

Big Data e Data for Good na Ciência da Informação: aproximação teórica

Patricia Siqueira Santos

Doutoranda em Ciência da Informação pelo Programa de Pós-Graduação em Ciência da Informação da Universidade Federal de Santa Catarina (UFSC). Florianópolis, Brasil.
patriciasiqueirass@gmail.com

Nathalia Berger Werlang

Doutora em Administração pela Universidade Federal de Santa Catarina (UFSC). Docente do Departamento de Ciência da Informação e do Programa de Pós-Graduação em Ciência da Informação da Universidade Federal de Santa Catarina (UFSC). Florianópolis, Brasil.
nathaliabw@gmail.com

Resumo

Objetivo: compreender de que forma os temas já estudados na Ciência da Informação estão inseridos nas discussões sobre Data for Good (dados para o bem), assim como explorar quais caminhos, com base nos GTs do Enancib, melhor dialogam com as publicações identificadas. **Metodologia:** é desenvolvida pesquisa qualitativa, com caráter exploratório e documental, por meio de revisão de literatura. A análise temática é utilizada para levantar as relações encontradas e são definidos quatro temas principais: Cidades, Educação, Saúde e Comunidade. Para completar a análise e atender ao objetivo proposto, é realizado cruzamento de dados dos artigos pesquisados com as informações das ementas dos grupos de trabalho do Encontro Nacional de Pesquisa em Ciência da Informação. **Resultados:** os trabalhos recuperados apresentam pesquisas com diversidade de métodos de coleta e análise. Tais pesquisas, mesmo sem fazer referência direta à Ciência da Informação, exploram temas que são transversais e de clara relevância para a área. Os quatro GTs que se destacam na relação feito com os artigos são GT 05 – Política e Economia da Informação; GT 11 – Informação & Saúde; GT 08 – Dados, Informação e Tecnologia e GT 06 – Informação, Educação e Trabalho, encontrando relação com os temas propostos. **Conclusões:** é possível traçar o paralelo de buscarem utilizar o Big Data como apoio à solução de problemas socioambientais, com levantamento de riscos, benefícios e indicação de boas práticas para tomada de decisões informadas.

Descritores: Big Data. Data for Good. Ciência da Informação. Revisão de Literatura.

Recebido em: 19.09.2025 | **Aceito em:** 13.05.2026

1 Introdução

A Ciência da Informação enquanto área de conhecimento investiga em seus

estudos as propriedades, comportamentos e fluxos da informação, além dos meios de processamento para otimização de acesso e uso (Borko, 1968).

O Big Data, caracterizado em última análise por Davenport (2014), como a coleta e interpretação de grandes conjuntos de dados, encontra interseção com a área em aspectos basilares, como a definição das ferramentas de análise, coleta, armazenamento, uso e descarte.

Dentro do escopo do Big Data e suas relações com a Ciência da Informação, investiga-se nesse trabalho os conceitos de Data for Good. É conduzida revisão de literatura para a busca de ligações entre os três temas através de artigos publicados e recuperados em três bases de dados: Scopus, Web of Science e ProQuest. As bases Scielo e Redalyc também foram consultadas, mas não apresentaram resultados às buscas.

De maneira geral, é bastante explorado na literatura científica na área de Ciência da Informação (CI) a utilização do Big Data para fins de geração de vantagem competitiva às organizações. Neste trabalho buscou-se por relações da CI e Big Data com o foco direcionado aos desafios e oportunidades da utilização de dados e tecnologias na sociedade, desenvolvimento e ação humanitária.

Para isso, definiu-se como pergunta de pesquisa a seguinte questão: quais áreas de estudos são proeminentes na relação entre os temas Big Data, Data for Good e Ciência da Informação? O objetivo foi compreender de que forma os temas já estudados na CI estão inseridos nas discussões sobre Data for Good, assim como compreender quais caminhos, com base nos Grupos de Trabalho (GTs) do Encontro Nacional de Pesquisa em Ciência da Informação (Enancib) melhor dialogam com as publicações identificadas na revisão de literatura.

Todos os artigos apresentados nos resultados tratam de conceitos sobre Data for Good, seja diretamente nomeando a iniciativa, ou então discutindo as aplicações, éticas ou governança de dados para benefícios públicos, comunitários ou humanitários.

2 Fundamentação Teórica

O aprofundamento teórico será apresentado para as temáticas de Big Data e Data for Good, considerando uma visão ampla das áreas e pontos específicos

encontrados nos artigos que compõem os resultados deste trabalho.

2.1 Big Data

O avanço da internet e uso de dispositivos foi o gerador do atual cenário da grande quantidade de dados disponíveis para análise, criando os padrões de navegação em sites, fluxo de cliques e compras realizadas, além dos conteúdos gerados nos diversos formatos pelos usuários em mídias sociais, blogs e plataformas (Silveira; Marcolin; Freitas, 2015). Existem também os dados gerados pelos próprios negócios, como bancos, financeiras, o setor da saúde, como exemplos.

Para que os dados sejam considerados Big Data é preciso questionar: o conjunto é gerado em alto volume, velocidade e com variedade? (Data Science Academy, 2025) Em caso positivo, encaixam-se nessa categoria.

Essa definição dos 3 Vs do Big Data foi criada por Doug Lancy (2001), quando o conceito ganhou relevância nos anos 2000. Volume significa a grande quantidade de dados gerados, Velocidade remete à rapidez de sua criação, em tempo real, e a Variedade se relaciona com as diversas fontes de onde chegam, como sensores, sistemas, aplicativos, redes sociais. Outros “Vs” podem ser encontrados na composição dessa definição de conceito, como Veracidade e Valor.

É necessário apontar, no entanto, a problematização no entorno do conceito, no que se refere à identificação do que é ou não Big Data. Como pergunta Davenport (2017, p. 17): “É verdade que essas características são importantes, mas e se você só tiver um ou dois dos Vs? Será que isso significa que você só tem um terço ou dois quintos do big data?”

O autor quer dizer, principalmente, que para aproveitar ao máximo esse recurso é preciso investigar alguns níveis abaixo. As organizações precisam muito mais entender como desconstruir o termo para definir o que é relevante para sua realidade e então poder direcionar suas estratégias de forma específica.

Como afirmam Patil e Mason (2015, p. 6), usar dados de forma eficaz não se trata apenas da escolha do melhor banco de dados ou quantos profissionais atuam na equipe, “mas sim uma interação complexa entre os dados que você possui, onde são armazenados, como as pessoas trabalham com eles e quais problemas são considerados dignos de solução”.

No contexto da utilização dos dados para o bem comum, a reflexão passa a ser a respeito da importância dos dados nos diversos contextos urbanos e sociais, sendo utilizados para a construção de cenários benéficos, com reflexões aprofundadas sobre privacidade, ética, correta governança e utilização.

2.2 Data For Good

Uma pesquisa na Web em busca de resultados sobre o Data for Good resgata informações sobre um ambiente internacional robusto em ações. A começar pela organização fundadora, a Data Kind (ICE, 2018), que desenvolve atividades em combate aos maiores desafios mundiais usando ciência de dados e inteligência artificial (Data Kind, 2025), contando com uma equipe ampla e diversa.

Nos Estados Unidos é possível encontrar até mesmo uma profissionalização da área, como a existência de consultorias especializadas ou projetos em modelo de desafios. Com a chamada "Competições em Data Science para construir um mundo melhor", a organização DrivenData trabalha com projetos que reúnem Data Science (DS) e impacto social em áreas como desenvolvimento internacional, saúde, educação, pesquisa e conservação e serviços públicos.

O processo funciona com o levantamento de práticas de ponta em DS para resolver os maiores desafios sociais. São ações que duram em média 2 a 3 meses, abertas a uma comunidade global de cientistas de dados que concorrem a premiações pelas melhores soluções (DrivenData, 2025).

No Brasil, a organização Social Good Brasil (SGB), criada em 2012 e ainda atuante, desenvolve atualmente um trabalho voltado a “metodologias em que dados e novas tecnologias são utilizados de forma consciente, ética e voltada para o bem, gerando impacto socioambiental positivo em todo o país” (SGB, 2025). São desenvolvidos projetos, produtos e formações, na busca de gerar impacto social e ambiental.

São exemplos de ações as formações Cidadão de Dados e Laboratório de Dados, para aplicação dos conhecimentos na vida pessoal, profissional, assim como por institutos, fundações, empresas, empreendedores(as) sociais e governo.

O Festival SGB é uma ação proeminente realizada anualmente em Florianópolis desde 2012, promovendo conexões através de palestras, apresentações

e discussões. Reúne pessoas e iniciativas referência em tecnologia, dados e impacto social.

O SGB já trabalhou com temáticas diversas, muito voltado no início a novas tecnologias e mídias para impacto socioambiental. Desde 2017 o trabalho tem sido voltado à tendência do uso dos dados para o bem.

Na literatura científica também é possível encontrar reflexões sobre o conceito Data for Good. O trabalho de Mann e Sahuguet (2018) reflete sobre como o Data Science pode ser aplicado para resolver interesses públicos. Trazem uma interessante visão sobre o quanto a ideia de Data Science para o bem social poderia ser vista como um pleonasma, considerando a ênfase redundante da ciência ser para o bem. No entanto, a observação das aplicações em Data Science demonstra um direcionamento maior das ações ao setor privado, deixando o setor público - com maior potencial em promover mudanças sociais - em desvantagem.

Sendo assim, afirmam que parte do grande interesse em torno do “Data Science for social good” se dá pela atenção que o movimento dá a problemas que não seriam atendidos pelos interesses do mercado, como os problemas do setor público, ou das pessoas que não possuem poder de compra.

No artigo os autores têm objetivo de apresentar a conferência Data for Good Exchange (D4GX), que reúne anualmente uma audiência composta por cientistas de dados, pesquisadores, indústria organizações sem fins lucrativos e governo. A conferência teve início em Nova York em 2014 e em 2021 foi lançada a edição asiática.

As discussões da Conferência defendem que o acesso aos dados é crucial para empoderar as comunidades. Permite a tomada de decisões informadas, promove participação, apoia o planejamento e alocação de recursos, promovendo inovação e desenvolvimento. O movimento busca quebrar as barreiras de acesso e promover inclusão, por meio dos diálogos, programas, estudos e desafios (D4GX India, 2025).

Outro estudo importante, de Chandy, Hassan e Mukherji (2017) analisa como as inovações envolvendo Big Data apoiam a resolução dos maiores desafios da atualidade, com foco nos países em desenvolvimento. Explora como o grande volume de dados pode ajudar tomadores de decisão a melhor encaminhar questões como a pobreza, doenças, conflitos, migração, poluição, entre outros.

Os autores partem do pressuposto de que o Big Data pode ter profundo impacto

para o bem e, mais do que isso, as inovações atingem maior impacto quando resolvem urgências de um grande número de pessoas. Para isso, tal impacto pretendido requer investimento em mercados emergentes no que tange ao armazenamento de dados, sua integração e disseminação.

Um ponto principal abordado trata do fato de que muitas discussões em torno do Big Data estão focadas no impacto gerado em economias desenvolvidas, muito porque a maior parte dos dados é gerada nesses locais onde as fontes digitais são totalmente presentes. Nesse aspecto, as oportunidades de lucro desse setor aumentam quanto mais desenvolvida é a economia, uma vez que existem clientes com poder aquisitivo para tal mercado.

É dada ênfase, no entanto, no fato de que as soluções baseadas em dados não possuem caráter meramente comercial, mas guardam potencial de ajudar a resolver os mais profundos e importantes problemas enfrentados pela humanidade. Tal articulação se dá pelas ações complementares de entidades governamentais, ONGs, pesquisadores(as), assim como organizações que buscam lucratividade.

3 Metodologia

Trata-se de uma pesquisa qualitativa, com caráter exploratório e documental. Para atender ao objetivo de sintetizar e avaliar criticamente o conhecimento existente no campo pesquisado, foi adotado o método de revisão de literatura, cujas buscas foram realizadas em fevereiro de 2025. Buscou-se responder a seguinte questão de pesquisa: Quais áreas de estudos são proeminentes na relação entre os temas Big Data, Data for Good e Ciência da Informação?

Os resultados a serem analisados foram recuperados de três bases de dados: Scopus, Web of Science e ProQuest. A base Scielo não apresentou resultados para as buscas e a base Redalyc resgatou um resultado, que já havia sido identificado, sendo então desconsiderado.

Na base de dados Scopus, a *string* de busca utilizada foi a seguinte: ALL ("big data" AND "data for good" AND "information science"). O resultado inicial foi de 99 trabalhos. O filtro *Document type* foi utilizado para a escolha de artigos, o que reduziu a quantidade recuperada para 59 documentos.

A base ProQuest resgatou 16 resultados através da seguinte *string*, com a

busca em todos os campos: ("data for good" AND "information science" AND "big data").

Na base Web of Science a *string All Fields* ("data for good" AND "information science" AND "big data") não obteve resultados. A busca por ("data for good" AND "information science") obteve 2 resultados.

O resumo dos resultados é apresentado no Quadro 01 abaixo:

Quadro 01 – Resultados recuperados por base

Base	Campos	Resultados
Scopus	ALL - Artigos	59
ProQuest	Todos	16
Web of Science	All Fields	2
Total		77

Fonte: Elaboração própria (2026).

A estratégia de análise iniciou com a exportação das informações dos artigos encontrados, para leitura dos títulos e resumos das publicações. Esta avaliação permitiu observação da presença ou não dos temas pesquisados. Os critérios de inclusão foram os seguintes: 1) Trabalhos com a temática Big Data, ou Data, ou variações no título ou resumo; 2) Trabalhos com a temática Data for Good ou variações no título ou resumo; 3) Trabalhos relacionando as temáticas Data for Good ou variações com campos de aplicação diversos.

Os artigos que não corresponderam a estes critérios foram retirados do filtro e os que atenderam receberam um código proveniente da análise temática.

Após a primeira leitura de títulos e resumos da base dados Scopus, dos 59 artigos recuperados, 14 foram considerados para a análise final. Entre os 16 resultados da base ProQuest, 3 foram selecionados para seguimento da análise. A busca na Base Web of Science recuperou dois resultados, sendo apenas um mantido para análise.

Para fortalecer a resposta à questão norteadora, realizou-se cruzamento de dados entre os termos dos resumos e as palavras-chave dos 18 artigos com os termos das ementas dos Grupos de Trabalho do Enancib, de acordo com informações disponibilizadas pela Associação Nacional de Pesquisa e Pós-graduação em Ciência da Informação (ANCIB, 2026).

Após fornecidas as fontes, foi construído o seguinte prompt no NotebookLM, ferramenta de análise utilizada na pesquisa: "Análise os resumos e palavras-chave

dos 18 artigos e compare os termos com as ementas dos Grupos de Trabalho (GT) do Enancib (fonte fornecida). Monte um quadro indicando em qual GT cada artigo se encaixa e quais foram os termos que se alinharam para esta definição”. Após conferência dos resultados apresentados e ajustes na tabela fornecida, as relações apontadas são trazidas na seção de resultados.

4 Resultados

Os resultados são apresentados a partir de códigos e temas identificados pelo método de análise temática. Este é um método qualitativo que permite a identificação de padrões, significados repetidos nos dados coletados. Os temas gerados são captados através de trechos (códigos) do material analisado que respondem à questão de pesquisa (Braun; Clarke, 2006).

De acordo com as autoras, a fase de planejamento é composta por seis etapas, que foram conduzidas neste trabalho como descritas a seguir: (1) familiarização com os dados; que significa uma imersão na coleta realizada. (2) geração de códigos iniciais, que são os rótulos que descrevem características interessantes dos dados, relevantes à pergunta de pesquisa. (3) busca por temas: a partir da lista de códigos, estes são agrupados em potenciais temas, que capturam uma ideia ou padrão importante sobre os dados. (4) revisão de temas: é necessário verificar se os temas estão conectados de forma lógica, sem sobreposições ou lacunas, sendo possível realizar descarte, junções, divisões, novos nomes. (5) definição e nomeação dos temas: descrição de sua abrangência. (6) produção do relatório: redação da análise.

Como já apontado, na base dados Scopus, dos 59 artigos recuperados, 14 foram selecionados, conforme apresentado no Quadro 02 abaixo, já com os códigos definidos para cada artigo. A coluna ID corresponde a um identificador dos artigos, para conexão à etapa de cruzamento de dados:

Quadro 02 – Resultados base Scopus

ID	Título	Código área de estudo
1	The data awareness framework as part of data literacies in K-12 education	Melhora de experiência educacional adolescente
2	'Don't make assumptions about me!': understanding children's perception of datafication online	Dataficação: percepção infantil
3	Critical datafication literacy – a framework for educating about datafication	Dataficação: educação crítica

4	The fading gloss of data science: towards an agenda that faces the challenges of big data for development and humanitarian action	Desafios para desenvolvimento e ação humanitária
5	The uneven burden of COVID-19 in the metropolitan region of São Paulo, Brazil – Risk analysis from a bottom-up perspective	Saúde e vulnerabilidade social
6	Smart citizens: from data providers to decision-makers? the case study of Barcelona	Cidades inteligentes e cidadania
7	The resilience to emergencies and disasters index: applying big data to benchmark and validate neighborhood resilience capacity	Gestão de emergência
8	Patterns of population displacement during mega-fires in California detected using Facebook Disaster Maps	Mobilidade populacional e gestão de desastres
9	A graph-based approach for population health analysis using Geo-tagged tweets	Análise de saúde populacional
10	Development strategy based on combination typologies of building carbon emissions and urban vibrancy – a multi-sourced data-driven approach in Beijing, China	Emissão de carbono e políticas de desenvolvimento urbano
11	The critical importance of citizen science data	Dados e cidadania
12	Data sharing for the common good – remarks on the data altruism framework established under the data governance act	Compartilhamento voluntário de dados para bem comum
13	Understanding social impact of data on local communities	Utilização comunitária de dados
14	Using crowdsourced data to assess the impact of shelter-in-place orders (SIPO) during COVID-19: Lessons from a small island developing state	Mobilidade populacional

Fonte: Elaboração própria (2026).

Entre os 16 resultados da base ProQuest, 3 foram selecionados para seguimento da análise, conforme apresentado no Quadro 03:

Quadro 03 – Resultados base ProQuest

ID	Título	Código área de estudo
15	Development of big data-analysis pipeline for mobile phone data with mobipack and spatial enhancement	Mobilidade populacional e tomada de decisões
16	Analysis of spatiotemporal mobility of shared-bike usage during COVID-19 pandemic in Beijing	Mobilidade populacional e tomada de decisões
17	Living in a pandemic: changes in mobility routines, social activity and adherence to COVID-19 protective measures	Mobilidade populacional e tomada de decisões

Fonte: Elaboração própria (2026).

A busca na Base Web of Science recuperou dois resultados, sendo apenas um mantido para análise, conforme indicado no Quadro 04:

Quadro 04 – Resultados base Web of Science

ID	Título	Código área de estudo
18	A framework for building pro-bono Data for Good projects	Importância projetos data for good

Fonte: Elaboração própria (2026).

Os códigos identificados na leitura dos títulos e resumos foram agrupados em temas e são apresentados no Quadro 05 a seguir:

Quadro 05 – Temas definidos para os códigos

Códigos	Tema
Cidades inteligentes e cidadania	Cidades
Gestão de emergência	
Deslocamento populacional e gestão de desastres	
Emissão de carbono e políticas de desenvolvimento urbano	
Dados e cidadania	
Movimentação populacional	
Mobilidade populacional e tomada de decisões	
Mobilidade populacional e tomada de decisões	
Mobilidade populacional e tomada de decisões	
Melhora de experiência educacional adolescente	Educação
Dataficação: educação crítica	
Dataficação: percepção infantil	
Saúde e vulnerabilidade social	Saúde
Análise de saúde populacional	
Desafios para desenvolvimento e ação humanitária	
Compartilhamento voluntário de dados para bem comum	Comunidade
Utilização comunitária de dados	
Importância projetos data for good	

Fonte: Elaboração própria (2026).

O tema Cidades foi o mais presente, com nove artigos compilados. Também foram identificados os temas Comunidade (quatro artigos), Saúde (dois artigos) e Educação (três artigos).

Abaixo são apresentadas as descrições resumidas das pesquisas recuperadas e agrupadas nesta análise. O software NotebookLM, do Google, foi utilizado para organização dos conteúdos e apontamento de relações possíveis entre os achados. Os textos foram construídos a partir dos resultados apresentados pela ferramenta, não apenas replicados.

Cidades

Os artigos que compõem o tema Cidades foram encaixados desta forma por

apresentarem pesquisas ligadas ao escopo urbano.

O artigo **“Smart citizens: from data providers to decision-makers? the case study of Barcelona”** (Calzada, 2018) explora como Barcelona está redefinindo seu modelo de cidade inteligente ao focar na soberania dos dados e na participação cidadã, analisando as iniciativas de governança de dados da cidade. Os resultados mostram que essas iniciativas conseguiram descentralizar a gestão de dados e promover maior transparência na administração pública.

“The resilience to emergencies and disasters index: applying big data to benchmark and validate neighborhood resilience capacity” (Kontokosta; Malik, 2018) apresenta o índice REDI, uma métrica que avalia a resiliência de bairros urbanos diante de desastres. Como caso de estudo, o furacão Sandy em Nova York foi analisado para validar o modelo. Os resultados indicam que bairros com menor escore no índice REDI tiveram mais dificuldades em se recuperar do desastre.

“Patterns of population displacement during megafires in California detected using Facebook Disaster Maps” (Jia *et al.*, 2020) analisou o deslocamento populacional durante os incêndios Mendocino Complex e Woolsey, na Califórnia, usando dados do Facebook Disaster Maps. O artigo avaliou a eficácia desses dados para monitorar evacuações e retornos populacionais, aplicando métodos estatísticos para identificar padrões espaciais e temporais de migração. Os resultados indicam que áreas densamente povoadas foram evacuadas mais rapidamente devido ao melhor acesso ao transporte, enquanto regiões menos povoadas tiveram deslocamentos mais prolongados.

“Development strategy based on combination typologies of building carbon emissions and urban vibrancy – A multi-sourced data-driven approach in Beijing, China” (Xia; Wang; Lai, 2024) examina a relação entre emissões de carbono e a vitalidade urbana em Pequim, utilizando dados de múltiplas fontes. O objetivo é desenvolver uma abordagem estratégica para equilibrar o crescimento urbano com a sustentabilidade ambiental. Os resultados mostram que diferentes áreas da cidade exibem combinações únicas de emissões de carbono e vitalidade urbana e o estudo destaca a importância de integrar planejamento urbano e políticas ambientais para um desenvolvimento sustentável.

“The critical importance of citizen science data” (Sherbinin *et al.*, 2021) discute como dados coletados por cidadãos podem contribuir significativamente para

a pesquisa científica e a formulação de políticas públicas. Os resultados indicam que a qualidade dos dados de ciência cidadã muitas vezes iguala ou supera a de fontes tradicionais, desde que sejam aplicadas boas práticas de validação.

“Using crowdsourced data to assess the impact of shelter-in-place orders (SIPO) during COVID-19: Lessons from a small island developing state” (Mohan; Ramsawak, 2024) investiga como as ordens de isolamento social afetaram a mobilidade populacional em Trinidad e Tobago durante a pandemia. Utilizando dados crowdsourced do Facebook Data for Good, os pesquisadores aplicaram métodos estatísticos para medir a adesão da população às restrições. Os resultados mostram que as medidas mais eficazes para reduzir a mobilidade foram o fechamento de negócios e espaços públicos.

“Development of big data-analysis pipeline for mobile phone data with mobipack and spatial enhancement” (Witayangkurn; Arai; Shibasaki, 2022) apresenta o desenvolvimento de um pipeline de análise de big data para dados de telefonia móvel, com objetivo de permitir a ingestão, processamento e análise de grandes volumes de dados. A arquitetura do pipeline foi projetada para ser flexível, garantindo suporte a diversas fontes de dados e permitindo sua utilização em diferentes aplicações, como monitoramento de mobilidade e planejamento urbano.

“Analysis of spatiotemporal mobility of shared-bike usage during COVID-19 pandemic in Beijing” (Chai *et al.*, 2021) analisa a mobilidade urbana em Pequim durante a pandemia de COVID-19 por meio do uso de dados de compartilhamento de bicicletas. O objetivo foi entender como as restrições impostas afetaram os deslocamentos e quais foram os padrões emergentes. Os resultados mostraram que o uso de bicicletas compartilhadas caiu 64,8% durante o pico da pandemia, refletindo as restrições e o medo do contágio. No entanto, houve um aumento gradual de 15,9% posteriormente, indicando uma recuperação parcial das atividades econômicas e sociais.

“Living in a pandemic: changes in mobility routines, social activity, and adherence to COVID-19 protective measures” (Lucchini *et al.*, 2021) examina como a pandemia da COVID-19 alterou os padrões de mobilidade e interação social das pessoas ao longo do tempo. Usando dados de GPS anonimizados de mais de 837.000 indivíduos nos EUA, os pesquisadores analisaram como as medidas de contenção (NPIs) influenciaram as rotinas diárias. Os dados revelaram que a pandemia reduziu

drasticamente as visitas a pontos de interesse, com quedas de até 69% em Nova York.

Educação

O tema Educação reúne três artigos que tratam diretamente de questões educacionais, ambos a respeito de literacia de dados.

No artigo **“The data awareness framework as part of data literacies in K-12 education”** (Höper; Schulte, 2023) é apresentada uma estrutura projetada para aumentar a compreensão dos alunos sobre práticas de coleta e processamento de dados em artefatos digitais. Os resultados indicam que os alunos que receberam a intervenção pedagógica baseada no framework demonstraram uma maior compreensão sobre a coleta e o processamento de dados. Além disso, o estudo sugere que o framework pode ajudar os alunos a desenvolverem uma postura mais reflexiva e crítica sobre suas interações digitais, aumentando sua capacidade de tomar decisões informadas e autodeterminadas.

“Critical datafication literacy – a framework for educating about datafication” (Sander, 2024) propõe um framework para a alfabetização crítica da dataficação, combinando pesquisas teóricas e empíricas. O objetivo é criar uma estrutura educacional que capacite os alunos a compreenderem sistematicamente a dataficação, estimular o pensamento crítico e permitir ações informadas. Os resultados mostram que muitos educadores e criadores de recursos educacionais promovem uma abordagem crítica para ensinar sobre dataficação, enfatizando a emancipação, a ação coletiva e a compreensão sistêmica.

O estudo **“‘Don't make assumptions about me!': Understanding Children's Perception of Datafication Online”** (Wang *et al.*, 2022) explora como crianças de 7 a 13 anos no Reino Unido percebem a dataficação online, focando especialmente em como plataformas digitais coletam, analisam e inferem informações sobre elas. Os resultados sugerem que as crianças não compreendem plenamente como seus dados são usados para gerar inferências sobre elas. O estudo destaca a necessidade de mecanismos mais transparentes e de um modelo de governança de dados que proteja melhor as crianças.

Saúde

No tema Saúde são agrupados trabalhos que tratam de questões de saúde populacional.

“The uneven burden of COVID-19 in the metropolitan region of São Paulo, Brazil – Risk analysis from a bottom-up perspective” (Santos *et al.*, 2024) analisou a desigualdade no impacto da COVID-19 na região metropolitana de São Paulo, utilizando uma abordagem mista de métodos quantitativos e qualitativos. Os resultados mostraram que a pandemia afetou desproporcionalmente populações vulneráveis, reforçando desigualdades sociais já existentes. Os grupos focais indicaram que moradores de áreas centrais possuíam mais recursos para lidar com a crise, enquanto comunidades periféricas enfrentaram desafios maiores, como acesso limitado a serviços de saúde e dificuldades econômicas.

“A graph-based approach for population health analysis using Geo-tagged tweets” (Nguyen *et al.*, 2021) propõe uma abordagem baseada em grafos para a análise de saúde populacional utilizando dados de mídias sociais, especificamente tweets georreferenciados. Os resultados mostraram que a abordagem utilizada superou os métodos convencionais na previsão de índices de saúde e classificação de regiões com base na situação de saúde. É destacado o potencial de mídias sociais como ferramenta de monitoramento em tempo real para políticas de saúde pública.

Comunidade

O tema Comunidade recebe pesquisas ligadas à ação humanitária, questões comunitárias e públicas.

O artigo **“The fading gloss of data science: towards an agenda that faces the challenges of big data for development and humanitarian action”** (Gutierrez; Bryant, 2022) discute as promessas e desafios do uso de ciência de dados para desenvolvimento e ação humanitária. Embora muitas organizações, como as Nações Unidas, tenham adotado o uso de big data para apoiar decisões políticas, o estudo aponta que dados podem excluir populações vulneráveis e, em alguns casos, aumentar sua exposição a riscos. Propõe um conjunto de princípios para garantir que a ciência de dados seja utilizada de forma equitativa e eficaz, evitando armadilhas como anonimização ineficiente e concentração de dados em plataformas privadas.

“Data sharing for the common good – remarks on the data altruism framework established under the data governance act” (Michałowicz, 2024)

examina o conceito de "altruísmo de dados" dentro do Data Governance Act (DGA) da União Europeia. A pesquisa explora como a legislação incentiva a partilha voluntária de dados pessoais e não pessoais para fins de interesse público, como saúde, pesquisa científica e resposta a desastres. O estudo sugere que um maior alinhamento regulatório e incentivos para a adesão voluntária podem ampliar o impacto positivo do altruísmo de dados na sociedade.

“Understanding social impact of data on local communities” (Yoon; Copeland, 2019) analisa o impacto social da utilização de dados em comunidades locais, explorando como o uso de dados pode influenciar decisões comunitárias, identificar problemas e promover mudanças. O estudo enfatiza a necessidade de minimizar desigualdades no acesso aos dados e capacitar comunidades para que possam usufruir plenamente do potencial informacional disponível.

“A framework for building pro-bono data for good projects” (Monteiro; Maia, 2022) apresenta o framework criado e implementado com sucesso pelo Data Science for Social Good Portugal (DSSG PT), uma comunidade aberta de entusiastas de dados que trabalham pro bono em projetos de Data for Good. Além disso, apresenta uma metodologia para medir o impacto social dos projetos e garantir padrões éticos, como privacidade de dados e justiça.

Para aprofundar a análise de relação dos artigos com as temáticas na Ciência da Informação no que se refere às ementas dos grupos de trabalho do Enancib, apresenta-se cruzamento de dados no Quadro 06 abaixo:

Quadro 06 – Cruzamento de dados artigos e GTs Enancib.

ID	Termos do Artigo (Resumo/Palavras-chave)	Termos da Ementa do GT	GT Enancib
13	<i>Social impact, community informatics, open data, data reuse.</i>	Apropriação da informação; mediação e protagonismo social; inclusão informacional e social.	GT 03 – Mediação, Circulação e Apropriação da Info.
18	<i>Data for good projects, framework, project management, formal organization.</i>	Gestão de ambientes, serviços e produtos de informação; processos de trabalho e gestão.	GT 04 – Gestão da Informação e do Conhecimento
2	<i>Datafication, algorithms, children, security and privacy, algorithmic literacy.</i>	Vigilância, privacidade e proteção de dados; usos políticos e econômicos de algoritmos.	GT 05 – Política e Economia da Informação

4	<i>Humanitarian action, algorithmic bias, data infrastructure, privacy.</i>	Ética da informação; vigilância; usos de algoritmos e IA; ecologia e transformação social.	
6	<i>Smart citizens, policy, GDPR, technopolitics, data ownership, data commons.</i>	Leis e políticas de informação; vigilância, privacidade e proteção de dados; ética da informação.	
10	<i>Carbon emissions, sustainability, urban policy, sustainable urban planning.</i>	Ecologia, meio ambiente e sustentabilidade; Estado, leis e políticas de informação.	
12	<i>Data altruism, Data Governance Act, laws, data sharing, general interest.</i>	Estado, leis e políticas de informação; economia política da informação; proteção de dados.	
1	<i>Data literacy, K-12 education, data awareness, computing education.</i>	Competência em informação (habilidades e dimensões); educação; experiências pedagógicas.	GT 06 – Informação, Educação e Trabalho
3	<i>Critical data literacy, educating about datafication, critical pedagogy, media literacy.</i>	Relações entre informação e educação; bases curriculares e experiências pedagógicas.	
11	<i>Citizen science data, open data, scientific research, data life cycle.</i>	Práticas científicas, ciência aberta, comunicação e uso da informação em Ciência, Tecnologia e Inovação.	GT 07 – Produção e Comunicação da Info. em CT&I
7	<i>Big data, resilience planning, spatial data, anomaly detection, data integration.</i>	Processamento, tratamento, análise, visualização de dados em ambientes informacionais.	GT 08 – Dados, Informação e Tecnologia
8	<i>Crowdsourced data, social media, spatial and temporal pattern, data generation.</i>	Processamento, geração, coleta, uso e análise de dados em ambientes informacionais.	
15	<i>Big data-analysis pipeline, CDR data, spatial enhancement, processing.</i>	Desenvolvimento de tecnologias de informação; armazenamento, gestão, processamento e tratamento de dados.	
5	<i>COVID-19, public health, vulnerability, inequality, focus group.</i>	Informação relacionada à saúde humana; população; informação e saúde.	GT 11 – Informação & Saúde
9	<i>Population health analysis, public health,</i>	Ecossistema da informação em saúde; tecnologias digitais no	

	<i>geo-tagged tweets, health surveys.</i>	âmbito da informação em saúde.
14	<i>COVID-19, public health policy, SIPO, crowdsourced data.</i>	Informação relacionada à saúde humana; atores da saúde, gestores e população.
16	<i>COVID-19 pandemic, epidemiological research, public health, spatial temporal analysis.</i>	Fenômenos e processos informacionais associados à saúde; pesquisa em informação em saúde.
17	<i>COVID-19, epidemic models, protective measures, human behavior.</i>	Comportamentos informacionais em saúde; processos informacionais associados a pacientes/população.

Fonte: Elaborado com base nos artigos pesquisados e ANCIB (2026).

5 Discussão

A pergunta de pesquisa trazida para a coleta de dados foi: quais áreas de estudos são proeminentes na relação entre os temas Big Data, Data for Good e Ciência da Informação?

Os estudos selecionados demonstram uma diversidade de aplicações para os dados em pesquisas que visam contribuir com soluções de problemáticas sociais. São pesquisas que, a princípio, não citam ciência da informação diretamente, mas abordam temáticas que perpassam a área.

Os artigos relacionados à temática de Cidades tratam de pontos como densidade e monitoramento de fluxo populacional, dinamismo e planejamento urbano, previsão de vulnerabilidade e tomadas de decisões em situações de risco, além de participação cidadã e transparência.

Os estudos na área da Educação estão focados em desenvolvimento de postura crítica, reflexiva, compreensão sistêmica do universo dos dados e emancipação digital, em busca de oferecer saídas para um melhor uso e aproveitamento dos recursos oferecidos pelo Big Data.

As pesquisas na área da Saúde apresentam diversidade entre sim, com temáticas como análise do impacto da COVID-19 na saúde pública e monitoramento geolocalizado para apoio a decisões em políticas públicas de saúde.

O tema Comunidade trata da utilização em si do Big Data, seu papel e importância, assim como pontos de atenção e riscos, como a exclusão de populações vulneráveis, desigualdade de acesso. São estudos que propõem caminhos para abordagens mais equitativas e éticas, privacidade e ética na utilização dos dados.

De forma geral, são áreas de estudo e atuação governamental que visam a descoberta de soluções para problemas socioambientais e grande escala, que conseguem ser amparados com o auxílio de Big Data, devido à sua característica de grande escala e possibilidade de decisões informadas.

No que se refere ao cruzamento de dados das palavras-chave dos artigos com os termos das ementas dos GTs do Enancib, a análise automatizada apontou relações com sete GTs, mais fortemente com três deles: GT 05 – Política e Economia da Informação e GT 11 – Informação & Saúde, com 5 artigos relacionados em cada; e GT 08 – Dados, Informação e Tecnologia, com 3 artigos. O GT 06 – Informação, Educação e Trabalho teve dois artigos relacionados e os demais apenas um artigo.

Correspondem à análise temática os temas de Saúde e Educação, sendo os temas Cidades e Comunidade também correspondentes ao contexto dos demais GTs relacionados.

Considera-se que as discussões sobre Big Data em articulação ao Data for Good não se mostram como uma temática externa à Ciência da Informação, mas sim como uma extensão tecnológica das problemáticas já tratadas e até mesmo enraizadas na área, conforme apontam as ementas dos grupos de trabalho do maior evento do campo, ambiente formal de discussão e fortalecimento da CI.

Um bloco significativo da pesquisa recente sobre Data for Good é aplicado à saúde pública (GT 11), lidando com o ecossistema e ferramentas informacionais em saúde aplicadas à gestão e monitoramento na pandemia de COVID-19, evento que impulsionou o avanço de muitas pesquisas.

Na mesma medida, são apontados textos que abordam as leis que regem os dados (como a Governança e Altruísmo de Dados), o viés algorítmico, os perigos da dataficação para populações vulneráveis (crianças) e as políticas de sustentabilidade (GT 05).

Como o que se propõe no GT 08, há artigos que discutem o tratamento, a infraestrutura técnica para lidar com desastres, arquitetura e processamento de grande volume de dados. Já no GT 06, encontra-se a relação sobre como o

movimento Data for Good não foca apenas na aplicação dos dados, mas em como ensinar cidadãos a consumi-los e compreendê-los.

6 Considerações Finais

A análise crítica dos resultados desta revisão de literatura aponta que a relação entre Big Data, Data for Good e Ciência da Informação deve ser compreendida como uma articulação de problemáticas já enraizadas na área, na busca de soluções de desafios sociais complexos e a necessidade de uma governança de dados mais ética e responsável.

Respondendo à questão de pesquisa, sobre as áreas de estudos proeminentes na relação entre os temas, identificou-se que os temas Cidades, Saúde, Educação e Comunidade, destacam-se na literatura atual. No entanto, embora os artigos analisados tratem de temáticas transversais e de clara relevância para a CI, esta não é citada diretamente, o que demonstra a natureza interdisciplinar da área, que pode ser reconhecida de maneira transversal.

Ao fazer o mapeamento e diálogo do Data for Good na CI com os Grupos de Trabalho (GTs) do Enancib, verificou-se que o foco da CI não reside apenas no aspecto técnico, como a infraestrutura e o processamento de grande volume de dados (GT 08). Pelo contrário, a maior relevância está nas implicações éticas, políticas e sociais desses fenômenos, evidenciada pela proeminência do GT 05 (Política e Economia da Informação), que aborda leis, altruísmo de dados e viés algorítmico, e do GT 11 (Informação & Saúde), impulsionado pela gestão e monitoramento na pandemia de COVID-19.

O ponto central da reflexão diante dos achados reside na problematização inerente ao movimento Data for Good: o uso de Big Data é um recurso essencial para a tomada de decisões informadas voltadas à solução de problemas sociais, mas pode, simultaneamente, exacerbar desigualdades. Enquanto as pesquisas avançam com o uso de novas fontes de dados (como tweets georreferenciados e mobilidade urbana), a realidade da invisibilidade de populações vulneráveis devido à ausência de infraestrutura digital se mantém.

A CI tem condições, portanto, de assumir um papel mediador, uma vez que suas pesquisas dialogam com aspectos diversos e campos que se complementam,

para buscar o equilíbrio entre o acesso e a disponibilidade total de dados. As discussões sobre literacia de dados e literacia de dataficação (GT 06) se tornam cruciais para promover a soberania de dados e o empoderamento dos cidadãos, garantindo que o Big Data para o bem comum seja, de fato, equitativo.

Para futuras pesquisas, é possível aprofundar o estudo de termos como responsabilidade algorítmica e altruísmo de dados, traduzindo bases teóricas da CI, de organização e uso da informação, em ações práticas com impacto social tangível.

Referências

ASSOCIAÇÃO NACIONAL DE PESQUISA E PÓS-GRADUAÇÃO EM CIÊNCIA DA INFORMAÇÃO (ANCIB). Coordenações e Ementas de GT. 2026. Disponível em: <https://ancib.org/coordenacoes-e-ementas-de-gt/>. Acesso em: 7 maio 2026.

BORKO, H. Information Science: what is it? **American Documentation**, New Jersey, v. 19, n. 1, p. 3-5, 1968. Disponível em: <https://doi.org/10.1002/asi.5090190103>. Acesso em: 21 ago. 2025.

BRAUN, V.; CLARKE, V. Using thematic analysis in psychology. **Qualitative Research in Psychology**, [S.l.], v. 3, n. 2, p. 77-101, 2006. Disponível em: <https://www.tandfonline.com/doi/abs/10.1191/1478088706qp063oa>. Acesso em: 13 mar. 2025.

CALZADA, I. Smart Citizens: From Data Providers to Decision-Makers? The Case Study of Barcelona. **MDPI Sustainability**, 10(9), 3252, p. 1-25, 2018. Disponível em: <https://doi.org/10.3390/su10093252>. Acesso em: 09 ago. 2026.

CHAI, X.; GUO, X.; XIAO, J.; JIANG, J. Analysis of spatiotemporal mobility of shared-bike usage during COVID-19 pandemic in Beijing. **Transactions in GIS**, [S.l.], v. 00, p. 1-22, 2021. Disponível em: <https://doi.org/10.1111/tgis.12784>. Acesso em: 22 fev. 2025.

CHANDY, R.; HASSAN, M.; MUKHERJI, P. Big Data for Good: Insights from Emerging Markets. **Journal of Product Innovation Management**, [S.l.], v. 34, n. 5, p. 703-713. Disponível em: <https://onlinelibrary.wiley.com/doi/abs/10.1111/jpim.12406>. Acesso em: 13 mar. 2025.

D4GX INDIA. **Data for Good Exchange 2024 India**. 2024. Disponível em: <https://www.d4gxindia.com/>. Acesso em: 13 mar. 2025.

DATA KIND. **Our Story**. 2025. Disponível em: <https://www.datakind.org/about-us/our-story/>. Aceso em: 13 mar. 2025.

DATA SCIENCE ACADEMY. **Big Data Fundamentos 3.0**. 2023. Disponível em: <https://www.datascienceacademy.com.br>. Acesso em: 10 ago. 2024.

DAVENPORT, T. H. **Big data no trabalho**: derrubando mitos e descobrindo oportunidades. Rio de Janeiro: Alta Books, 2017.

DAVENPORT, T. H. How strategists use “big data” to support internal business decisions, discovery and production. **Strategic Leadership**, Leeds, v. 42, n. 4, p. 45-50, 2014. Disponível em: <https://doi.org/10.1108/SL-05-2014-0034>. Acesso em: 22 ago. 2025.

DRIVENDATA. **Data science competitions to build a better world**. 2024. Disponível em: <https://www.drivendata.org/>. Acesso em: 13 mar. 2025.

GUTIERREZ, M.; BRYANT, J. The Fading Gloss of Data Science: Towards an Agenda that Faces the Challenges of Big Data for Development and Humanitarian Action. **Development**, [S.l.], v. 65, p. 80-93, 2022. Disponível em: <https://doi.org/10.1057/s41301-022-00327-2>. Acesso em: 22 fev. 2025.

HÖPER, L.; SCHULTE, C. The data awareness framework as part of data literacies in K-12 education. **Information and Learning Sciences**, [S.l.], v. 125, n. 7/8, p. 491-512, 2023. Disponível em: <https://doi.org/10.1108/ILS-06-2023-0075>. Acesso em: 22 fev. 2025.

INSTITUTO DE CIDADANIA EMPRESARIAL - ICE. **Tecnologias e uso de dados para transformação social**: Movimento Data for Good começa a se estruturar no Brasil. 2018. Disponível em: <https://ice.org.br/2018/10/17/tecnologias-e-uso-de-dados-para-transformacao-social-movimento-data-for-good-comeca-a-se-estruturar-no-brasil/>. Acesso em: 13 mar. 2025.

JIA, S.; KIM, S. H.; NGHIEM, S. V.; DOHERTY, P.; KAFATOS, M. C. Patterns of population displacement during mega-fires in California detected using Facebook Disaster Maps. **Environmental Research Letters**, [S.l.], v. 15, n. 7. 2020. Disponível: <https://doi.org/10.1088/1748-9326/ab8847>. Acesso em: 07 maio 2026.

KONTOKOSTA, C. E.; MALIK, A. The resilience to emergencies and disasters index: applying big data to benchmark and validate neighborhood resilience capacity. **Sustainable Cities and Society**, v. 36, p. 272–285. Disponível em: <http://dx.doi.org/10.1016/j.scs.2017.10.025>. Acesso em: 09 ago. 2026.

LANCY, D. 3D Data Management: Controlling Data Volume, Velocity, and Variety. **Application Delivery Strategies**. Stamford: Meta Group. 2001. Disponível em: <https://studylib.net/doc/8647594/3d-data-management--controlling-data-volume--velocity--a>. Acesso em: 13 mar. 2025.

LUCCHINI, L.; CENTELLEGER, S.; PAPPALARDO, L.; GALLOTTI, R.; PRIVETERA, F.; LEPRI, B.; NADAI, M. Living in a pandemic: changes in mobility routines, social activity and adherence to COVID-19 protective measures. **Scientific Reports**, [S.l.], v. 11, n. 24452, p. 1-14, 2021. Disponível em: <https://doi.org/10.1038/s41598-021-04139-1>. Acesso em: 22 fev. 2025.

MANN, G.; SAHUGUET, A. With Great Data Comes Great Responsibility: How Data Science Can Be Applied to Solve Public Interest Problems Without Losing Its Soul.

Journal of Technology in Human Services, [S.l.], v. 36, 2018. Disponível em: <https://doi.org/10.1080/15228835.2017.1419157>. Acesso em: 13 mar. 2025.

MICHAŁOWICZ, A. Data Sharing for the Common Good – Remarks on the Data Altruism Framework Established under the Data Governance Act. **Yearbook of Antitrust and Regulatory Studies**, [S.l.], v. 17, n. 29, p. 209-234, 2024. Disponível em: <https://doi.org/10.7172/1689-9024.YARS.2024.17.29.7>. Acesso em: 22 fev. 2025.

MOHAN, P. S.; RAMSAWAK, R. Using Crowdsourced Data to Assess the Impact of Shelter-in-Place Orders (SIPO) during COVID-19: Lessons from a Small Island Developing State. **Journal of Health Management**, [S.l.], v. 26, n. 3, p. 555–570, 2024. Disponível em: <https://doi.org/10.1177/09720634241244422>. Acesso em: 22 fev. 2025.

MONTEIRO, M.J., MAIA, P. A Framework for Building *pro-bono* Data for Good Projects. In: KAMP, M., KOPRINSKA, I.; BIBAL, A.; BOUADI, T.; FRÉNAY, B.; GALÁRRAGA, L.; ORAMAS, J.; ADILOVA, L.; KRISHNAMURTHY, Y.; KANG, B.; VIARD, T.; WELKE, P.; RUPCCP, M.; AUNE, E.; GALLICCHIO, C.; SCHIELE, G.; PERNKOPF, F.; BLOTT, M.; FRÖNING, H.; SCHINDLER, G.; GUIDOTTI, R.; MONREALE, A.; RINZIVILLO, S.; BIECEK, P.; NTOUTSI, E.; PECHENIZKIY, M.; ROSENHAHN, B.; BUCKLEY, C.; CIALFI, D.; LANILLOS, P.; RAMSTEAD, M.; VERBELEN, T. FERREIRA, P. M.; ANDRESINI, G.; MALERBA, D.; MEDEIROS, I.; FOURNIER-VIGER, P.; NAWAZ, M. S.; VENTURA, S.; SUN, M.; ZHOU, M.; BITETTA, V.; BORDINO, I.; FERRETTI, A.; GULLO, F.; PONTI, G.; SEVERINI, L.; RIBEIRO, R.; GAMA, J.; GAVALDÀ, R.; COOPER, L.; GHAZALEH, N.; RICHIARDI, J.; ROQUEIRO, D.; MIRANDA, D. S.; SECHIDIS, K.; GRAÇA, G. **Machine Learning and Principles and Practice of Knowledge Discovery in Databases**. Springer: [S.l.], 2021. v. 1525. Disponível em: https://doi.org/10.1007/978-3-030-93733-1_19. Acesso em: 22 fev. 2025.

NGUYEN, H.; NGUYEN, T.; NGUYEN, D. T. A Graph-Based Approach for Population Health Analysis Using Geo-Tagged Tweets. **Multimedia Tools and Applications**, v. 80, p. 7187-7204, 2021. Disponível em: <https://doi.org/10.1007/s11042-020-10034-0>. Acesso em: 22 fev. 2025.

PATIL, D. J; MASON, H. **Data Driven: Creating a Data Culture**. California: O'Reilly Media, 2015.

SANDER, I. Critical datafication literacy – a framework for educating about datafication. **Information and Learning Sciences**, [S.l.], v. 125, n. 3/4, p. 270-292, 2024. Disponível em: <https://doi.org/10.1108/ILS-06-2023-0064>. Acesso em: 22 fev. 2025.

SANTOS, A. P.; HEIDER, K.; GRESSE JUNIOR, S.; RODRIGUEZ LOPEZ, J. M. The uneven burden of COVID-19 in the metropolitan region of São Paulo, Brazil – Risk analysis from a bottom-up perspective. **Applied Geography**, [S.l.], v. 162, 2024. Disponível em: <https://doi.org/10.1016/j.apgeog.2023.103146>. Acesso em: 22 fev. 2025.

SHERBININ, A.; BOWSER, A.; CHUANG, T. R.; COOPER, C.; DANIELSEN, F.; EDMUNDS, R.; ELIAS, P.; FAUSTMAN, E.; HULTQUIST, C.; MONDARDINI, R.; POPESCU, I.; SHONOWO, A.; AIVAKUMAR, K. The Critical Importance of Citizen Science Data. **Frontiers in Climate**, [S.l.], v. 3, p. 1-12, 2021. DOI:

10.3389/fclim.2021.650760. Disponível em:

<https://www.frontiersin.org/articles/10.3389/fclim.2021.650760/full>. Acesso em: 22 fev. 2025.

SILVEIRA, M.; MARCOLIN, C. B.; FREITAS, H. Uso Corporativo do Big Data: Uma Revisão de Literatura. **Revista de Gestão e Projetos**, [S.l.], v. 6, n. 3. set/dez. 2015.

Disponível em: <https://periodicos.uninove.br/gep/article/view/9627/4372>. Acesso em: 13 mar. 2025.

SOCIAL GOOD BRASIL – SGB. **Quem somos**. 2025. Disponível em:

<https://socialgoodbrasil.org.br/quem-somos/>. Acesso em: 13 mar. 2025.

WANG, G.; ZHAO, J.; VAN-KLEEK, M. G.; SHADBOLT, N. R. 'Don't make assumptions about me!': Understanding children's perception of datafication online.

In: ACM CONFERENCE ON COMPUTER-SUPPORTED COOPERATIVE WORK AND SOCIAL COMPUTING, 22, 2022, Nova Iorque. **Anais eletrônicos [...]** Nova Iorque: Association for Computing Machinery, 2022. p. 1-24. Disponível em:

<https://dl.acm.org/doi/10.1145/3555144>. Acesso em: 22 fev. 2025.

WITAYANGKURN, A.; ARAI, A.; SHIBASAKI, R. Development of Big Data-Analysis Pipeline for Mobile Phone Data with Mobipack and Spatial Enhancement. **ISPRS International Journal of Geo-Information**, [S.l.], v. 11, n. 196, p. 1-21, 2022.

Disponível em: <https://doi.org/10.3390/ijgi11030196>. Acesso em: 22 fev. 2025.

XIA, J.; WANG, J.; LAI, Y. Development Strategy Based on Combination Typologies of Building Carbon Emissions and Urban Vibrancy—A Multi-Sourced Data-Driven Approach in Beijing, China. **Land**, [S.l.], v. 13, n. 1062, p. 1-19, 2024. Disponível em: <https://doi.org/10.3390/land13071062>. Acesso em: 22 fev. 2025.

YOON, A.; COPELAND, A. Understanding social impact of data on local communities. **Aslib Journal of Information Management**, [S.l.], v. 71, n. 4, p. 558-567, 2019. Disponível em: <https://doi.org/10.1108/AJIM-12-2018-0310>. Acesso em: 22 fev. 2025.

Big Data and Data for Good in Information Science: theoretical approximation

Abstract

Introduction: based on the scope of Big Data and its relationship with Information Science,

this work investigates the concepts and applications of Data for Good, an approach to using data for the common good. **Objective:** To understand how the themes already studied in Information Science are embedded in discussions about Data for Good, as well as to understand which paths, based on the Enancib Working Groups, best engage with the identified publications. **Methodology:** A qualitative research with an exploratory and documentary character is developed through a literature review. Thematic analysis is used to identify the relationships found, and four main themes are defined: Cities; Education, Health, and Community. Data from the researched articles are also cross-referenced with information from the abstracts of the working groups of the National Meeting of Research in Information Science (Enancib). **Results:** The retrieved works present research with a diversity of data collection and analysis methods. These studies, even without directly referencing Information Science, explore themes that are transversal and clearly relevant to the area. The four working groups that stand out in the relationship made are WG 05 – Information Policy and Economics; WG 11 – Information & Health; GT 08 – Data, Information and Technology and GT 06 – Information, Education and Work, finding a relationship with the proposed themes. **Conclusions:** it is possible to draw a parallel in their attempts to use Big Data to support the solution of socio-environmental problems, with the identification of risks, benefits and indication of good practices for informed decision-making.

Descriptors: Big Data. Data for Good. Information Science. Literature Review.

Big Data y Datos para el Bien en la Ciencia de la Información:

un enfoque teórico

Resumen

Introducción: basado en el alcance del Big Data y su relación con las Ciencias de la Información, este trabajo investiga los conceptos y aplicaciones de Data for Good, un enfoque de uso de datos para el bien común. **Objetivo:** comprender cómo los temas ya estudiados en Ciencias de la Información se integran en los debates sobre Data for Good, así como identificar qué enfoques, basados en los Grupos de Trabajo de Enancib, permiten una mejor interacción con las publicaciones identificadas en la revisión bibliográfica. **Metodología:** se utiliza el análisis temático para identificar las relaciones encontradas y se definen cuatro temas principales: Ciudades; Educación, Salud y Comunidad. Los datos de los artículos investigados se referencian cruzadamente con la información de los resúmenes de los grupos de trabajo de la Reunión Nacional de Investigación en Ciencia de la Información (Enancib). **Resultados:** os trabajos recuperados presentan investigación con una diversidad de métodos de recopilación y análisis de datos. Estos estudios, incluso sin referenciar directamente a la Ciencia de la Información, exploran temas que son transversales y claramente relevantes para el área. Los cuatro grupos de trabajo que destacan en la relación hecha con los artículos son GT 05; GT 11; GT 08 y GT 06, encontrando una relación con los temas propuestos. **Conclusiones:** es posible establecer un paralelismo entre sus intentos de utilizar Big Data para apoyar la solución de problemas socioambientales, con la identificación de riesgos, beneficios e indicación de buenas prácticas para la toma de decisiones informadas.

Descriptores: Big Data. Datos para el bien. Ciencia de la Información. Revisión de Literatura.