

Visualização de Dados de Saúde Pública: um estudo de caso sobre a Covid-19

Public Health Data Visualization: a case study on COVID-19

Michele Krieger Bohnert

Mestranda em Ciência da Informação da Universidade Federal do Rio Grande do Sul, Porto Alegre, RS, Brasil.

E-mail: michele.bohnert@gmail.com

Caterina Marta Groposo Pavão

Doutora em Comunicação e Informação pela Universidade Federal do Rio Grande do Sul – UFRGS, Brasil e doutorado sanduíche na Universidad Complutense de Madrid, Espanha; Professora da Universidade Federal do Rio Grande do Sul, Porto Alegre, RS, Brasil.

ORCID: <http://orcid.org/0000-0003-3712-7200>

E-mail: caterina@ufrgs.br

Fabiano Couto Corrêa da Silva

Doutor em Información y documentación Sociedad Conocimiento pela Universitat de Barcelona, Espanha; Professor Adjunto da Universidade Federal do Rio Grande do Sul – UFRGS, Porto Alegre, RS, Brasil.

ORCID: <https://orcid.org/0000-0001-5014-8853>

E-mail: fabianocc@gmail.com

Resumo

A visualização de dados tornou-se uma ferramenta fundamental para exploração e comunicação de mensagens complexas, em especial na comunicação científica. Este artigo discute as vantagens da visualização e também alguns de seus principais aspectos no escopo da Ciência Aberta, como consequência do fato de que há uma certa tendência a mostrar conjuntos de dados, elaborados de forma gráfica, próxima e didática. Esta foi uma pesquisa de análise qualitativa, com procedimentos de levantamento bibliográfico e estudo de caso, tendo sido analisadas três iniciativas de visualizações representativas de dados estatísticos atribuídos à Covid-19. As análises demonstram a utilidade comunicativa dessas visualizações de dados para comunicar informações de forma clara e eficiente através de gráficos estatísticos e gráficos informativos para diversos públicos. Mediante os casos selecionados, constatou-se que a visualização ajuda a analisar e raciocinar sobre dados e evidências, tornando dados complexos digeríveis, compreensíveis e utilizáveis.

Palavras-chave: visualização; infografia; narração visual de histórias; grandes dados.

Abstract

Data visualization has become a fundamental tool to explore and communicate complex messages, especially in scientific communication. This study discusses the advantages of visualization and some of its main aspects in Open Science as a consequence of its tendency for graphic, near, and didactic datasets. This was a qualitative analysis, with a literature survey and case study. In total, three initiatives to representatively visualize statistical data on Covid-19 were analyzed. Analyses showed the communicative usefulness of these data visualizations to communicate information clearly and efficiently via statistical graphs and information graphics to several audiences. Via the chosen cases, visualization helped to analyze and reason about data and evidence, making complex data digestible, understandable, and usable.

Keywords: visualization; infographics; visual storytelling; big data.

1. Introdução

Ciência aberta designa um conjunto de conceitos relacionados, envolvendo vários movimentos, dimensões e práticas não isoladas, mas que se complementam. Em síntese, busca romper barreiras de fluxo do conhecimento científico produzido e agregar novos atores para essa produção. Uma das abordagens refere-se ao aumento da participação cidadã no acesso e nos processos de desenvolvimento científico, o que envolve a ampliação de canais de disseminação da informação científica, a simplificação da linguagem, a aproximação de um público mais amplo aos resultados das pesquisas (FECHER; FRIESIKE, 2014). Para Estevão (2019, p. 300), “[...] a ciência aberta, o acesso aberto e os dados científicos abertos promovem uma abordagem mais participativa e cidadã da Ciência, evidenciando princípios democráticos e de interesse público da produção do conhecimento”.

Dentre as razões para ciência aberta elencadas por FOSTER (2021), estão incluídas a apropriação social do conhecimento, a promoção da transparência do processo científico, o envolvimento da sociedade no processo de criação do conhecimento e a democratização do acesso ao conhecimento científico. Essa proposta de uma ciência para todos, no entanto, esbarra na complexidade de acesso e de uso da produção científica. Democratizar significa tornar acessível. Contudo, acessar conjuntos de dados complexos sem ter condições de interpretá-los, acaba inviabilizando a participação de cidadãos comuns. A disponibilização de dados, ainda que em formato aberto, pode não ser suficiente. Embora o acesso seja importante, este não basta. Mais que ter acesso à informação, é preciso considerar seu uso, utilidade e relevância. Faz-se necessário, então, utilizar recursos que permitam aos usuários a compreensão desses dados. Neste contexto, uma das formas mais eficazes de explorar e de compreender conjuntos de dados é por meio da visualização.

A visualização de dados é definida por Kirk (2016), como representação e apresentação do dado para facilitar a compreensão, sendo aplicada em diversas áreas de estudo. Utiliza-se representações visuais para explorar, analisar e comunicar conjuntos de dados. O grande volume e a heterogeneidade de dados disponibilizados na *web* são um fenômeno contemporâneo. Porém, a massiva disseminação de dados gera dificuldades de compreensão. Dados, mesmo que em grande volume, são apenas dados: é preciso produzir informação e conhecimento para explorar os benefícios que essa matéria-prima bruta pode trazer (AMARAL, 2016a). Dessa forma, os dados brutos somente são úteis quando aplicados métodos em que se possa derivar conhecimento a partir deles.

Este artigo propõe, por meio de três estudos de caso relacionados à pandemia Covid-19, demonstrar as potencialidades da visualização de dados para ampliação, difusão e divulgação do conhecimento científico, sobretudo visando os agentes envolvidos além dos muros acadêmicos. Tem como objetivo explorar a visualização de dados como recurso que promova a apropriação social da Ciência.

2. Ciência Aberta

Ciência Aberta é a atividade científica praticada de modo aberto, colaborativo e transparente. Trata-se da prática da ciência de tal forma que outros podem colaborar e contribuir, devendo ser entendida por todos os agentes envolvidos: pesquisadores, instituições, formuladores de políticas, editores, empresas e sociedade em geral. Parte do entendimento que o conhecimento científico é um produto da colaboração social, logo sua propriedade pertence à comunidade (FOSTER, 2021). A Ciência Aberta é uma prática que pressupõe maior participação cidadã no desenvolvimento de processos científicos e tecnológicos, ampliando sua ação na geração de conhecimento científico e no uso de uma maior variedade de canais de comunicação científica, a partir de linguagens mais simples e que visem aproximar o público em geral daquilo que resulta dos trabalhos de pesquisa (FECHER; FRIESIKE, 2014). A Ciência Aberta, portanto, representa uma nova abordagem para o processo científico, baseada no trabalho cooperativo e nas novas formas de difusão e de divulgação de conhecimentos por meio de tecnologias de comunicação digital. De acordo com Sato (2017), o intuito da divulgação científica é servir como um primeiro contato com os conteúdos tratados, permitindo ao usuário aprofundar-se no tema conforme seu interesse. Nesse contexto, a representação visual da ciência é uma forma de tornar a percepção dos dados científicos mais clara, principalmente por meio das Infografias, comumente utilizadas como ferramentas que potencializam a comunicação e a visibilidade da informação científica. Seu uso como complemento informacional valida estudos ao relacionar os dados científicos, ampliando a compreensão sobre o alicerce teórico que sustenta as pesquisas.

Comumente, Ciência Aberta é definida como um termo guarda-chuva, uma vez que envolve diversos movimentos com o objetivo de remoção de barreiras para o compartilhamento dos processos de pesquisa, como acesso aberto, dados abertos, ciência cidadã, revisão por pares aberta, código aberto, caderno aberto de laboratório, recursos educacionais abertos e redes sociais científicas. Para Fecher e Friesike (2014), o mesmo termo evoca entendimentos bastante

diferentes, que vão desde o direito democrático de acesso ao conhecimento financiado publicamente ao desenvolvimento de ferramentas de colaboração disponíveis gratuitamente. Visando estruturar as diversas abordagens da produção e da disseminação do conhecimento científico, os autores propõem Cinco Escolas do Pensamento da Ciência Aberta. Este artigo discorre sobre os preceitos das escolas Democrática e Pública, uma vez que ambas incluem os cidadãos entre os grupos envolvidos.

A escola Democrática preocupa-se com o acesso ao conhecimento. Representa um conjunto amplo de argumentos que acompanham a discussão sobre acesso gratuito aos produtos de pesquisa. “[...] todos deveriam ter o mesmo direito de acesso ao conhecimento, especialmente quando é financiado pelo Estado” (FECHER; FRIESIKE, 2014, p. 25, grifo nosso). Machado (2015), complementa que o “acesso ao conhecimento e à informação são fundamentais para o desenvolvimento humano, social e econômico” (MACHADO, 2015, p. 207, grifo nosso). Dentre as ferramentas e métodos utilizados em busca da acessibilidade, este artigo destaca os Dados Abertos.

De acordo com a Open Knowledge Foundation (2021), são dados abertos aqueles que podem ser livremente utilizados, reutilizados e redistribuídos por qualquer pessoa. Para Machado (2015), a abertura de dados científicos, seguindo os princípios dos dados abertos, traz enormes vantagens para a toda a sociedade, pois a cadeia de informação se enriquece com a maior propagação e amplia as possibilidades de usos e reusos da informação. No entanto, o autor observa que para alcançar comunidades mais amplas é necessário que os dados sejam publicados de forma adequada e inteligível.

Por sua vez, a escola Pública se preocupa com a acessibilidade da criação de conhecimento, argumentando que a “[...] ciência precisa ser acessível para um público mais amplo” (FECHER; FRIESIKE, 2014, p. 19, grifo nosso). Nesta escola, podem ser reconhecidas duas correntes distintas: acessibilidade do processo de pesquisa (produção) e compreensibilidade do resultado da pesquisa (produto). O avanço da tecnologia de comunicação permite a inclusão de indivíduos externos ao meio acadêmico, tornando questões antes ocultas em transparentes e acessíveis ao cidadão comum. O envolvimento do público não acadêmico no processo de pesquisa científica sintetiza o conceito de Ciência Cidadã.

Albagli (2015) relata que nessa vertente estão contempladas iniciativas orientadas à maior participação, intervenção e empoderamento de cidadãos.

Ciência cidadã refere-se ao engajamento do público em geral nas atividades de pesquisa científica quando os cidadãos contribuem ativamente para a ciência, seja com seu esforço intelectual, com seu conhecimento local ou com suas ferramentas e recursos. (PARRA, 2015, p. 126)

Conforme Parra (2015), a diversificação nas formas de colaboração entre cientistas e cidadãos reinventa a dimensão pública da ciência, amplia as possibilidades de produção, coleta, compartilhamento e análise de dados e transforma as próprias dinâmicas de produção, validação, difusão e apropriação dos conhecimentos produzidos. O autor ressalta, no entanto, que há dificuldades nas condições de participação cidadã, devido às diferentes condições de acesso à informação entre os cidadãos e cientistas, assim como possibilidades desiguais de apropriação, aplicação e ressignificação dos conhecimentos produzidos. Logo, é necessário criar condições de acesso e uso dessas informações, por meio da utilização de ferramentas que permitam a compreensão e que atendam as expectativas do público. A visualização de dados se apresenta como uma alternativa no avanço da transparência e na acessibilidade do conhecimento científico.

3. Visualização de Dados

A visualização de dados é uma área multidisciplinar, uma vez que é explorada em distintas áreas do conhecimento, e que tem se destacado nos últimos anos devido ao aumento da produção de dados e aos avanços das tecnologias digitais. Utilizada especialmente na mineração, análise e comunicação, tem papel relevante, possibilitando a leitura dos dados e oferecendo a estes uma nova interpretação e significado de dimensão social, informacional e comunicacional. Estes passam, assim, de meros elementos brutos a geradores de conhecimento. Para Rininsland *et al.* (2017, p. 532), “[...] a visualização de dados é a disciplina que estuda como usar a percepção visual para comunicar e analisar dados”.

Segundo Few (2009), os dados, por si só, não são valiosos. Os dados só se tornam valiosos quando os entendemos. Para Amaral (2016b, p. 3), “[...] dados são fatos coletados e normalmente armazenados. Informação é o dado analisado e com algum significado. O conhecimento é a informação interpretada, entendida e aplicada para um fim”. Uma maneira eficiente de encontrar informações importantes em conjuntos de dados é a aplicação de técnicas de visualização. A principal vantagem do uso da visualização se dá pela capacidade humana de processar informações visuais muito mais rapidamente do que informações verbais. Pode-se detectar com facilidade atributos básicos, como diferenças de comprimento, tamanho, cor,

ângulo, textura. Visualização é a forma mais rápida de transformar dados que estão à disposição em algo útil. Transformar, por fim, dados em oportunidades.

Existem diversas definições para o termo Visualização de Dados. Apesar das variantes, a essência permanece a mesma – consiste na representação visual de dados com propósito de atribuição de sentido e comunicação.

Card, Mackinlay e Shneiderman (1999) definem como uma representação visual de dados abstratos de forma interativa, a fim de amplificar a cognição humana. Kirk (2012), conceituou como a representação e apresentação de dados que exploram nossas habilidades de percepção visual para amplificar a cognição. O mesmo autor, em 2016, sintetiza essa definição como a representação e apresentação de dados para facilitar a compreensão.

Kirk (2016), descreve as funções e as relações de cada componente. Dados são nomes, quantidades, valores estatísticos, datas, comentários etc., sendo textuais e numéricos em formato, normalmente mantidos em conjuntos de dados em forma de tabela, com linhas de registros e colunas de variáveis para os diferentes registros mantidos. No entanto, não se pode estabelecer facilmente o tamanho comparativo e a relação entre vários pontos de dados. Nossos olhos e mente não estão preparados para traduzir facilmente os valores textuais e numéricos para um significado quantitativo e qualitativo.

Para derivar a compreensão dos dados, é preciso vê-los representados de uma forma visual e este é o ato da representação de dados. A representação, então, diz respeito às escolhas de formas físicas em que os dados são retratados visualmente: linha, barra, círculo ou qualquer outra variável visual, tomando os dados como matéria-prima e criando uma representação para melhor retratar seus atributos, explorando as capacidades de percepção visual do cérebro. A apresentação dos dados se refere a todas as outras decisões que compõem a anatomia geral da visualização, incluindo escolhas sobre as possíveis aplicações de interatividade, recursos de anotação, uso de cores e a composição do trabalho. Por fim, a visualização de dados visa facilitar a compreensão. O objetivo é fazer com que o usuário fique mais bem informado sobre o assunto.

3.1 Objetivos da Visualização

Segundo Daniel Keim (1997 *apud* GRÉGIO *et al.*, 2009), a visualização de dados pode ser usada para os seguintes objetivos:

- a. *Análise exploratória*, utilizando dados de natureza conhecida sem uma hipótese definida sobre fenômenos que podem ocorrer com os dados. Envolve a busca visual por tendências, exceções, estruturas etc. O resultado da análise exploratória pode ser a definição de hipóteses.
- b. *Análise para confirmação*, usando dados de natureza conhecida e hipóteses sobre fenômenos relacionados a esses dados. Através da visualização, a hipótese pode ser confirmada ou rejeitada.
- c. *Apresentação*, ao usar os dados com o objetivo de demonstração dos dados, fenômenos relacionados a estes ou hipóteses. Deve-se definir uma técnica para apresentação apropriada e que permita fácil interpretação.

Iliinsky e Steele (2011) afirmam que a visualização é benéfica para analisar, compreender e comunicar informação. Por meio da visualização é possível identificar padrões, relacionamento e seus significados e identificar tendências. É, geralmente, parte de um processo maior, que envolve desde a análise exploratória de dados, passando pela descoberta de conhecimento e finalizando pela apresentação de resultados de técnicas ou análise visual.

3.2 Técnicas de Visualização

Existem diversas maneiras de visualizar dados. As técnicas de visualização procuram apresentar graficamente dados de forma que a representação visual gerada explore a capacidade de percepção do visualizador. A partir das imagens exibidas, o visualizador deve interpretar e compreender as informações apresentadas para, enfim, buscar novos conhecimentos. Para Freitas *et al.* (2001), uma das principais considerações a ser feita no processo de visualização é determinar qual técnica deve ser empregada em determinada aplicação ou situação, o que depende do tipo de informação que está sendo tratada e das tarefas que precisam ser realizadas pelo usuário.

Segundo Tufte (2001), para representar visualmente dados de maneira efetiva, todo gráfico deve:

- a. Exibir os dados;
- b. Induzir o visualizador (ou público-alvo) a pensar sobre o dado e não sobre metodologia, design gráfico ou tecnologia utilizada para gerar o gráfico;
- c. Evitar distorções que o dado não contém;
- d. Tornar coerentes conjuntos de dados volumosos;
- e. Encorajar o olho do visualizador a comparar pedaços dos dados;
- f. Revelar o dado em vários níveis de detalhes;
- g. Servir a um propósito claro.

Iliinsky e Steele (2011) declaram que dados diferentes necessitam, consequentemente, diferentes abordagens e técnicas. A definição da estrutura visual de um dado envolve o conhecimento da natureza do problema e do objetivo que se almeja representar. Devem, portanto, ser observadas as especificidades em relação à forma e à função em cada tipo de escolha visual.

Nos últimos anos, diversas iniciativas têm surgido, trazendo novas abordagens na classificação de técnicas (Figura 1). São, geralmente, utilizadas como guias ou bibliotecas e têm se tornado referências para auxiliar o processo de visualização. Dentre essas iniciativas, se destacam: *From Data to Viz*¹; *Data Viz Project*²; *Visual Vocabulary Tableau*³ e *Data Viz Catalogue*⁴. A escolha da técnica de visualização depende, sobretudo, do tipo de dado e do que se deseja expressar.

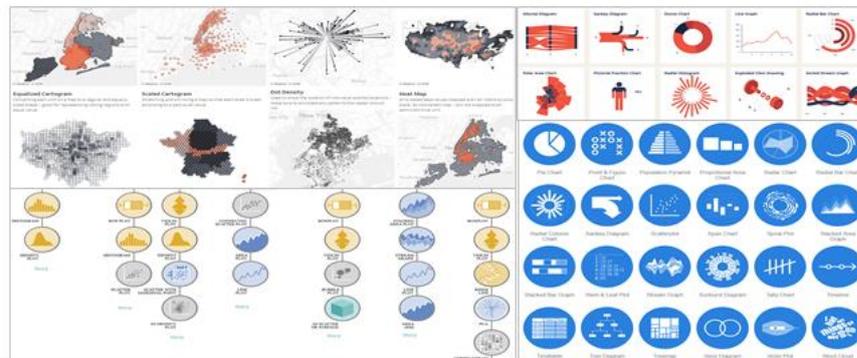
¹ <https://www.data-to-viz.com>

² <https://datavizproject.com>

³ <https://public.tableau.com/en-us/gallery/visual-vocabulary>

⁴ <https://datavizcatalogue.com/>

Figura 1 – Exemplos de técnicas de visualização



Fonte: Compilação dos autores⁵

Os dados somente são úteis quando consegue-se comunicar o que eles significam. Apresentá-los de maneira equivocada pode levar a uma interpretação incorreta ou confundir o público a que se destina. Por isso, a técnica de visualização é fundamental para comunicar claramente as informações.

4. Procedimentos Metodológicos

Do ponto de vista da natureza, trata-se de uma pesquisa aplicada, pois objetiva gerar conhecimentos para aplicação prática, dirigidos à solução de problemas específicos, envolvendo verdades e interesses locais.

A abordagem desse estudo é qualitativa. Segundo Richardson (2017), os estudos que empregam uma metodologia qualitativa podem, por exemplo, descrever a complexidade de determinado problema, analisar a interação de certas variáveis, compreender e classificar processos dinâmicos.

Em relação aos objetivos do artigo, esta classifica-se como exploratória, pois há necessidade de identificar, conhecer, levantar ou descobrir informações sobre determinado tema, objetivando proporcionar maior familiaridade com o problema, a fim de torná-lo mais explícito ou de construir hipóteses (GIL, 2010). Esse tipo de pesquisa permite se obter novos conhecimentos, bem como a ampliação e complementação do tema abordado.

Como métodos de pesquisas, utilizou-se levantamento bibliográfico e estudo de caso. De acordo com Lakatos e Marconi (2009), a pesquisa bibliográfica abrange todo o referencial

⁵ Montagem a partir de imagens coletadas nos sites *From Data to Viz*, *Data Viz Project*, *Visual Vocabulary Tableau* e *Data Viz Catalogue*.

teórico já tornado público em relação ao tema de estudo. Já o estudo de caso é a estratégia escolhida para examinar-se acontecimentos contemporâneos. Trata-se de uma investigação empírica que pesquisa um fenômeno dentro de seu contexto da vida real. “A investigação de um estudo de caso baseia-se em várias fontes de evidências e beneficia-se do desenvolvimento prévio de proposições teóricas para conduzir a coleta e a análise de dados” (YIN, 2010, p. 39).

Objetivando demonstrar as potencialidades da visualização de dados no contexto científico, tornando conjuntos de dados complexos em informação de fácil compreensão, foram identificadas as iniciativas expostas a seguir. Optou-se pela revisão narrativa de análise das visualizações de dados de casos relativos à pandemia Covid-19 nas bases de dados Johns Hopkins Coronavirus Resource Center, Our World in Data e o Painel de Coronavírus da Organização Mundial de Saúde.

4.1 Estudos de Caso

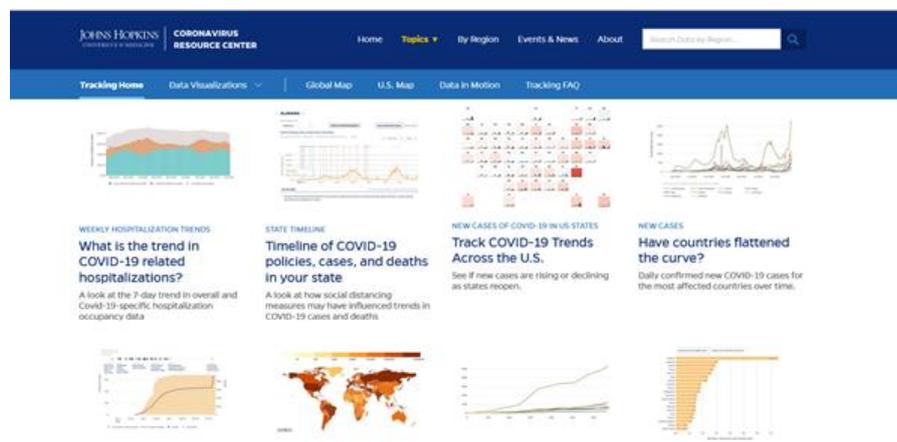
De acordo com a Organização Mundial de Saúde (OMS), Covid-19 é a doença causada pelo novo coronavírus denominado Sars-Cov-2. Primeiramente identificada na China em dezembro de 2019, foi elevada à estado de pandemia em março de 2020. Esta classificação se deu pela rápida disseminação geográfica, o que demandava ações conjuntas de governos a fim de frear os níveis de contaminação ocasionados pela doença. Segundo a OMS, o principal componente de disseminação é o contato entre as pessoas. Neste sentido, fez-se necessária a ampla divulgação das maneiras de prevenção, entre as quais se destaca o distanciamento social, assim como a disponibilização de informações que pudessem contribuir no desenvolvimento e na avaliação de estratégias epidemiológicas eficazes para prevenção da propagação da doença.

A pandemia de Covid-19 fez emergir a necessidade de abertura e transparência de dados, sejam estes governamentais ou científicos, uma vez que a falta de dados limitava a compreensão do verdadeiro impacto da pandemia, prejudicando o combate ao vírus. Fez surgir, também, a ampliação de canais de disseminação de informação científica através do uso de linguagem simples, visando o entendimento sobre a doença por todos os públicos.

4.1.1 Johns Hopkins Coronavirus Resource Center

Especialistas da Universidade Johns Hopkins (Estados Unidos) criaram o Johns Hopkins Coronavirus Resource Center (CRC), que se tornou uma das principais fontes de informação mundiais sobre o assunto (Figura 2). Tem por objetivo coletar e analisar dados disponíveis sobre casos, mortes, testes, hospitalizações e vacinas, auxiliando o público, legisladores e profissionais de saúde em todo o mundo a responder à pandemia Covid-19. A plataforma foi criada para que o público entenda a situação do vírus e como ele evolui, de forma transparente e permitindo o download dos dados.

Figura 2 – Painel de informações sobre pandemia Covid-19 CRC



Fonte: plataforma *online* CRC Johns Hopkins (2021)

O CRC é o resultado de uma colaboração interdisciplinar que abrange todo o sistema da Universidade e Medicina Johns Hopkins: Escola de Saúde Pública Bloomberg, Departamento de Medicina Johns Hopkins, Centro de Ciência e Engenharia de Sistemas, Centros para Impacto Cívico e Centro de Segurança Sanitária. É, ainda, apoiado pela Bloomberg Philanthropies e pela Stavros Niarchos Foundation. Uma equipe de cientistas e engenheiros de dados constrói e mantém a infraestrutura de dados do CRC. Obtém dados de nível local a global e, por meio desses conjuntos de dados complexos, fornece todas as visualizações e análises do CRC com informações constantemente atualizadas.

4.1.2 Our World In Data

Trata-se de um esforço colaborativo entre os pesquisadores da Universidade de Oxford (Inglaterra), que são os contribuidores científicos, e da organização sem fins lucrativos Global Change Data Lab, mantenedora do site e das ferramentas de dados. A equipe de pesquisa é afiliada ao Programa Oxford Martin sobre Desenvolvimento Global, cuja missão é produzir pesquisa acadêmica sobre os maiores problemas do mundo com base na análise empírica de dados globais. As pesquisas e dados publicados, no entanto, não se resumem ao trabalho desta equipe, pois contam com o trabalho de uma comunidade global de acadêmicos. O portal aborda temas diversos, não se limitando à exposição de informações sobre a pandemia Covid-19.

O principal objetivo dessa plataforma (Figura 3) é tornar o conhecimento, acerca dos grandes problemas globais, acessível e compreensível, a fim de possibilitar o uso de pesquisas e dados acadêmicos para fazer progresso contra essas mazelas. Parte-se do pressuposto que um dos principais motivos pelos quais não se alcança esse progresso se deve ao uso insuficiente das pesquisas e dados existentes, pois muitas vezes estes se encontram em bancos de dados inacessíveis e atrelados a jargões de trabalhos acadêmicos. Fundamentalmente a plataforma pretende construir uma infraestrutura que torne as pesquisas e os dados disponíveis abertamente e úteis para todos. Objetiva-se apresentar e sintetizar com precisão os dados sobre o desenvolvimento global para comunicá-los a um público mais amplo possível. A plataforma pública Our World in Data torna os resultados de pesquisas acadêmicas acessíveis e compreensíveis por meio de visualização de dados interativa.

Figura 3 – Painel de informações sobre pandemia Covid-19 *Our World in Data*



Fonte: plataforma *online Our World in Data* (2021)

As publicações de pesquisa são dedicadas a uma ampla gama de problemas globais: mudança demográfica, saúde, alimentação e agricultura, energia e meio ambiente, inovação e

mudança tecnológica, pobreza e desenvolvimento econômico, condições de vida, comunidade e bem-estar, direitos humanos e democracia, educação e conhecimento. O site de acesso aberto apresenta essa perspectiva global de como o mundo está mudando por meio de visualizações e explicações interativas sobre os dados. Ao produzir gráficos e tornar os dados de fato acessíveis, estes podem ser utilizados por diversos atores para os mais variados fins. Todas as saídas, incluindo dados, visualizações e ferramentas são de código aberto. Isso significa que são gratuitas para todos. As visualizações são licenciadas sob CC BY, podendo ser livremente utilizadas para qualquer propósito. Os dados estão disponíveis para *download* e todos os códigos são abertos sob licença MIT e disponibilizados no GitHub.

4.1.3 Painel Coronavírus da Organização Mundial da Saúde

O painel de coronavírus da OMS (Figura 4) apresenta contagens diárias oficiais de casos, mortes e utilização de vacinas de Covid-19 relatados por países, territórios e áreas. Através do painel, pretende fornecer uma visualização de dados atualizados diariamente, conectando os usuários a recursos úteis e informativos.

Os dados utilizados para a construção do painel estão disponíveis para download em acesso aberto. O acesso aberto da OMS aplica-se a todas as publicações licenciados sob CC BY-NC-SA 3.0 IGO. Essa licença permite aos usuários copiar, reproduzir, reimprimir, distribuir, traduzir e adaptar livremente o trabalho para fins não comerciais, desde que a OMS seja reconhecida como fonte.

Figura 4 – Painel de informações sobre pandemia Covid-19 OMS



Fonte: plataforma *online* OMS (2021)

5. Análise dos Estudos de Caso

Com o objetivo de criar instrumentos de apoio ao enfrentamento da crise sanitária e, assim, auxiliar autoridades responsáveis na tomada de decisões e informar o público a respeito do tema, os portais estudados apresentam conjuntos de visualizações de dados sobre a pandemia Covid-19. Os casos analisados possibilitam a visualização de dados em diversos aspectos. Cada nível de desagregação é caracterizado por um amplo conjunto de indicadores: número de casos diagnosticados, número de óbitos, taxa de contaminação (número de casos por 100 mil habitantes), taxa de letalidade (número de óbitos por número de casos), entre outros. As fontes de dados e os tipos de arquivo, utilizados para compor as visualizações dos painéis sobre a Covid-19, são especificados pelas plataformas e integram os processos comunicacionais.

Considerando a grande quantidade de dados produzida no contexto da pandemia, é possível aplicar diversas técnicas de visualização e interação. A fim de que as representações visuais possam ser alinhadas em torno de uma lógica comum, foram levantados critérios de análise entre essas iniciativas. Os portais referenciados apresentam duas dimensões principais nas representações visuais de dados: geoespacial e temporal.

5.1 Representações Geoespaciais

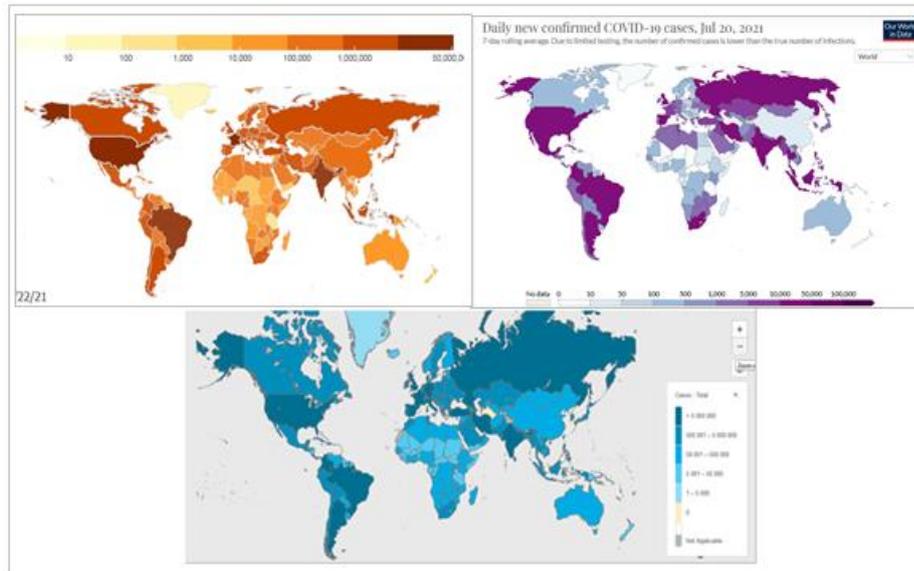
As representações geoespaciais são usadas quando os dados estão associados a uma localização geográfica específica. Sendo a pandemia Covid-19 um fenômeno global, diversas organizações de pesquisa e órgãos políticos e governamentais utilizam mapas temáticos para visualizar e compreender a propagação espacial do vírus e, também, para informar e orientar os cidadãos. Seu uso se dá para análises e representações de distribuição (localização), comparação e proporção. Dentre as representações verificadas, evidenciamos que os gráficos foram usados para comparar e exibir relacionamentos entre categorias, pelo uso de posicionamento e proporções. A visão geral dos gráficos foi usada para analisar padrões e correlações.

5.1.1 *Choropleth Map*

Trata-se de um mapa temático no qual as áreas são demarcadas com diferentes cores e padrões de sombreamento dentro de limites geográficos para mostrar a proporção de uma variável de interesse. Isso fornece uma maneira de visualizar valores em uma área, que pode mostrar variações ou padrões em todo o local exibido. A variável de dados usa progressão de cores para se representar em cada região do mapa.

Os três portais analisados utilizam essa técnica de visualização de dados (Figura 5). Tem sido empregada, especialmente, para representar os dados de casos e de mortes acumulados ao longo dos meses. Através dessa técnica pode-se observar a evolução da pandemia: dos primeiros casos diagnosticados no Oriente, o avanço pela Europa até a expansão pela América, abrangendo todas as regiões. Observa-se, ainda, a concentração de casos e de mortes em alguns países. O Brasil, por exemplo, foi fortemente impactado, sendo um dos países como maior registro de casos e de óbitos ocasionados pelo coronavírus. Considerando os números absolutos, o Brasil é o terceiro país com maior número de casos, tendo registrado mais de 22 milhões de contaminados desde o início da pandemia até final de dezembro de 2021, sendo superado apenas pelos Estados Unidos e Índia. Já em relação ao número de óbitos, foram registrados mais de 619 mil no mesmo período, quantitativo inferior somente ao dos Estados Unidos, cujas vítimas fatais ultrapassam 825 mil. Se forem consideradas as medidas relativas de casos, o Brasil encontra-se na escala de maior intensidade, com índice superior a 10 mil casos confirmados por 100 mil habitantes. Da mesma forma, o índice de mortes encontra-se na faixa de maior magnitude, superando os 200 óbitos por 100 mil habitantes.

Figura 5 – *Choropleth Maps*



Fonte: compilação dos autores⁶

Os mapas verificados apresentam conexões com relacionamentos geográficos usados para mostrar rotas através de duas categorias de visualização: casos e óbitos confirmados por países. Desse modo, os três mapas analisados alcançam o propósito de revelar padrões espaciais por meio da distribuição de conexões e concentrações, tanto pela densidade populacional como por números absolutos, através da variação de cores.

5.1.2 *Bubble Map*

Bubble map, ou mapa de bolhas, pode representar duas variáveis simultaneamente em uma visualização. É utilizado para visualizar localização e proporção de forma simples. Nesse mapa, os círculos são exibidos em uma região geográfica designada com a área do círculo proporcional ao seu valor no conjunto de dados.

Dentre os *sites* estudados, apenas CRC utiliza essa técnica. Através da imagem (Figura 6) pode-se observar a expansão de casos e mortes dentro dos espaços geográficos. O uso de círculos (bolhas) de diferentes tamanhos permite ao usuário identificar quais regiões em um território têm sido mais afetadas e, assim, direcionar políticas públicas, recursos e medidas de isolamento de forma mais efetiva no combate ao vírus.

⁶ Montagem a partir de imagens coletadas nas plataformas on-line CRC John Hopkins, Our World in Data e OMS.

Figura 6 – *Choropleth Maps*

Fonte: plataforma *online* CRC Johns Hopkins (2021)

Com este tipo de mapa de dados, foi possível verificar que os círculos são exibidos pela CRC em uma região geográfica designada com a área proporcional ao seu valor no conjunto de dados. Os mapas de bolhas que constam na plataforma da Johns Hopkins são adequados para comparar proporções associadas a regiões geográficas sem os problemas causados pelo tamanho da área regional. Nesse sentido, não constatamos sobreposições entre regiões diferentes do mapa, tendo em vista que essa é uma característica que requer cuidado para a escolha desse tipo de visualização.

5.2 Representações Temporais

As representações temporais são métodos de visualização que mostram os dados ao longo do tempo como uma forma de encontrar tendências ou mudanças nesse período. Pela característica contínua da pandemia, com início em dezembro de 2019 e final indeterminado, faz-se necessário o acompanhamento ao longo do tempo para que possam ser traçadas estratégias de controle e combate ao vírus. Devida, também, às atualizações diárias de casos, mortes, vacinação, entre outras variáveis, o uso dessas representações visuais é fundamental para a compreensão do cenário atual.

Todos os portais analisados utilizam essas representações visuais através de diversas técnicas, especialmente por meio de gráficos de linha e suas variações.

Os gráficos de linha são usados para exibir o desenvolvimento de valores quantitativos em um período. São comumente utilizados para mostrar tendências ao invés de comunicar valores específicos. A direção da linha no gráfico revela a tendência em um conjunto de dados: uma inclinação para cima indica que os valores estão em ascensão, enquanto que uma

inclinação para baixo revela queda. Quando agrupadas com outras linhas (outras séries de dados), as linhas individuais podem ser comparadas entre si. No entanto, o uso excessivo de séries de dados distintas em um mesmo gráfico pode poluir a representação visual, tornando sua leitura confusa.

Através dessas visualizações, é possível analisar o estágio em que se encontra determinado local no contexto da pandemia e, assim, instituir quais as medidas sanitárias mais adequadas a serem tomadas naquele momento, como o isolamento social ou a reabertura de espaços públicos. Essas orientações baseadas em dados, ou seja, em evidências, provaram ser fundamentais para o controle da disseminação do coronavírus.

Os portais analisados fazem uso de diversas representações visuais através de gráficos de linhas simples ou suas variações, como gráfico de área (Figura 7). Essas representações se tornaram populares, uma vez que foram referenciados por diversos veículos de comunicação e autoridades públicas, especialmente quando se referiam a termos como “achatamento da curva”, “onda” e “platô”. Esses termos, se utilizados sem o apoio visual, dificilmente poderiam ser compreendidos pelos mais diversos públicos a quem a mensagem era destinada.

Figura 7 – Gráficos de linha



Fonte: compilação dos autores⁷

Considerando a heterogeneidade e a desigualdade existente entre os diversos países, características que também se enquadram internamente em países como o Brasil, cada território tem sua própria curva epidemiológica. Há que se considerar, no entanto, que a falta ou a desatualização de dados oficiais podem provocar distorções na análise, gerando um

⁷ Montagem a partir de imagens coletadas nas plataformas on-line CRC John Hopkins, Our World in Data e OMS.

entendimento equivocado. Países com subnotificação em seus registros, provocada especialmente por ausência de políticas de testagem em massa e por deficiências nos setores de controle sanitário e de saúde, têm distorções entre a representação visual e a realidade.

Ao apresentar dados de forma ilustrativa e dinâmica, é possível analisar o cenário de forma rápida e compreensível pelos mais diversos perfis de usuários. É possível interagir com os gráficos apresentados, podendo o usuário escolher o que visualizar e, assim, customizar as comparações. Os painéis de visualização dos portais estudados são uma ferramenta eficiente e eficaz para a tomada de decisões baseada em evidências. O conhecimento fundamentado na análise, no entendimento e na divulgação dos dados permite estimar o impacto causado pelo vírus e como as medidas protetivas, como o isolamento social, podem diminuir a quantidade de vítimas.

6. Considerações finais

Apesar dos movimentos que avançam em busca de uma Ciência mais acessível, transparente, colaborativa, participativa e democrática, lidar com dados abertos pode significar um desafio para o cidadão comum. Assim, não basta apenas a disponibilização de conjuntos de dados se parte dos atores envolvidos não possui conhecimentos de mineração, tratamento e análise de dados brutos. Sem interpretação, os dados, ainda que acessíveis, não possuem qualquer valor. É necessário, portanto, ir além da democratização do acesso à Ciência. É fundamental que o cidadão seja capaz de se apropriar dos dados abertos, tomando-os para si e fazendo uso da forma que lhe for conveniente.

Nesse cenário, a visualização se manifesta como um importante mecanismo que favorece o entendimento dos conjuntos de dados disponibilizados. A visualização de dados é uma maneira eficiente de comunicar uma informação complexa. Os dados, inseridos em um contexto visual, podem sugerir relações entre as variáveis estudadas que não seriam facilmente identificadas se analisadas de outra forma. Padrões, tendências e correlações são reconhecidos mais facilmente por meio do uso de técnicas e ferramentas de visualização. Para Kirk (2012), a visualização é o meio possível de retratar os dados de maneiras que nos permitem vê-los sob nova luz, observar padrões, exceções e as possíveis histórias que estão por trás do seu estado bruto. É, em sintaxe, uma ferramenta de descoberta, um instrumento vantajoso no entendimento dos dados, seja na exploração, na análise ou na comunicação.

A visualização de dados sobre a Covid-19 apresenta novas formas de ontologia relacional. Atua no sentido de promover reconfigurações intersubjetivas no ambiente sociocultural das imagens apresentadas, servindo como uma contribuição para e-Science. Por meio dos casos expostos, buscou-se apresentar as potencialidades da visualização de dados na divulgação dos processos e dos resultados de pesquisa, expandindo o conhecimento científico para além dos muros da academia. Identificaram-se padrões nas representações visuais, assim como na disponibilização dos conjuntos de dados utilizados. Mediante os casos analisados, verifica-se que os infográficos são visualmente atraentes e seu impacto visual agrega valor aos dados disponíveis. As ilustrações são bem elaboradas e coerentes com o estilo visual e temático dos temas tratados. As informações são contextualizadas com a quantidade necessária de dados para a assimilação do conteúdo. As análises demonstram a utilidade comunicativa dessas visualizações de dados para comunicar informações de forma clara e eficiente através de gráficos estatísticos e gráficos informativos. Por meio dos casos selecionados, constatou-se que a visualização ajuda a analisar e raciocinar sobre dados e evidências. Também torna os dados complexos mais digeríveis, compreensíveis e utilizáveis. Em um contexto pandêmico, em que os dados se modificam de maneira frequente, é fundamental que as informações que chegam aos usuários, sejam estes pesquisadores, autoridades públicas, imprensa ou cidadãos comuns, possam ser compreendidas de forma célere. A visualização adequada permite a capacidade de síntese, convergindo para a compreensão de algo científico e auxiliando na tomada de decisão, se desdobrando em políticas apoiadas na ciência, tão essenciais para o enfrentamento ao vírus.

Transformar dados em conhecimento e conhecimento em ação são necessidades urgentes no contexto de uma pandemia. A visualização de dados mostrou ter um papel ativo de ajudar a entender questões complexas, sendo um recurso eficaz para a produção e disseminação de informação, podendo trazer novas possibilidades para o movimento da Ciência Aberta.

Referências

- ALBAGLI, Sarita. Ciência aberta em questão. *In*: ALBAGLI, Sarita; MACIEL, Maria Lucia; ABDO, Alexandre Hammud (Orgs.). **Ciência aberta, questões abertas**. Brasília: IBICT; Rio de Janeiro: UNIRIO, 2015. p. 9-25. Disponível em: http://livroaberto.ibict.br/bitstream/1/1060/1/Ciencia%20aberta_questoes%20abertas_PORTUGUES_DIGITAL%20%285%29.pdf. Acesso em: 10 jul. 2021.
- AMARAL, Fernando. **Aprenda mineração de dados: teoria e prática**. Rio de Janeiro: Alta Books, 2016a.
- AMARAL, Fernando. **Introdução à ciência de dados: mineração de dados e big data**. Rio de Janeiro: Alta Books, 2016b.
- CARD, Stuart; MACKINLAY, Jock; SHNEIDERMAN, Ben. **Readings in information visualization: using visualization to think**. [S. l.]: Morgan Kaufmann, 1999.
- ESTEVIÃO, Janete Saldanha Bach. **Letramento informacional para reuso de dados nas ciências sociais em ambientes virtuais de pesquisa: proposta de requisitos e competências**. 2019. Tese (Doutorado em Tecnologia e Sociedade) – Universidade Tecnológica Federal do Paraná, Curitiba, 2019. Disponível em: <http://repositorio.utfpr.edu.br/jspui/handle/1/4675>. Acesso em: 15 jul. 2021.
- FECHER, Benedikt; FRIESIKE, Sascha. Open science: One term, five schools of thought. *In*: BARTLING, Sönke; FRIESIKE, Sascha. (Orgs.). **Opening Science: the evolving guide on how the internet is changing research, collaboration and scholarly publishing**. Springer Open, 2014. p. 17-47. Disponível em: <https://link.springer.com/book/10.1007/978-3-319-00026-8>. Acesso em: 15 jul. 2021.
- FEW, Stephen. **Now you see it: simple visualization techniques for quantitative analysis**. [Estados Unidos]: Analytics Press, 2009.
- FOSTER. [Site institucional]. [2021]. Disponível em: <https://www.fosteropenscience.eu>. Acesso em: 15 jul. 2021.
- FREITAS, Carla *et al.* Introdução à visualização de informações. **Revista de informática teórica e aplicada**, Porto Alegre, v. 8, n. 2, p. 143-158, 2001. Disponível em: <https://lume.ufrgs.br/handle/10183/19398>. Acesso em: 22 maio 2021.
- GIL, Antônio Carlos. **Como elaborar projetos de pesquisa**. 5. ed. São Paulo: Atlas, 2010.
- GRÉGIO, André R. A. *et al.* Técnicas de visualização de dados aplicadas à segurança da informação. *In*: SIMPÓSIO BRASILEIRO DE SEGURANÇA DA INFORMAÇÃO E DE SISTEMAS COMPUTACIONAIS, 9., 2009, Campinas. **Anais [...]**. Campinas: UNICAMP, 2009. Disponível em: <http://www.lac.inpe.br/~rafael.santos/Docs/SBSEG/2009/sbseg2009.pdf>. Acesso em 23 maio 2021.
- ILIINSKY, Noah; STEELE, Julie. **Designing data visualizations: representing informational relationships**. [S. l.]: O'Reilly Media, 2011.

JOHNS HOPKINS UNIVERSITY. **Coronavirus Resource Center**. 2021. Disponível em: <https://coronavirus.jhu.edu/>. Acesso em: 15 jul. 2021.

KIRK, Andy. **Data visualisation**: a handbook for data driven design. [S. l.]: Sage, 2016.

KIRK, Andy. **Data visualization**: a successful design process. [S. l.]: Packt, 2012.

LAKATOS, Eva Maria; MARCONI, Marina de Andrade. **Técnicas de pesquisa**: planejamento e execução de pesquisas, amostragens e técnicas de pesquisas, elaboração, análise e interpretação de dados. 7. ed. São Paulo: Atlas, 2009.

MACHADO, Jorge. Dados abertos e ciência aberta. In: ALBAGLI, Sarita; MACIEL, Maria Lucia; ABDO, Alexandre Hannud (Orgs.). **Ciência aberta, questões abertas**. Brasília: IBICT; Rio de Janeiro: UNIRIO, 2015. p. 201-227. Disponível em: http://livroaberto.ibict.br/bitstream/1/1060/1/Ciencia%20aberta_questoes%20abertas_PORTUGUES_DIGITAL%20%285%29.pdf. Acesso em: 10 jul. 2021.

OPEN KNOWLEDGE FOUNDATION. [Site institucional]. 2021. Disponível em: <http://opendefinition.org/>. Acesso em: 3 jul. 2021.

OUR WORLD IN DATA. [Site institucional]. 2021. Disponível em: <https://ourworldindata.org/>. Acesso em: 15 jul. 2021.

PARRA, Henrique Z. M. Ciência cidadã: modos de participação e ativismo informacional. In: ALBAGLI, Sarita; MACIEL, Maria Lucia; ABDO, Alexandre Hannud (Orgs.). **Ciência aberta, questões abertas**. Brasília: IBICT; Rio de Janeiro: UNIRIO, 2015. p. 121-141. Disponível em: http://livroaberto.ibict.br/bitstream/1/1060/1/Ciencia%20aberta_questoes%20abertas_PORTUGUES_DIGITAL%20%285%29.pdf. Acesso em: 10 jul. 2021.

RICHARDSON, Roberto Jarry. **Pesquisa social**: métodos e técnicas. 4. ed. São Paulo: Atlas, 2017.

RININSLAND, Andrew et al. **D3.js**: cutting-edge data visualization. [S. l.]: Packt, 2017.

SATO, Susana Narimatsu. **A infografia na divulgação científica**: um estudo de caso da revista Pesquisa FAPESP. 2017, 155f. Dissertação (Mestrado em Ciência da Comunicação) – Escola de Comunicações e Artes, Universidade de São Paulo, São Paulo, 2017. Disponível em: <https://www.teses.usp.br/teses/disponiveis/27/27153/tde-07112017-155938/publico/SusanaNarimatsuSato.pdf>. Acesso em: 25 ago. 2019.

TUFTE, Edward. **The visual display of quantitative information**. 2. ed. [S. l.]: Graphics, 2001.

WORLD HEALTH ORGANIZATION. **WHO Coronavirus (COVID-19) dashboard**. 2021. Disponível em: <https://covid19.who.int/>. Acesso em: 17 dez. 2021.

YIN, Robert. **Estudo de caso**: planejamentos e métodos. 4. ed. Porto Alegre: Bookman, 2010.

Artigo submetido em: 24 fev. 2022

Artigo aceito em: 04 nov. 2022