

Como falam os dados? Como se fala de dados?

Alternativas para a produção social do conhecimento a partir dos dados

Bernardo Beiriz e Daniel Lannes

Da Ciência Cidadã à Geração Cidadã de Dados Definições, papéis e futuros possíveis

O avanço da digitalização traz consigo inúmeras transformações para a produção, distribuição e consumo do conhecimento. A multiplicação das fontes de dados é nesse contexto simultaneamente um **desafio** e uma **oportunidade**. Com o objetivo de explorar caminhos para a navegação nesse cenário de mudança, o presente trabalho analisa formas alternativas de produção de conhecimento a partir dos dados. A **Ciência Cidadã**, a **Geração Cidadã de Dados** e conceitos associados são discutidos como iniciativas capazes de promover um reestruturação do processo de construção do conhecimento, sendo capazes também de potencializar os impactos políticos e sociais que o conhecimento pode ter. Para isso, começamos por duas perguntas que permitem repensar o papel e a função social dos dados em uma cadeia de geração de conhecimento: **como falam os dados** e **como se fala deles**?



Termos comumente associados à Geração Cidadã de Dados e à Ciência Cidadã



Diagrama elaborado pelos autores com base em Dykes (2016)

Como falam os dados?

Os dados falam. Falam sobre realidades explícitas ou implícitas, associam-se a narrativas e contam histórias, ajudam a construir gráficos e a preencher tabelas. Fortalecem argumentos ou derrubam hipóteses. **Os dados, entretanto, não se expressam sozinhos: sua capacidade de "falar" advém de sua interação com aqueles que participam no processo de produção, tratamento, curadoria e divulgação dos mesmos.**

Os dados falam, mas o fazem em um determinado contexto, estabelecendo redes de interações com diferentes interlocutores e realidades. A **Geração Cidadã de Dados (GCD)** e a **Ciência Cidadã**, entendidos como termos guarda-chuva sob os quais estão abrigados conceitos semelhantes, oferecem caminhos para **fazer com que os dados falem propositadamente**, isto é, com um objetivo. **Como engajar com os dados?** Diferentes respostas para essa pergunta gerarão diferentes resultados, algo que exploraremos mais à frente.

Por que dados?

Dados falam, mas para que se expressem precisam antes ser gerados, coletados e/ou construídos. A digitalização provoca uma multiplicação e diversificação das fontes possíveis de produção de dados, tornando-as mais acessíveis. Sensores de baixo custo e remotamente configuráveis podem ser distribuídos para gerar dados sobre a qualidade do ar em diferentes localidades, assim como smartphones podem registrar e enviar informações sob demanda para diferentes projetos. Esse universo de possibilidades, no contexto da produção social do conhecimento, é regido pelos **propósitos** e pelas **necessidades** identificadas pela "comunidade de pesquisa". Em outras palavras, definir o porquê de usar dados auxilia também a pensar formas de **como utilizá-los**. Em um sentido mais amplo, **dados podem ser entendidos como uma das diversas ferramentas disponíveis para promover a produção social do conhecimento; como instrumentos de intervenção política, social e científica em diferentes contextos.** Trabalhar com dados requer uma valorização dos mesmos para **além do papel de "matéria-prima", enxergando-os como atores ativos na produção do conhecimento.** A construção e a legitimação de metodologias que potencializem a agência, a possibilidade de agir política e socialmente, dos dados é um dos desafios abordados nas páginas seguintes.

Como falar com dados?

Existem diferentes iniciativas, abordagens e metodologias para o uso de dados na produção do conhecimento. Entre o momento do planejamento até a coleta/construção dos dados, passando pela análise e consolidação dos resultados das pesquisas, **a tarefa de "falar com os dados" atravessa diferentes desafios**: tanto **materiais** como **imateriais**.

Um dos principais desafios imateriais enfrentados nesse contexto é obter a **"autorização"** e o **reconhecimento** para que seja possível **se expressar como produtores de conhecimento**.

Gwen Ottinger (2017) afirma que "devido à autoridade estrutural concedida à ciência e aos cientistas na política ambiental" (p. 2), por exemplo, podemos pensar nos **"leigos"** ou nos **"não-cientistas"** como uma **categoria que é marginalizada no contexto da justiça epistêmica**. Em outras palavras, aqueles que se encontram fora da Ciência "formal" são injustiçados/não são reconhecidos quanto a sua condição de "conhecedores".

A política dos/nos dados

Falar com os dados, portanto, faz parte de um universo de disputas muito mais amplo sobre conhecimento, legitimidade e justiça. Tendo consciência dessas disputas, quando passamos a utilização concreta e objetiva dos dados no desenvolvimento das pesquisas, devemos tecer perguntas capazes de unir esses desafios materiais e imateriais.

Gilberto Vieira (2018) delimita algumas das questões comumente abordadas pela organização data_labe, da qual é co-fundador, e que podem ser utilizadas para pensar os desafios sobre os quais nos debruçamos anteriormente:

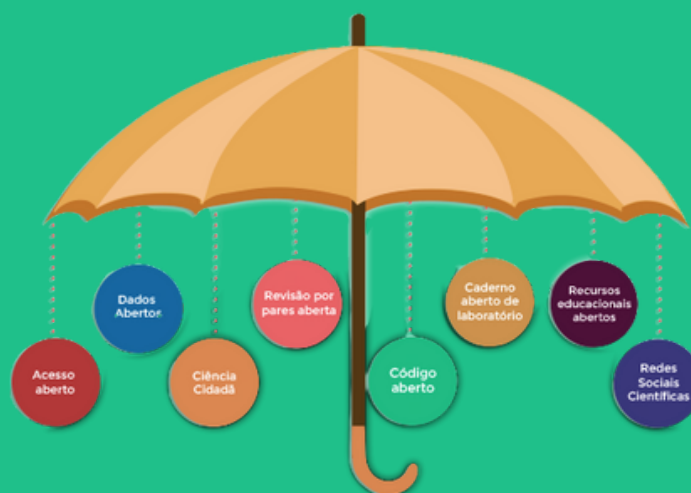
"A quem pertencem os dados? A favor de quem eles estão mais ou menos disponíveis? Por quem são produzidos? Como tornar os dados visíveis a ponto de garantir novas narrativas que permitam o desenvolvimento de comunidades plenas de direitos?"



Ciência Aberta

Um dos conceitos que podemos utilizar para contextualizar a discussão sobre Geração Cidadã de Dados e Ciência Cidadã é o de **Ciência Aberta**. Fabiano Couto Corrêa da Silva e Lúcia da Silveira oferecem uma definição para o termo, afirmando que:

"A Ciência Aberta é um **movimento** que incentiva a **transparência da pesquisa científica** desde a concepção da investigação até o uso de softwares abertos. Também promove esclarecimento na elaboração de metodologias e gestão de dados científicos, para que estes possam ser **distribuídos**, **reutilizados** e estar **acessíveis** a todos os níveis da sociedade, sem custos. Propõe, ainda, a **colaboração de não cientistas na pesquisa**, ampliando a participação social por meio de um conjunto de elementos que dispõem de novos recursos para a formalização da comunicação científica" (SILVA; SILVEIRA, 2019, p. 2)



O guarda-chuva da Ciência Aberta
Diagrama elaborado pela CAPES sobre
Ciência Aberta

Como se fala de dados?

A Ciência Cidadã

Afinal, o que é?

Existe uma grande variedade de definições sobre o que é a ciência cidadã, sendo difícil limitá-la a uma caracterização específica. Normalmente, essas definições remetem à uma **ampliação da participação na produção científica** em relação aos modelos convencionais, com o envolvimento de cidadãs e cidadãos em qualquer etapa do projeto científico (Eitzel et al., 2017). Assim, a ciência cidadã pode ser um caminho para suscitar **maior acessibilidade** à informação e aos dados científicos, assim como para tornar a produção científica **mais democrática** e **transparente** (de Sherbinin et al., 2021). Eitzel et al. (2017) identificam três tipos de caracterizações da ciência cidadã:

Quando caracterizada de forma instrumental, a ciência cidadã é descrita como uma **ferramenta metodológica** que permite a pesquisa alcançar maiores escalas do que o comum, por exemplo, com o uso das ferramentas de big data e open data, e a partir da colaboração voluntária da sociedade civil. Contudo, apesar de gerar engajamento da comunidade com a pesquisa científica, essa caracterização ainda insere a ciência cidadã nos modelos tradicionais e hierárquicos da produção científica e dos processos decisórios.

A ciência cidadã também é caracterizada como um **movimento político**, sendo descrita como meio de democratização do processo de pesquisa científica que aproxima a produção de conhecimento da comunidade. Quando assim qualificada, a ciência cidadã faz parte de uma luta por maior inclusão da comunidade nas agendas de pesquisa e nos processos decisórios. Dessa forma, é vista como uma resistência ao exclusivismo e à opacidade dos institutos de pesquisa e órgãos públicos que implementam políticas públicas

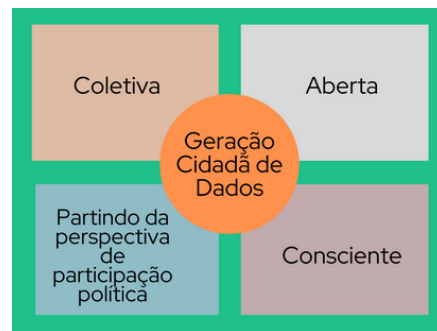
Se caracterizada como **ciência pública**, a ciência cidadã é compreendida como uma forma de aproveitar a capacidade de produção de conhecimento da sociedade, com o intuito de aumentar o conjunto de evidências que orienta, por exemplo, a elaboração de políticas públicas. A partir da colaboração entre pesquisadores acadêmicos e não-acadêmicos, a ciência cidadã está associada a uma promoção do ativismo comunitário, que busca publicar dados e informações locais para o público geral e influenciar processos decisórios.

A Geração Cidadã de Dados

A GCD se refere à produção de dados, por pessoas ou organizações, que tem o objetivo de **agir sobre problemas que as afetam diretamente**, monitorando, reivindicando ou conduzindo as mudanças necessárias. Dessa forma, a GCD passa pela **apropriação da produção científica pela própria comunidade local, produzindo representações diretas de suas perspectivas** (DataShift, 2015).

Envolve, portanto, a **participação voluntária** de pessoas na coleta de dados sobre **questões que atendam aos interesses das comunidades locais** e não apenas sobre os interesses de científicos de pesquisadores acadêmicos (Ponti, 2020). Essas pessoas **atuam para que suas vozes sejam ouvidas** pela comunidade e pelo poder público e **criam uma base de dados que podem complementar ou contestar** dados e narrativas oficiais (DataShift, 2015).

Características da Geração Cidadã de Dados



Qualquer pessoa pode iniciar um projeto de GCD, e as motivações para tal podem ser diversas. Em alguns casos **a falta de precisão dos dados oficiais** ou mesmo a **falta de confiança nas autoridades públicas** por parte da comunidade pode incentivar cidadãos e cidadãs a se organizarem. Em outros, a necessidade de **complementação de dados oficiais** ou de **conscientização** sobre tópicos que não recebem a devida atenção pode mobilizar as pessoas. Alguns exemplos:

Originalmente chamado de DefeZap, o projeto foi criado como um canal de denúncia à violência de agentes do Estado contra civis via whatsapp. Após apuração, as denúncias eram enviadas aos órgãos responsáveis. Tendo êxito em sua operação, a iniciativa passou a integrar a Comissão de Direitos Humanos da ALERJ.



A iniciativa foi criada frente à ausência de indicadores sobre tiroteios no Rio de Janeiro. Os dados eram coletados manualmente na internet e na imprensa. Hoje se tornou uma organização que produz mais de 20 indicadores sobre violência armada com dados abertos.



Trata-se de um canal de denúncia, via whatsapp, que busca mapear os problemas de saneamento básico no Complexo da Maré. O projeto cria uma base de dados complementar à oficial com o intuito de gerar debates que pressionem o poder público por soluções mais adequadas à realidade local.

Onde os dados falam diferente: o Sul Global na produção de conhecimento a partir dos dados

A ciência "passiva" não é suficiente para o Sul Global. A ciência e a produção de conhecimento de maneira geral, especialmente no contexto dos países "em desenvolvimento", devem buscar ativamente a transformação da realidade ao seu entorno. Garantir a **transparência** e a **colaboração** é uma condição necessária, porém não suficiente, para a construção de uma ciência e para uma produção social do conhecimento de verdadeiro impacto.

Os princípios da Ciência Aberta, por exemplo, embora norteiem o desenvolvimento de um ecossistema científico mais inclusivo, transparente e que dialoga com o restante da sociedade, devem ser "**subvertidos**" e portanto **interpretados como deveres**: não basta que a pesquisa esteja disponível/acessível para todos. É necessário um **movimento ativo de socialização**, de transformação de resultados em guias para **incidir sobre os contextos e temas estudados**.

A relevância do conhecimento que é gerado a partir dos dados no Sul Global está relacionada a outros processos, muitas vezes invisibilizados e/ou apartados do discurso científico. Dinâmicas como o racismo ambiental, o colonialismo de dados, o extrativismo epistêmico e as desigualdades sociais encontradas nas periferias econômicas, políticas e científicas do Sistema Internacional tornam ainda mais urgente a transformação da ciência na direção de sua abertura, proatividade e socialização.

A urgência da conservação de modos de vida, de proteção e garantia dos direitos básicos a partir do conhecimento que pode ser gerado por metodologias colaborativas e que utilizam os dados conferem novos argumentos para os debates sobre a legitimidade das transformações no campo científico.

"O nível de desenvolvimento científico de um campo, país ou região não é medido simplesmente por publicações indexadas em bancos de dados científicos convencionais e pelo impacto de suas citações. É igualmente importante avaliar os resultados do trabalho de pesquisas locais e regionais para compreender as configurações da ciência e sua importância em cada contexto." É o que dizem as pesquisadoras Dirce Santini e Sônia Caregnato; Elas alegam que as periferias (entendidas em sua dimensão mais ampliada) têm seus próprios sistemas de geração e uso de conhecimento e avaliação, o que pode exigir diferentes tipos de indicadores. A simples transposição de epistemologias da "ciência principal" para os espaços periféricos tende a gerar análises inadequadas e efeitos nocivos à própria ideia de ciência".

Citação retirada de texto de **Gilberto Vieira (2022)**

Entre dados e sujeitos: Dados como ferramentas de res/existência

Aquilo que pode ser "apagado" ou **invisibilizado** pela falta de informações também pode **reexistir**, existir novamente, e **resistir** a novas tentativas de apagamento. As formas de produção de conhecimento aqui analisadas, portanto, mostram-se como ferramentas que podem ser utilizadas com inúmeros propósitos e resultados. Seus efeitos mais importantes são referentes a sua capacidade de dar voz aos sujeitos, efetivamente garantindo seu "**direito de existência**", de protesto e de saber. As discussões aqui levantadas inserem-se em um contexto mais abrangente de lutas pela transformação social a partir do conhecimento. O presente trabalho foi desenvolvido com o intuito de fazer frente à urgente necessidade de pensar **novas formas de participação social e de incidência**.



Sobre o PET IRI PUC-Rio

Financiado pelo Ministério da Educação (PET) e pela PUC/Rio (TEPP), o Programa de Educação Tutorial do Instituto de Relações Internacionais (IRI) da Pontifícia Universidade Católica do Rio de Janeiro (PUC-Rio) foi criado para promover a formação em pesquisa, além de oferecer densa experiência na produção de conteúdo diverso e de alta especialidade entre os alunos do curso de graduação em Relações Internacionais da PUC-Rio.

Referências

- CLÍNIO, A. Ciência aberta na América Latina: duas perspectivas em disputa. *Transformação*, v.31, e190028. <http://dx.doi.org/10.1590/238180889201931e190028>
- DE SHERBININ, A. et al. The Critical Importance of Citizen Science Data. *Frontiers in climate*, v. 3, mar. de 2021.
- DYKES, B. 2016. CC BY-SA 2.5 <<https://creativecommons.org/licenses/by-sa/2.5/>>, via Wikimedia Commons
- EITZEL, M.V. et al. Citizen science terminology matters: exploring key terms. *Citizen Science: Theory and Practice*, v. 2, n. 1, p.:1-20, 2007.
- OTTINGER, Gwen. Making sense of citizen science: Stories as a hermeneutic resource. *Energy Research & Social Science*, 2017.
- PONTI, M.; CRAGLIA, M. Citizen-generated data for public policy. *European Commission*, Ispra, 2020, JRC120231. Disponível em: <https://ec.europa.eu/jrc/communities/sites/jrccties/files/jrc120231_citizen-generated_data_for_public_policy.pdf> Acesso em 05 dez. 2022.
- SILVA, F. 2017. Mas o que é geração cidadã de dados?. <<https://medium.com/data-labe/mas-o-que-%C3%A9-gera%C3%A7%C3%A3o-cidad%C3%A3-de-dados-fdac93c8fd70>>
- SILVA, F.C.C.; Silveira, L. O ecossistema da Ciência Aberta. *Transformação*, v.31, e190001, 2019. <http://dx.doi.org/10.1590/2318-0889201931e190001>
- VIEIRA, G. Geração Cidadã de Dados: um fazer político. 2018. <<https://medium.com/data-labe/gera%C3%A7%C3%A3o-cidad%C3%A3-de-dados-um-fazer-pol%C3%ADtico-c6b0450babfa>>
- VIEIRA, G. Geração de dados como ciência cidadã. 2022. <<https://medium.com/data-labe/gera%C3%A7%C3%A3o-de-dados-como-ci%C3%A4ncia-cidad%C3%A3-f5ffe3d976b3>>
- WHAT is citizen-generated data and what is the DataShift doing to promote it? *DataShift*, 2015. Disponível em: <http://civicus.org/images/ER%20cgd_brief.pdf> Acesso em: 03 dez. 2022.