de reconhecer que as experiências vividas vão muito além dos dados online (estruturados); a urgência de criar, promover e expandir iniciativas em torno da alfabetização e educação sobre dados e seus impactos sobre as pessoas, e as possibilidades criadas pela centralização de comunidades marginalizadas no projeto, propriedade e liderança de plataformas online.

Veja um resumo dessas conversas:



Feliz por participar das salas simultâneas da Descolonizando os Dados Estruturados da Internet, da @WhoseKnowledge. O resultado mais importante aqui é pensar que a busca da justiça em #GráficosdeConhecimento não deve ser restrita apenas a fatos, mas também a modelos de dados e limitações lógicas.

HOUCMEDDINE TURKI

@CSISC1994 ON TWITTER, 13 DE OUTUBRO DE 2021

Escuta e aprendizado

Essa sessão, chamada *Escuta e Aprendizado*, se seguiu às nossas discussões em grupos menores. Convidamos ouvintes de cada grupo a se reportarem ao plenário e, mais uma vez, tivemos um grupo incrível de participantes: Thomas Hervé Mboa Nkoudou, um cientista social, bioquímico e criador camaronês; Maya Indira Ganesh, uma pesquisadora e educadora indiana sobre tecnologia, especialmente sobre ética em inteligência artificial; Jani Pereira, uma bióloga baseada em São Paulo, Brasil; Constanza Verón, uma historiadora e educadora baseada em Buenos Aires, que está atualmente na Wikimedia Argentina; Asaf Bartov, um Wikimedista e entusiasta por dados estruturados, e bibliotecário digital; Amanda Jurno, professora e pesquisadora brasileira em estudos de comunicação, do Wiki Movimento Brasil, e Alex Hanna, socióloga e diretora de pesquisa do Distributed AI Research Institute.

Essas pessoas participantes explicaram o que as surpreendeu, entusiasmou ou

desafiou e quais foram as ideias mais radicalmente imaginativas e interessantes — assim como sugestões de como chegar lá. Nossas pessoas ouvintes e outras participantes ofereceram essas ideias-chave:

ACESSO E CONTROLE DE DADOS

- ▶ Reconhecer que a descolonização de dados estruturados e dos sistemas construídos sobre eles é uma questão crítica e urgente de direitos humanos, e não simplesmente uma conversa opcional, feita à margem.
- ▶ Reconhecer que as comunidades marginalizadas — a maioria minorizada do mundo — têm pouco ou nenhum acesso ou controle sobre os dados que governam suas vidas atualmente.
- ▶ Compartilhar conhecimentos com diferentes comunidades sobre dados, ferramentas técnicas e como habitar e usar espaços online e ferramentas digitais de forma segura e completa.
- ▶ Estar ciente de que uma visão para dados estruturados centrada na comunidade pode ser inconciliável com as agendas de alguns indivíduos

e organizações envolvidos com dados estruturados, especialmente aqueles que os utilizam e que priorizam o lucro em vez do bem-estar humano.

AGÊNCIA/ENGAJAMENTO

- ▶ Negar às pessoas as desculpas para não se envolverem na descolonização de dados estruturados, criando e compartilhando recursos, e gradualmente criar um ecossistema onde não haja desculpas razoáveis para não reconhecer e representar uma diversidade de contextos e epistemologias.
- ▶ Reconhecer que as comunidades indígenas e outras comunidades marginalizadas devem ter o direito de se recusar a ter seus conhecimentos online, ou a serem *datificadas*, e o direito à agência sobre quais informações elas compartilham com o mundo.

▶ DADOS DISTRIBUÍDOS

- ▶ Desenvolver maneiras de criar conjuntos de dados menores e conectados que sejam governados pelas próprias comunidades marginalizadas, ao invés de empresas pautadas apenas no lucro.

▶ REDESENHAR SISTEMAS COLOCANDO O PLANETA AO CENTRO

- ▶ Estar consciente e incorporar nas discussões os impactos ambientais da estrutura necessária para o processamento de grandes quantidades de dados (por exemplo, centros de dados).
- ▶ Reestruturar sistemas enquanto nos perguntamos constantemente: para que servem? Qual é o propósito?

▶ PLURALIDADE DE DADOS

- ▶ Redesenhar sistemas e modelá-los com base em diferentes conhecimentos e das comunidades e suas gloriosas complexidades.
- ▶ Reconhecer que a maioria dos sistemas de conhecimento do mundo não são baseados em texto (e podem incluir um meio que inclui canções, contos, obras de arte de diferentes tipos), e ir além do texto para encontrar diferentes formas de codificação de dados.
- ▶ Aprender com exemplos locais e suas especificidades para criar mudanças mais amplas, especialmente porque "local" raramente significa "pequeno", principalmente no Sul Global.
- ▶ Comprometer-se a escutar e acomodar diferentes perspectivas para mudar a natureza dos dados estruturados, sabendo que a descolonização é um processo que leva tempo e esforço constante de múltiplas pessoas e comunidades.

Quais os próximos passos?

Para muitas pessoas participantes, *Descolonizando os Dados Estruturados da Internet* foi uma oportunidade muito necessária para parar por um momento e imaginar possibilidades radicais de dados estruturados. Também significou uma oportunidade de fazê-lo em comunidade e entre regiões e áreas de estudo e de trabalho. Daqui em diante, muitas delas manifestaram seu interesse em organizar mais espaços coletivos como esse e avançar com passos mais concretos em direção a práticas emancipatórias.

Como organizadoras, reconhecemos a enormidade do desafio, e sabíamos desde o início que essa conversa precisaria ser a primeira de muitas. Ao chegarmos em 2022, estamos ansiosas para criar e reunir mais oportunidades para reimaginar radicalmente e redesenhar os dados estruturados da internet desde uma perspectiva feminista, anticolonial e anti-racista. Com esses próximos passos em mente, permanecemos conectadas através de uma lista de emails com todas as

ASSISTA ÀS GRAVAÇÕES DO PAINEL

ESCUA E APRENDIZADO 

E DA

SESSÃO DE ENCERRAMENTO 

pessoas que se juntaram a nós para essa primeira conversa. Também manteremos nossas comunidades informadas e envolvidas através de atualizações [em nossa newsletter](#).

Gratidão

A *Descolonizando os Dados Estruturados da Internet* não teria sido possível sem todas as pessoas que se juntaram à conversa. Gratidão a todas as incríveis pessoas panelistas, ouvintes e participantes. Gratidão também às nossas fabulosas pessoas intérpretes, sem as quais não haveria uma conversa multilíngue. Também queremos expressar nossa gratidão às nossas pessoas co-conspiradoras desse evento, Amanda Jurno e Érica Azzellini do [Wiki Movimento Brasil](#), e Sabine Müller e Dominik Scholl da [Wikimedia Deutschland](#). Por último, mas não menos importante, aproveitamos essa oportunidade para agradecer às incríveis pessoas que nos apoiaram, parceiras e aliadas, cujo suporte diário fez com que essa conversa acontecesse, e fará com que muitas mais se tornem realidade.

