

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
FACULDADE DE BIBLIOTECONOMIA E COMUNICAÇÃO
DEPARTAMENTO DE CIÊNCIAS DA INFORMAÇÃO
PÓS-GRADUAÇÃO EM CIÊNCIA DA INFORMAÇÃO

LAURA REGINA DO CANTO LEAL

OS REPOSITÓRIOS DE DADOS DE PESQUISA EM PROJETOS DE CIÊNCIA
CIDADÃ: participação do cidadão não-cientista em artigos acadêmicos



PORTO ALEGRE

2023

LAURA REGINA DO CANTO LEAL

**OS REPOSITÓRIOS DE DADOS DE PESQUISA EM PROJETOS DE CIÊNCIA
CIDADÃ: participação do cidadão não-cientista em artigos acadêmicos**

Dissertação apresentada ao Programa de Pós-Graduação em Ciência da Informação da Universidade Federal do Rio Grande do Sul como requisito parcial à obtenção do título de Mestra em Ciência da Informação.

Área de Concentração: Informação e Ciência

Orientador: Prof. Dr. René Faustino Gabriel Junior

Porto Alegre

2023

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL

Reitor: Prof. Dr. Carlos André Bulhões Mendes

Vice-Reitora: Prof.^a Dr.^a. Patrícia Helena Lucas Pranke**FACULDADE DE BIBLIOTECONOMIA E COMUNICAÇÃO**Diretora: Prof.^a Dr.^a. Ana Maria de Moura

Vice-diretora: Vera Regina Schmitz

PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA INFORMAÇÃO

Coordenador: Prof. Dr. Thiago Henrique Bragato Barros

Coordenador Substituto: Prof. Dr. Moisés Rochemback

CIP - Catalogação na Publicação

Leal, Laura Regina do Canto
OS REPOSITÓRIOS DE DADOS DE PESQUISA EM PROJETOS DE
CIÊNCIA CIDADÃ: participação do cidadão não-cientista
em artigos acadêmicos / Laura Regina do Canto Leal. --
2023.

128 f.

Orientador: René Faustino Gabriel Junior.

Dissertação (Mestrado) -- Universidade Federal do
Rio Grande do Sul, Faculdade de Biblioteconomia e
Comunicação, Programa de Pós-Graduação em Ciência da
Informação, Porto Alegre, BR-RS, 2023.

1. Ciência Cidadã. 2. Crowdsourcing. 3. Dados de
pesquisa. 4. Repositório de dados abertos. I. Gabriel
Junior, René Faustino, orient. II. Título.

Elaborada pelo Sistema de Geração Automática de Ficha Catalográfica da UFRGS
com os dados fornecidos pela Autora.

PPGCIN – UFRGS

Rua Ramiro Barcelos, 2705, Prédio 22201 CEP: 90035-007 Porto Alegre – RS

Telefone: (51) 3308-5067

E-mail: ppgcin@ufrgs.br

LAURA REGINA DO CANTO LEAL

OS REPOSITÓRIOS DE DADOS DE PESQUISA EM PROJETOS DE CIÊNCIA

CIDADÃ: participação do cidadão não-cientista em artigos acadêmicos

Dissertação apresentada ao Programa de Pós-Graduação em Ciência da Informação da Universidade Federal do Rio Grande do Sul como requisito parcial à obtenção do título de mestre em Ciência da Informação.

Aprovada em Porto Alegre, 25/08/2023.

BANCA EXAMINADORA

Prof^a. Dr^a. Paula Carina de Araújo - Docente do Departamento de Ciência e Gestão da Informação e do Programa de Pós-graduação em Gestão da Informação da Universidade Federal do Paraná (UFPR).

Prof. Dr. Fabiano Couto Corrêa da Silva – Programa de Pós-Graduação em Ciência da Informação da Universidade Federal do Rio Grande do Sul (PPGCIN/UFRGS)

Prof. Dr. René Faustino Gabriel Junior – Programa de Pós-Graduação em Ciência da Informação da Universidade Federal do Rio Grande do Sul (PPGCIN/UFRGS)

Prof^a. Dr^a. Caterina Marta Groposo Pavão – Programa de Pós-Graduação em Ciência da Informação da Universidade Federal do Rio Grande do Sul (PPGCIN/UFRGS)

Prof. Dr. Thiago Henrique Bragato Barros – Programa de Pós-Graduação em Ciência da Informação da Universidade Federal do Rio Grande do Sul (PPGCIN/UFRGS)

AGRADECIMENTOS

É com imensa gratidão e alegria que me sinto compelida a expressar meus sinceros agradecimentos a todas as pessoas que estiveram ao meu lado, oferecendo apoio, incentivo e sabedoria ao longo dessa caminhada desafiadora.

Em primeiro lugar, quero dedicar um agradecimento todo especial aos meus filhos, Carolina e Lucas. Vocês foram a força motriz que me impulsionou a superar obstáculos, manter o foco e perseverar diante das adversidades. Amo vocês!

Ao meu orientador, René Faustino Gabriel Junior, expresso minha profunda gratidão pela orientação dedicada, paciência inesgotável e compromisso com a excelência acadêmica. Suas orientações e conselhos foram fundamentais para o desenvolvimento desta dissertação, e aprendi muito com sua experiência e conhecimento.

Agradeço também à minha banca de qualificação, composta por professores e pesquisadores cujas contribuições foram inestimáveis para o aprimoramento do meu trabalho. Suas sugestões valiosas e análises criteriosas enriqueceram esta dissertação, possibilitando-me aprimorar minha pesquisa de forma significativa.

Não poderia deixar de agradecer à Universidade Federal do Rio Grande do Sul (UFRGS), representada pelos seus professores e funcionários.

Que este momento de gratidão e celebração seja apenas o início de novas conquistas e aprendizados em nossas vidas.

Muito obrigada a todos!

RESUMO

A colaboração entre cientistas e voluntários não especializados permite que cidadãos comuns se envolvam ativamente na pesquisa científica. A Ciência Cidadã, com a participação de cidadãos na coleta de dados, tem se destacado como uma abordagem inclusiva e colaborativa. O compartilhamento de dados abertos em projetos de Ciência Cidadã é crucial para desenvolver o conhecimento e embasar decisões com evidências, embora ainda existam lacunas na compreensão sobre como esse compartilhamento ocorre em artigos científicos e se os resultados alcançam ampla divulgação na sociedade. Este estudo tem como foco uma pesquisa bibliográfica, visando a compreensão do processo de compartilhamento de dados abertos em repositórios relacionados à Ciência Cidadã e o seu impacto na sociedade. Tem-se como objetivo principal em investigar o processo de compartilhamento de dados abertos em repositório de projetos com vertentes na Ciência Cidadã, analisando como esse compartilhamento é abordado nos artigos científicos. Caracterizado como estudo de natureza básica e de caráter descritivo. Para descrever e interpretar os dados coletados sobre Ciência Cidadã e os seus dados abertos, o estudo utiliza metodologia e técnica de abordagem qualitativa e para atender os objetivos propostos utiliza-se da revisão bibliográfica. Dos resultados da pesquisa, entende-se que a literatura desempenha um papel importante ao disponibilizar os estudos em Ciência Cidadã, permitindo que a sociedade se aproprie do conhecimento gerado, por meio do envolvimento ativo dos cidadãos em projetos relevantes, os quais contribuem com suas habilidades e experiências nas pesquisas científicas. Na análise da relação entre a Ciência Cidadã e a participação do cidadão em artigos acadêmicos, a partir do armazenamento de dados de pesquisa, evidenciou a importância da participação ativa do cientista cidadão em projetos de Ciência Cidadã. Os repositórios desempenham um papel crucial na promoção da transparência, colaboração e reutilização de conjuntos de dados, fomentando avanços em diversas áreas do conhecimento. Apesar dos resultados tímidos na identificação do armazenamento de projetos em repositórios, em que apenas sete permitiram a análise dos dados de pesquisa, ainda é evidente o impacto positivo quando esses recursos são bem utilizados.

Palavras-chave: Ciência Cidadã; *Crowdsourcing*; Dados de pesquisa; Repositório de dados abertos.

ABSTRACT

Collaboration between scientists and non-expert volunteers allows ordinary citizens to become actively involved in scientific research. Citizen Science, with the participation of citizens in data collection, has stood out as an inclusive and collaborative approach. Sharing open data in Citizen Science projects is crucial for developing knowledge and supporting decisions with evidence, although there are still gaps in understanding how this sharing occurs in scientific articles and whether the results achieve wide dissemination in society. This study focuses on bibliographical research, aiming to understand the process of sharing open data in repositories related to Citizen Science and its impact on society. The main objective is to investigate the process of sharing open data in a repository of projects with aspects of Citizen Science, analyzing how this sharing is approached in scientific articles. Characterized as a study of a basic and descriptive nature. To describe and interpret the data collected on Citizen Science and its open data, the study uses a qualitative approach methodology and technique and to meet the proposed objectives, it uses a bibliographic review. From the research results, it is understood that literature plays an important role in making studies in Citizen Science available, allowing society to appropriate the knowledge generated, through the active involvement of citizens in relevant projects, which contribute to their skills and experiences in scientific research. In the analysis of the relationship between Citizen Science and citizen participation in academic articles, based on the storage of research data, the importance of the active participation of citizen scientists in Citizen Science projects was highlighted. Repositories play a crucial role in promoting transparency, collaboration and reuse of datasets, fostering advances in various areas of knowledge. Despite the timid results in identifying the storage of projects in repositories, in which only seven allowed the analysis of research data, the positive impact when these resources are used well is still evident.

Keywords: Citizen Science; Crowdsourcing; Search data; Open data repository.

LISTAS DE FIGURAS

| | |
|---|----|
| Figura 1 - Valores fundamentais e princípios orientadores da Ciência Aberta..... | 21 |
| Figura 2 - Representação dos 10 princípios da Ciência Cidadã, representados em diferentes idiomas. | 28 |
| Figura 3 - Nuvens de palavras para vários termos dos participantes da Ciência Cidadã | 31 |
| Figura 4 - Modelo de engajamento do cientista cidadão na Ciência Cidadã | 34 |
| Figura 5 – Página inicial do Projeto Galaxy Zoo..... | 37 |
| Figura 6 – Página inicial do Projeto de transcrição dos documentos de John Torrey | 39 |
| Figura 7 - Projetos de Ciência Cidadã da NASA | 41 |
| Figura 8 - Representação de como as empresas podem se beneficiar do engajamento do público | 43 |
| Figura 9 – Página do iNaturalist no Cadastro de Repositórios de Dados de Pesquisa..... | 46 |
| Figura 10 – Página do Repositório Arca Dados | 47 |
| Figura 11 - <i>Research Data Repository</i> (RDR) da University College London..... | 74 |
| Figura 12 - Visualização dos projetos em grade e por primeira data <i>online</i> | 75 |
| Figura 13 - Dados do Projeto <i>Kuyruktobe monument, Otrar Oasis, aerial photography of 2016 and digital derivatives</i> , armazenados no <i>Research Data Repository</i> | 76 |
| Figura 14 - Imagem do Repositório do Museum für Naturkunde Berlin, já na página do artigo onde estão os resultados dos dados coletados..... | 77 |
| Figura 15 - Ilha do Barro Colorado, primeiro observatório do ForestGeo | 80 |
| Figura 16 - <i>Smithsonian Research Online</i> - BCI 50 ha Plot 1982-2005 Census Data..... | 81 |
| Figura 17 - <i>Salish Sea Marine Survival Project</i> | 83 |
| Figura 18 - Oceans 3.0 Data Portal..... | 84 |
| Figura 19 - Visualização da localização do Oceano Atlântico, no <i>Cape Bonavista Buoy</i> e informação do sensor de temperatura e umidade do ar. | 85 |
| Figura 20 - <i>Site Snapshot Wisconsin</i> | 86 |
| Figura 21 - Fluxo de dados por meio do Projeto <i>Snapshot Wisconsin</i> | 87 |
| Figura 22 - Coleções instantâneas de Wisconsin e Guia de campo das espécies. | 88 |
| Figura 23 - Coleção instantânea dos favoritos e os metadados | 89 |
| Figura 24 - Descrição da classificação das espécies com os exemplo de metadados | 89 |
| Figura 25 - Painel de dados da <i>Snapshot Wisconsin</i> | 90 |
| Figura 26 - Projeto Global Ozone (GO3)..... | 91 |
| Figura 27 - Projeto GO3, arquivado no <i>Internet Archive WayBack Machine</i> | 92 |
| Figura 28 - Representação do mapa e do resumo de dados do Projeto GO3 | 93 |
| Figura 29 - Imagens recortadas do vídeo do <i>YouTube</i> sobre os projetos | 94 |
| Figura 30 - Dados do Projeto AQTreks (AirQualityTreks), apresentado na <i>SciStarter</i> | 95 |
| Figura 31 - Página do Centro de Estudos Transculturais de Heidelberg (HCTS) onde está disponibilizado o Repositório <i>Open Dachs</i> | 96 |

| | |
|---|-----|
| Figura 32 - Dados gerados do artigo de Arnold e outros (2020), no repositório <i>GiHub</i> , relacionado ao primeiro Estudo de caso. | 97 |
| Figura 33 - Dados disponibilizados no Repositório <i>GiHub</i> | 98 |
| Figura 34 – Spipoll e as Últimas coleções..... | 101 |

LISTA DE GRÁFICOS

| | |
|--|----|
| Gráfico 1 - Representação dos dados selecionados após primeira etapa da Revisão Bibliográfica..... | 63 |
| Gráfico 2 - Número de publicações no período de 2012 a 2021 | 65 |

LISTA DE QUADROS

| | |
|--|----|
| Quadro 1 - Princípios de Dados Governamentais Abertos..... | 50 |
| Quadro 2 - Estratégia de Busca para Revisão Bibliográfica..... | 57 |
| Quadro 3 - Modelo base dos procedimentos para Análise Sistemática de Literatura na ferramenta <i>StArt</i> | 58 |
| Quadro 4 - Critérios de inclusão e exclusão utilizados no modelo base da Revisão Sistemática de Literatura..... | 58 |
| Quadro 5: Base de dados Scopus: termos e campos de buscas utilizados e documentos recuperados (2012 até dez. 2021) com frequência superior a duas incidências sobre Ciência Cidadã | 63 |
| Quadro 6 - Base de dados Scopus: termos e campos de buscas utilizados e documentos recuperados (2012 até 2021) com frequência superior a duas incidências sobre Dados de pesquisa | 64 |

LISTAS DE TABELAS

| | |
|---|-----|
| Tabela 1 - Lista de autores mais produtivos..... | 66 |
| Tabela 2 - Dispositivos que armazenam dados de pesquisa de projetos em Ciência Cidadã, mencionados nos artigos analisados. | 69 |
| Tabela 3 - Identificação dos Títulos por meio do ID disponibilizado pelo <i>software StArt</i> | 120 |
| Tabela 4 - Dados informativos dos artigos selecionados para estudo sobre projetos em Ciência Cidadã | 123 |

LISTA DE SIGLAS

| | |
|----------------|--|
| AA | Acesso Aberto |
| BOAI | <i>Budapest Open Access Initiative</i> |
| CA | Ciência Aberta |
| CC | Ciência Cidadã |
| ECSA | <i>European Citizen Science Association</i> |
| EUA | Estados Unidos da América |
| FOSTER | <i>Facilitate Open Science Training for European Research</i> |
| FIOCRUZ | Fundação Oswaldo Cruz |
| IAS | Espécies exóticas invasoras |
| IBICT | Instituto Brasileiro de Informação em Ciência e Tecnologia |
| IoT | Internet das coisas |
| NASA | <i>National Aeronautics and Space Administration</i> |
| ONGs | Organizações não governamentais |
| OD | Dados abertos |
| ODS | Objetivos de Desenvolvimento Sustentável |
| OKF | <i>Open Knowledge International</i> |
| ONU | Organização das Nações Unidas |
| ONC | <i>Ocean Networks Canada</i> |
| RDR | <i>Research Data Repository</i> |
| StArt | <i>State of the Art through Systematic Review</i> |
| UCL | <i>University College Londres</i> |
| UNESCO | Organização das Nações Unidas para Educação, Ciência e Cultura |
| USA | <i>United State of American</i> |
| TICs | Tecnologias de Informação e Comunicação |
| WEB | Variação de <i>World Wide Web</i> |

SUMÁRIO

| | |
|---|------------|
| 1 INTRODUÇÃO | 15 |
| 1.1 IDENTIFICAÇÃO DO PROBLEMA | 17 |
| 1.2 OBJETIVOS | 17 |
| 1.2.1 Objetivo geral..... | 17 |
| 1.2.2 Objetivos específicos | 18 |
| 1.3 JUSTIFICATIVA | 18 |
| 2 REFERENCIAL TEÓRICO | 20 |
| 2.1 CIÊNCIA ABERTA (<i>OPEN SCIENCE</i>)..... | 20 |
| 2.2 CIÊNCIA CIDADÃ (<i>CITIZEN SCIENCE</i>): HISTÓRIA E CARACTERÍSTICAS . | 22 |
| 2.2.1 Cientistas Cidadãos | 30 |
| 2.2.2 Projetos de Ciência Cidadã | 35 |
| 2.2.3 Crowdsourcing | 42 |
| 2.3 REPOSITÓRIOS DE DADOS DE PESQUISA | 44 |
| 2.4 DADOS ABERTOS DE PESQUISA EM CIÊNCIA CIDADÃ..... | 48 |
| 3 METODOLOGIA | 55 |
| 4 RESULTADOS E ANÁLISES | 62 |
| 4.1 ANÁLISE DOS DADOS DE PESQUISA NO CONTEXTO DA CIÊNCIA CIDADÃ NA BASE DE DADOS <i>SCOPUS</i> | 62 |
| 4.2 IDENTIFICAÇÃO DE ARMAZENAMENTO DE DADOS DE PESQUISA EM PROJETOS EM CIÊNCIA CIDADÃ EM REPOSITÓRIOS | 67 |
| 4.2.1 Análises dos compartilhamentos de dados em repositórios | 72 |
| 4.2.1.1 RDR - Research Data Repository (ID 13) | 73 |
| 4.2.1.2 MFN Data Repository (ID 16)..... | 77 |
| 4.2.1.3 Smithsonian Research Online e Plataforma Dryad (ID.17)..... | 79 |
| 4.2.1.4 Oceans 3.0 Data Portal (ID. 67)..... | 82 |
| 4.2.1.5 Zooniverse e a Snapshot Wisconsi (ID.70) | 85 |
| 4.2.1.6 <i>Internet Archive WayBack Machine</i> (ID.69) | 91 |
| 4.2.1.7 Repositório de citações <i>Open Dachs</i> (ID 21385) | 95 |
| 4.3 INFORMAÇÕES RELEVANTES AO ESTUDO | 98 |
| 5 CONSIDERAÇÕES FINAIS | 103 |
| REFERÊNCIAS | 109 |
| APÊNDICE - TABELA A | 120 |
| APÊNDICE - TABELA B | 123 |

1 INTRODUÇÃO

A participação ativa de cidadãos comuns no processo científico envolve a colaboração entre cientistas e voluntários não especializados, permitindo que as pessoas fora do campo acadêmico também se envolvam com a ciência. Muitos dados são coletados e utilizados em pesquisas científicas, gerando trabalhos acadêmicos e em alguns casos gerando um retorno social para a comunidade.

A ideia de compartilhar dados se concretizou efetivamente nas últimas três décadas, principalmente com o movimento do Acesso Aberto¹ (AA) à publicações científicas, movimento este que surgiu em 1991, na reunião de *Budapest Open Access Initiative* (BOAI)², que objetivava a abertura e acesso de forma gratuita e irrestrita da produção literária científica³, apresentando como vantagens a visibilidade e o fator de impacto nas pesquisas, ampliando, assim, o número de citações.

A abertura de dados de pesquisa amplia a visibilidade, transparência e confiabilidade, facilitando a colaboração e o aumento no ritmo das descobertas, além de promover um acesso maior e igualitário ao conhecimento científico e do próprio processo de pesquisa.

Nos últimos anos, mais especificamente a partir da Década de 2000, o movimento para disponibilizar informação científica ampliou-se com o movimento da Ciência Aberta (CA)⁴ que apresentou uma nova perspectiva nas questões e demandas da sociedade na construção do conhecimento, não só na academia, mas na sociedade em geral.

Entende-se que, com a Ciência Aberta existe um compromisso social da academia, mais especificamente com a Ciência Cidadã (CC), em socializar os conhecimentos produzidos pela ciência com a participação de pessoas comuns promovendo vantagens para ambos. A Ciência Cidadã, por meio da consciência social e global, pode e deve se mover para uma contribuição plena, onde cidadãos do mundo

¹ Acesso aberto ou acesso livre significa a disponibilização online e sem limitações dos resultados de investigação científica.

² <https://www.budapestopenaccessinitiative.org/>

³ Literatura científica, é aquela revisada por pares, isto é, disponibilizada sem restrições na internet, onde qualquer usuário pode ler, baixar, copiar, distribuir, imprimir, pesquisar ou referenciar, em seus textos, com credibilidade (BOAI, 2001, *online*).

⁴ Ciência aberta é a atividade científica praticada de modo aberto, colaborativo e transparente, em todos os domínios do conhecimento, desde as ciências fundamentais até as ciências sociais e humanidades: <https://portal.fiocruz.br/ciencia-aberta>.

todo possam contribuir mais do que nunca para a ciência e para a democratização do conhecimento.

Uma das formas de promover o retorno social é pelo compartilhamento dos dados coletados pelos cidadãos com a sociedade em geral. Portanto, vale salientar que a abertura de dados em projetos de Ciência Cidadã se faz necessária, pois existe um interesse comum entre sistema governamental e privado nessas iniciativas, seja como apoiadores ou/e financiadores.

Evidenciar a abertura de dados em projetos de Ciência Cidadã é permitir que a sociedade entenda a importância desses projetos, bem como se sinta estimulada a contribuir e participar em projetos de interesse próprio ou da sociedade, ou mesmo na criação de novos conhecimentos, com base nas informações geradas na abertura dos dados ou reuso dos mesmos.

A Ciência Cidadã, apesar de na prática não ser nova, mas na academia ainda não está plenamente consolidada, entretanto vem ganhando confiabilidade e visibilidade, à medida que a sociedade toma conhecimento das ações colaborativas que estão surgindo e que podem fazer parte desse processo (GUANAES, 2020). Portanto o aumento de sua produção ganha espaço devido a um envolvimento maior de voluntários nos projetos de Ciência Cidadã, além do envolvimento de instituições governamentais e privadas, bem como no compartilhamento de seus resultados de pesquisa.

Para a Leal (2011, p. 25), Ciência Cidadã [...] é o envolvimento de cidadãos que não são cientistas profissionais, na colaboração de investigações de caráter científico”. Em uma definição mais formal Ciência Cidadã é a coleta e análise metódica de informações, o avanço tecnológico, a validação de características naturais e a ocorrência dessas atividades por pesquisadores em grande parte de forma voluntária (OpenScientist, 2011).

Este projeto em sua morfologia está estruturado em cinco capítulos, um de introdução explicando sobre a problemática, o objetivo geral, os específicos e as justificativas do estudo.

No segundo capítulo, apresenta o referencial teórico sobre a Ciência Aberta, Ciência Cidadã e seus desdobramentos, Repositórios de Dados de pesquisa e Dados abertos de pesquisa.

No terceiro capítulo, descreve-se a metodologia quanto a sua abordagem, finalidade do estudos, estratégias e técnicas de análises, visando atender os objetivos

do estudo.

No quarto capítulo, são apresentadas as análises estruturadas de acordo com os objetivos específicos da pesquisa, abrangendo análises dos artigos recuperados, dos repositórios de dados de pesquisa provenientes dos artigos e as tipologias dos dados que envolvem a ciência cidadã.

No último capítulo, são apresentadas as percepções do pesquisador sobre os resultados obtidos na pesquisa, que incluem a identificação e análise de sete repositórios mencionados nos artigos, bem como sugestões para futuros estudos e continuidade da pesquisa.

1.1 IDENTIFICAÇÃO DO PROBLEMA

A Ciência Cidadã tem ganhado destaque como uma abordagem colaborativa e inclusiva para a pesquisa científica, envolvendo a participação ativa de cidadãos comuns no processo de coleta de dados. Nesse contexto, o compartilhamento de dados abertos em repositórios de projetos de Ciência Cidadã é fundamental para o avanço do conhecimento e para a tomada de decisões baseadas em evidências. No entanto, ainda há lacunas no entendimento sobre como o processo de compartilhamento está sendo explicitado nos artigos científicos, e se realmente os resultados estão sendo compartilhados com a sociedade como um todo.

Neste contexto, o problema de pesquisa concentra-se em compreender como está acontecendo o processo de compartilhamento de dados abertos procedente dos artigos científicos, em repositórios de dados (plataformas) que têm vertentes na Ciência Cidadã?

1.2 OBJETIVOS

Para melhor compreensão da temática e responder ao problema de pesquisa, foram estruturados os objetivos desta pesquisa, como descritos abaixo.

1.2.1 Objetivo geral

O objetivo geral desta pesquisa é investigar o processo de compartilhamento de dados abertos em repositórios de projetos com vertentes na Ciência Cidadã, buscando compreender como esse compartilhamento está sendo apresentado nos artigos científicos.

1.2.2 Objetivos específicos

- a) Identificar na literatura trabalhos que envolvem Ciência Cidadã e compartilhamento de dados.
- b) Analisar os repositórios de dados de pesquisa identificados nos textos selecionados na literatura.
- c) Verificar como estão armazenados os dados de pesquisa oriundos de projetos de Ciência Cidadã, nos repositórios.
- d) Identificar as tipologias de dados disponibilizados pelos repositórios de dados que os disponibilizam.

1.3 JUSTIFICATIVA

As motivações da escolha do tema podem ser justificadas em três abordagens, uma pessoal, uma profissional e outra acadêmica. A motivação pessoal, busca mitigar a curiosidade acadêmica de reconhecer projetos de Ciência Cidadã que compartilhem dados que possam ser utilizados em projetos sociais e ou educacionais. Alinhando-se à vertente profissional, pois como professora do estado do Rio Grande do Sul, prospectar o uso de dados em projetos pedagógicos com os alunos.

Da vertente acadêmica, oriunda da continuidade dos estudos abordados no trabalho de conclusão da graduação em Biblioteconomia, do qual observou-se uma carência de literatura científica sobre Ciência Cidadã, principalmente envolvendo as áreas de informação. Desta forma, ampliar os estudos sobre o tema, visa compreender melhor os modelos existentes e gerar novos conhecimentos para a sociedade.

Neste contexto, busca-se reconhecer as práticas da Ciência Cidadã e compreender como seus dados abertos são disponibilizados, bem como a importância e os desafios de suas contribuições para a ciência. Essas indagações levam a investigar práticas de Ciência Cidadã em outros países, compreendendo os possíveis desafios apresentados em seus projetos, bem como as perspectivas de abertura de seus dados de pesquisa.

Justifica-se a compreensão do processo da Ciência Cidadã, uma vez que esse modelo visa a interação e comunicação do meio acadêmico de forma colaborativa e participativa com a sociedade em geral, ou seja, com os não-cientistas, que passam a atuar de maneira ativa dentro das perspectivas da Ciência Cidadã.

O tema por si só se justifica e se sustenta, por ser atual, amplo e relevante para as pesquisas científicas, além do que estudos sobre a Ciência Cidadã e dados abertos serem um desafio e de temática ainda em construção na produção literária científica, ademais, ainda existe a necessidade de remoção dos entraves ao acesso às informações/dados científicos.

2 REFERENCIAL TEÓRICO

Essa seção apresenta a fundamentação teórica das temáticas abordadas neste estudo em quatro tópicos: Ciência Aberta, Ciência Cidadã, repositório de dados de pesquisa e as tipologias dos dados abertos de pesquisa. O referencial teórico é constituído para uma melhor compreensão da terminologia dos conceitos abordados, dos processos culturais e metodológicos da abertura de dados envolvendo estudos e projetos com a participação de cidadãos, bem como para destacar sua importância tanto para a ciência quanto para a sociedade.

2.1 CIÊNCIA ABERTA (*OPEN SCIENCE*)

A Ciência Aberta tem por finalidade tornar a pesquisa científica mais transparente, acessível, colaborativa e reprodutível. Busca, também, remover barreiras tradicionais que limitam o acesso ao conhecimento científico, promovendo a disponibilidade aberta de dados, metodologias e resultados de pesquisa para todos (CHAN; OKUNE; SAMBULI, 2015).

O documento Recomendação da UNESCO sobre Ciência Aberta, define Ciência Aberta como:

[...] um construto inclusivo que combina vários movimentos e práticas que têm o objetivo de disponibilizar abertamente conhecimento científico multilíngue, torná-lo acessível e reutilizável para todos, aumentar as colaborações científicas e o compartilhamento de informações para o benefício da ciência e da sociedade, e abrir os processos de criação, avaliação e comunicação do conhecimento científico a atores da sociedade, além da comunidade científica tradicional. Abrange todas as disciplinas científicas e todos os aspectos das práticas acadêmicas, incluindo ciências básicas e aplicadas, ciências naturais, sociais e humanas, e se baseia nos seguintes pilares-chave: conhecimento científico aberto, infraestrutura científica aberta, comunicação científica, envolvimento aberto dos atores sociais e diálogo aberto com outros sistemas de conhecimento (UNESCO, 2022).

De acordo com a Unesco (2022), os valores fundamentais e princípios orientadores da Ciência Aberta apresentam implicações como questões jurídicas, éticas, epistemológicas, econômicas, políticas, sociais, tecnológicas, abertura da ciência à sociedade, assim como o alargamento dos princípios de abertura de todo o ciclo da pesquisa científica, descritos na Figura 1.

Figura 1 - Valores fundamentais e princípios orientadores da Ciência Aberta



Fonte: Unesco (2022)

Para a FIOCRUZ (2018, *online*), entende-se como Ciência Aberta a “[...] atividade científica praticada de modo aberto, colaborativo e transparente, em todos os domínios do conhecimento, desde as ciências fundamentais até as ciências sociais e humanidades [...]”.

Outra definição é a apresentada no Portal Foster⁵, do projeto *Facilite Open Science Training for European Research Plus* (FOSTER PLUS), que expõe o conceito de Ciência Aberta como sendo:

[...] a prática da ciência de tal forma que outros podem colaborar e contribuir, na qual os dados de pesquisa, as notas de laboratório e outros processos de pesquisa estão disponíveis livremente, em condições que permitem a reutilização, redistribuição e reprodução da pesquisa e dos dados e métodos subjacentes (FOSTER, 2014, *online*).

A definição fornecida por Albagli e outros autores (2014) destaca que a Ciência Aberta é um termo guarda-chuva que vai além do simples acesso aberto a publicações e resultados de pesquisa. Enfatizam a importância de compartilhar não apenas os resultados, mas também todo o processo da pesquisa, incluindo metodologias, dados brutos, códigos e materiais, tornando-o transparente e reprodutível. Os autores complementam a definição do Portal Foster, ao valorizar a participação direta de não

⁵ Portal FOSTER é uma plataforma de *e-learning* que reúne os melhores recursos de treinamento dirigidos àqueles que precisam saber mais sobre o Open Science: <https://www.fosteropenscience.eu/>.

cientistas no fazer da ciência, permitindo que leigos e amadores contribuam ativamente em projetos científicos, enriquecendo as pesquisas e promovendo uma maior diversidade de perspectivas.

Outro aspecto destacado por Albagli e outros autores (2014) é o uso de tecnologia digital e ferramentas colaborativas, que possibilitam a partilha, colaboração e reutilização das informações em tempo real, aumentando a visibilidade da pesquisa e proporcionando novas oportunidades de parcerias e colaborações entre pesquisadores e a comunidade em geral. Com uma definição mais abrangente de Ciência Aberta, os autores evidenciam a mudança de paradigma na prática científica, onde a colaboração, a transparência e o envolvimento da sociedade são valorizados, buscando uma ciência mais inclusiva, responsável e relevante para a sociedade.

2.2 CIÊNCIA CIDADÃ (*CITIZEN SCIENCE*): HISTÓRIA E CARACTERÍSTICAS

O debate e iniciativas sobre estudos e projetos em Ciência Cidadã vem crescendo nas duas últimas décadas, mesmo que definir sua origem e conceituá-la pareça uma tarefa simples, mas, na verdade, tem se tornado um desafio e se pode verificar essa demanda na própria literatura, onde há uma propagação de sua origem, tipologia e definições (MORESI *et al.*, 2017).

É possível identificar na produção literária, que os primeiros relatos sobre o “fazer” Ciência Cidadã não são novidades. Ao longo da história se pode identificar diferentes tipos de projetos e abordagens, que consideramos na atualidade como sendo de Ciência Cidadã, sem o uso de uma definição científica.

Segundo Ryan e outros autores (2018) algumas formas de fazer Ciência Cidadã já foram utilizadas há aproximadamente 3.500 anos. Citam como exemplo os registros de observações de surtos de gafanhotos pela população chinesa, no período de plantações. Os autores afirmam que os melhores e mais ricos exemplos de Ciência Cidadã estão ligados às áreas de agricultura e ciência alimentar.

Estudos de Cohen e outros autores (2015) enfatizam que entre o século XVII e XIX, cientistas de diversas áreas do conhecimento, como dos campos da arqueologia, ornitologia e astronomia apropriaram-se do conhecimento e do uso de cidadãos cientistas na produção de suas pesquisas, mas enfatiza que nesses estudos não era exigido treinamento ou qualificação específica ou qualquer tipo de remuneração para seus coletores de dados. Colaboradores famosos como Charles Darwin, Benjamin

Franklin e Isaac Newton, também contribuíram para essa prática científica (PERKINS, 2013).

No contexto histórico, projetos de Ciência Cidadã já eram possíveis de identificar na virada do século XIX para o século XX, como projetos de biodiversidade envolvendo a sociedade, destacando-se o “*Christmas Bird Count*”, que consistia na preocupação de cientistas e naturalistas com a redução da população de pássaros, devido à caça lateral⁶. O projeto foi realizado no período de Natal (15/dez. a 05/jan.), por 27 observadores em 25 locais diferentes, o resultado registrou o total de 90 espécies. Como resultados, o ornitologista Frank Chapman, da incipiente Audubon Society⁷ (EUA), propôs uma nova forma de substituir a antiga tradição da caça lateral, pela caça de observação, com o registro das aves, este projeto mantém-se ativo até os dias atuais. (AUDUBON, *online*).

Outro projeto que se destaca é o do meteorologista Cleveland Abbe, no mapeamento das auroras boreais⁸, década de 1870, nos avistamentos no Polo Norte. Tinha por objetivo registrar as latitudes médias. Contou com um grupo de voluntários residentes dos EUA, que faziam seus registros do fenômeno a partir de seus horários locais. O projeto é considerado uma das primeiras formas de colaboração entre o cidadão comum em pesquisas científicas, pois a equipe de observadores era composta de 80 voluntários e 20 especialistas. As dificuldades apresentadas nas observações, em função dos diferentes horários locais de cada voluntário, fizeram com que o país fosse dividido em vários fusos. Em 1883, o Congresso dos Estados Unidos oficializou os fusos horários no país e nas décadas seguintes outros países passaram a adotá-lo com base no Meridiano de Greenwich (BARTKY, 1989).

A Ciência Cidadã tem como característica principal a participação de voluntários em projetos científicos – com essa parceria buscam incrementar a produção científica –, e seus desdobramentos são largamente explorados fora do meio acadêmico, como em *sites* de instituições, de ONGs (Organizações não governamentais), em *blogs* de cidadãos, entre outros (LEAL, 2019, p. 26).

⁶ Escolha do lado onde iriam caçar

⁷ Audubon Society é uma organização não-governamental de conservação da natureza, fundada em 1896, por Harriet Hemenway e Minna B. Hall e em 1905, passa a ser oficial, com o nome de National Audubon Society. É uma das mais antigas organizações dedicadas à temática da conservação da natureza recorrendo à ciência e ao ativismo ambiental.

⁸ A aurora polar é um fenômeno óptico composto de um brilho observado nos céus noturnos nas regiões polares, em decorrência do impacto de partículas de vento solar com a alta atmosfera da Terra, canalizadas pelo campo magnético terrestre.

Por meio da literatura é possível perceber a pluralidade de compreensão e interpretação do que é Ciência Cidadã, pois cada área do conhecimento ou projetos desenvolvidos, criam ou fazem adaptações dessa definição, adequando ao contexto de interesse a ser utilizado, para tanto se faz necessário entender quais as atividades que fazem parte do campo, que possam ser comuns a todas as áreas e assim produzir uma definição ampla que atenda a todos os campos/contexto/área, a nível internacional.

Perkins (2013) menciona que o uso da terminologia Ciência do Cidadão (uma das variantes da Ciência Cidadã) foi forjada pela primeira vez na década de 1970, no Cornell Lab of Ornithology, da Cornell University⁹ (EUA), instituição que desenvolve diversos projetos voltados à Ciência Cidadã, com ênfase na apreciação e conservação de aves.

No entanto, é somente nos meados da década de 1990, que o termo, tal como conhecemos na atualidade, passa a ser utilizado como terminologia científica, para representar a participação do cidadão em pesquisas científicas.

Em 1995, o britânico Adam Irwin cunha o termo Ciência Cidadã como uma ciência desenvolvida e aprovada pelos cidadãos não cientistas, isto é, uma ciência democrática, onde devem ser abertos os processos científicos para o público geral. Ele considerava “as possibilidades de uma abordagem da ciência e da expertise que ofereça, pelo menos, o potencial de diálogo entre grupos de cientistas e de cidadãos” (IRWIN, 1995, p. 33).

Nos Estados Unidos da América (EUA), Rick Booney, em 1996, liderava o Cornell Lab of Ornithology, trabalhava com o mesmo tema que Irwin, embora de forma distinta, também apresenta uma definição para Ciência Cidadã como sendo pragmática e instrumental, voltada para a participação ativa do público geral nos projetos de investigação científica e na preservação do meio ambiente, seria uma ciência participativa (COSTA, 2020).

De acordo com Ryan e outros autores (2018), não existe uma única definição para Ciência Cidadã, ela vem sendo construída ao longo das últimas décadas. Enfatiza que a cooperação dos cientistas cidadãos em pesquisas científicas não é nova, pois fizeram contribuições fundamentais para o entendimento do clima, da astronomia, geologia, questões ambientais, entre outros.

⁹ *Christmas Side Hunt*. <https://www.cornell.edu>

Por atuar em diversas áreas do conhecimento, a terminologia “Ciência Cidadã” pode ser identificada de diversas formas, como: ciência das multidões, ciência do voluntariado, ciência colaborativa, ciência participativa, *crowdsourcing*, *crowdsourcing*, entre outros; bem como variantes da sua própria definição.

A Ciência Cidadã é um dos principais pilares da Ciência Aberta. Yamamoto (2020) ressalta sua importância ao afirmar que ela é uma “ação transformadora para a ciência”, visto que contribui de maneira formal em pesquisas científicas. Existem diferentes interpelações e interpretações para o conceito e o fazer Ciência Cidadã.

Eitzel e outros autores (2017), abordam o conceito de Ciência Cidadã como sendo amplo e ao procurar aumentar sua participação na ciência, pode levar a muitos significados, visto ter de analisar os termos que vão ser utilizados para descrever as atividades, projetos e participantes, assim como pode permitir a inclusão ou exclusão do conhecimento dos cidadãos.

Diante do exposto e da flexibilidade de um único conceito, que pode ser adaptado a diferentes áreas do conhecimento, optou-se pelo entendimento de que Ciência Cidadã é a prática do envolvimento de cidadãos que não são cientistas profissionais, em projetos científicos, isto é, na colaboração de investigações de caráter científico, ou seja, cooperação entre o cientistas profissionais e não-cientistas, nos novos achados da ciência por meio dessa interação, ou ainda, pesquisa científica conduzida, no todo ou em parte, por cientistas amadores ou não profissionais.

Ignat e outros autores (2018, tradução nossa, *online*) entendem como Ciência Cidadã:

[...] à colaboração entre cientistas profissionais e o resto da sociedade. Por favor, note que vemos essa colaboração como uma atividade de mão dupla, que capacitará os cidadãos a desempenhar um papel na atividade de pesquisa e, como resultado, produzir melhorias e fazer descobertas o que será benéfico para a sociedade como um todo. A ciência cidadã efetivamente realoca atividades científicas em uma comunidade mais ampla e oferece motivos para mais interação entre ciência e sociedade.

Nesse contexto, os mesmos autores, acima citados, destacam a importância das bibliotecas no apoio e participação junto à Ciência Cidadã, mas ressaltam que esse apoio depende do interesse e do comprometimento da comunidade em se envolver nas atividades da Ciência Cidadã, assim como o entendimento dos benefícios e recursos financeiros alocados que, sem dúvida, motivariam as bibliotecas a se engajarem completamente com a Ciência Cidadã. Afirmando, ainda, que as

bibliotecas de pesquisa estão bem posicionadas para contribuir para a Ciência Cidadã.

Heigl e outros autores (2019), fazem uma análise sobre as tipologias apresentadas para a Ciência Cidadã e argumentam sobre a necessidade de uma definição internacional. Aponta que muitos profissionais buscam conceituá-la no passado, mas que logo percebem que isso não é mais possível, pois não se consegue abranger todo o campo e extensão do que é a Ciência Cidadã, visto estar diretamente ligada a projetos e defini-la no passado levaria a uma restrição no conjunto da qualidade de critérios a serem utilizados nos projetos.

Apesar dos argumentos dos autores acima citados sobre a definição de Ciência Cidadã, algumas instituições governamentais e privadas, entre outros, expõem suas definições, como por exemplo é o caso da *National Aeronautics and Space Administration*¹⁰ (NASA), que define Ciência Cidadã como sendo uma forma de colaboração aberta na qual indivíduos ou organizações participam voluntariamente do processo científico de várias maneiras. Complementa que os projetos de Ciência Cidadã dependem dos cidadãos para obterem sucesso. Atualmente a NASA tem 36 projetos de Ciência Cidadã em andamento, que estão abertos a todos no planeta para participarem, e abrangem algumas divisões das ciências.

A *Socientize Consortium*, desenvolveu o Livro Verde, onde aborda os principais temas de discussão e recomendações para políticas europeias, nacionais e institucionais voltadas para a Ciência Cidadã, bem como busca estimular o envolvimento dos cidadãos em projetos. Apresenta a definição de Ciência Cidadã como sendo o “[...] engajamento do público em geral em atividades de pesquisa científica, quando os cidadãos contribuem ativamente para a ciência, seja com seu esforço intelectual ou conhecimento circundante, seja com suas ferramentas e recursos” (EU-CITIZEN.SCIENCE, 2013, *online*).

No entender de Rollo (2020), a Ciência Cidadã constitui uma evolução na maneira como a ciência é pensada, criada, comunicada e apropriada e define como sendo uma das formas de aproximação da ciência com a sociedade, por meio da apropriação por parte dos cidadãos e de suas instituições no conhecimento gerado pela ciência, para além do diálogo, como por exemplo a criação e/ou implantação de políticas públicas.

¹⁰ Administração Nacional de Aeronáutica e Espaço (NASA)

Para Parra (2015), a Ciência Cidadã refere-se a uma “ciência de colaboração”, onde a dinâmica de produção na coleta de dados, informações e métodos aplicados em seu processo é entendida como uma forma de democratizar a ciência, onde cidadãos cientistas se apossam ou criam conhecimento científico, podendo intervir até mesmo no processo de tomada de decisão de políticas públicas, envolvendo temas de ciência e tecnologia, e levando à produção de conhecimento em ambientes externos à comunidade científica.

No que se refere à produção científica brasileira sobre Ciência Cidadã, pode-se perceber que está em construção, mas já há abordagens interessantes, como da prof.^a. Sarita Albagli (2015), ao discorrer sobre as iniciativas da Ciência Cidadã, apresentando duas vertentes. A primeira é a pragmática ou instrumental, que é a busca e contribuição do público não acadêmico nas pesquisas incluindo desde compartilhamento de recursos computacionais até coleta de informações de relevância científica, no que já se chama de *crowdsourcing science*. A segunda vertente é a democrática:

[...] são as iniciativas orientadas para maior participação, intervenção e empoderamento de cidadãos não só nas formas de produção e uso, mas nos próprios rumos da pesquisa. É o caso do desenvolvimento de ferramentas abertas e descentralizadas em favor da democratização e apropriação cidadã da ciência e da tecnologia em favor da inovação social (ALBAGLI, 2015, p. 15).

McKinley e outros autores (2017), colocam que a Ciência Cidadã pode contribuir com o engajamento de cidadãos em seus projetos, como por exemplo: promovendo a colaboração, criando sinergias e melhorando seus resultados; trazendo novas perspectivas e informações na tomada de decisões; promovendo a gestão ambiental; disseminando o conhecimento e compartilhando o aprendizado adquirido; respondendo questões de interesse da comunidade, entre outros. Salientam que a Ciência Cidadã pode melhorar na alfabetização científica e construir conhecimentos, ajudando o cientista cidadão a aprender como a ciência é feita, assim como entender e acessar melhor as informações científicas, o que pode ser uma experiência interessante e transformadora.

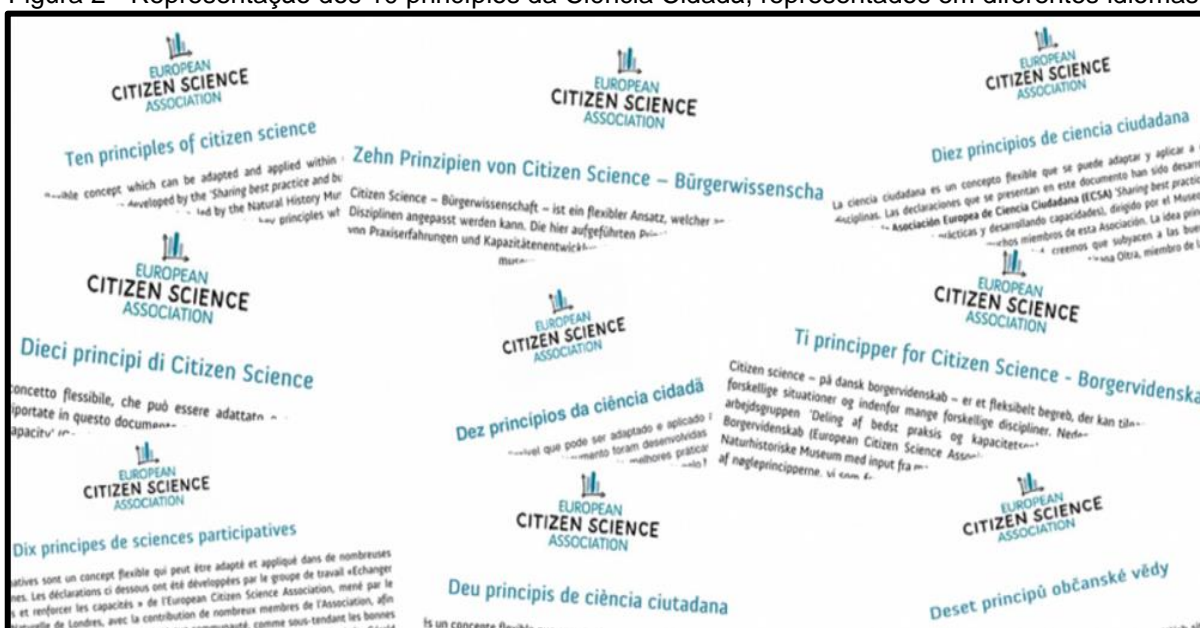
A Associação Europeia de Ciência Cidadã (ECSA – *European Citizen Science Association*)¹¹, em 2015, na cidade de Barcelona, elaborou uma diretriz definindo quais seriam as práticas do que constitui uma boa Ciência Cidadã e redigiu os 10

¹¹ <https://ecsa.citizen-science.net/>

princípios da Ciência Cidadã, mostrando as melhores práticas para um projeto de Ciência Cidadã, levou em consideração as questões de como envolver os cidadãos, a qualidade dos dados coletados, a ética e a legalidade dessas demandas, assim como a avaliação e o resultado dos projetos levando em conta seu impacto social.

A ECSA, permite que façam adaptações nos 10 princípios, desde que os interessados entrem em contato com a sede. Representou-se, na Figura 2, os 10 princípios da Ciência Cidadã, em diferentes idiomas.

Figura 2 - Representação dos 10 princípios da Ciência Cidadã, representados em diferentes idiomas.



Fonte: ECSA - Associação Europeia de Ciência Cidadã (2016)

Os 10 princípios da Ciência Cidadã da ECSA¹² são:

1. Os projetos de Ciência Cidadã envolvem ativamente os cidadãos nas atividades científicas o que gera novo conhecimento e compreensão. Os cidadãos podem atuar como contribuidores, colaboradores ou como líderes de projetos e assumir um papel significativo no projeto.
2. Os projetos de Ciência Cidadã produzem genuínos resultados científicos. Por exemplo, respondendo a uma pergunta de investigação ou colocando em prática as ações de conservação, decisões de gestão ou políticas ambientais.
3. Tanto os cientistas como os cidadãos cientistas beneficiam da sua participação nos projetos de Ciência Cidadã. Os benefícios podem incluir a publicação de resultados da investigação, oportunidades de aprendizagem, prazer pessoal,

¹² <https://zenodo.org/record/5127534#.YPrkNEBCRnE>.

benefícios sociais, satisfação através do contributo em evidências científicas para, por exemplo, encontrar respostas para questões com relevância local, nacional ou internacional e, desta forma, influenciar políticas nesta área.

4. Os cidadãos cientistas podem, caso queiram, participar em várias etapas do processo científico. O que pode incluir o desenvolvimento de uma questão científica, o delinear dos métodos a utilizar, a recolha e análise dos dados e a comunicação dos resultados.
5. Os cidadãos cientistas recebem *feedback* do projeto. Sobre, por exemplo, como os dados recolhidos estão a ser usados e quais os resultados no campo da investigação, política e sociedade.
6. A Ciência Cidadã é considerada como abordagem de investigação como qualquer outra, com limitações e enviesamentos que devem ser considerados e controlados. Contudo, ao contrário das abordagens científicas tradicionais, a Ciência Cidadã providencia oportunidades para um maior envolvimento do público e uma democratização da ciência.
7. Dados e metadados resultantes de projetos de Ciência Cidadã são tornados públicos e sempre que possível publicados num formato de acesso livre. A partilha de dados pode acontecer durante ou depois do projeto, a menos que existam motivos de segurança e privacidade que o impeçam.
8. O contributo dos cidadãos cientistas é reconhecido publicamente nos resultados dos projetos e nas publicações.
9. Os programas de Ciência Cidadã são avaliados pelos seus resultados científicos, qualidade dos dados, experiência para os participantes e abrangência dos impactos sociais e políticos.
10. Os responsáveis de projetos de Ciência Cidadã têm em consideração, questões legais e éticas relativas ao *copyright*, propriedade intelectual, acordos sobre partilha de dados, confidencialidade, atribuição e impacto ambiental de qualquer atividade.

A ECSA (2015), ao identificar as principais características da Ciência Cidadã, por meio de seus 10 princípios, argumenta que alguns desses aspectos podem ser relevantes ou não, isso vai depender dos contextos e objetivos, nos quais vai ser aplicado.

A Ciência Cidadã, ao se apropriar das tecnologias digitais em rede, pode divulgar, por meio de diferentes canais, suas pesquisas a públicos não acadêmicos. O que antes era difundido por meio de conferências, artigos, livros, agora utiliza plataforma de compartilhamento de vídeos, *blogs*, redes sociais digitais, entre outros. Em suma, as tecnologias digitais têm impulsionado a pessoas comum em se tornarem Cientistas Cidadãos ao proporcionar novas formas de engajamento público.

2.2.1 Cientistas Cidadãos

Ao buscar entender o que é Ciência Cidadã, volta-se diretamente para o termo de “cidadão ou cidadã”, que vai indicar, justamente, o papel que os indivíduos ou grupos de pessoas não especializadas desempenham na ciência por meio de sua participação/contribuição em projetos de interesses comuns, valendo-se para isso de seus conhecimentos e experiências nas atividades que vão desempenhar. Sua colaboração pode ser por meio de monitoração, observação, inclusão de dados em aplicativos, coleta e análise de dados e relato de resultados.

Ao participarem dos projetos de Ciência Cidadã, os cidadãos, além de contribuir com o processo de pesquisas científicas, por meio de observações, monitoramento, coleta de dados, entre outros, promovem o aprendizado e ao mesmo tempo aprendem sobre ciência na troca de informações com os cientistas.

Conforme Leal (2019), cientista cidadão é aquele indivíduo que não esteve na academia ou que não tem conhecimento científico do objeto que vai observar/coletar. Podem também ser referenciados de diversas formas, como: cientista cidadão, voluntários, cidadãos voluntários, membros públicos, colaboradores, amadores, leigos, cidadãos, atores não científicos, público, público geral, cientistas públicos, cidadãos locais, não-cientista, participantes leigos, cientista não acadêmico, cientista não profissional.

Para maior compreensão da Figura 3, os termos correspondentes a cidadãos são coloridos em tons de roxo, e os termos correspondentes a cientistas são coloridos em tons de azul-verde. O termo cientista cidadão aparece duas vezes, uma para cidadãos cientistas e outra para cientistas.

Figura 3 - Nuvens de palavras para vários termos dos participantes da Ciência Cidadã



Fonte: EITZEL, M. V e outros autores (2017, ONLINE)

Uma das perguntas mais interessantes que é efetuada sobre cientista cidadão ou voluntário leigo, é qual é a motivação para participar de projetos de Ciência Cidadã? Grey (2009) apresenta como uma das respostas à comunidade, a interação entre os cidadãos, o compartilhamento das informações sobre a "ciência por trás do projeto", seria algo parecido com uma rede social, mas de cunho científico, levando a ciência cidadã a um nível superior.

Validando esse argumento, Bruyere e Rappe (2007), identificam como motivações na Ciência Cidadã, além da interação social/fortalecimento de laços, o aprendizado de novos conhecimentos/projetos, a adequação do projeto aos seus interesses, desejo de ajudar o meio ambiente, valores individuais altruístas e humanitários e autoestima. Nessa mesma linha Eitzel e outros autores (2017), ressaltam que muitos cidadãos acreditam que a ciência deve servir ao interesse público e tem interesse em acrescentar ideais igualitários ao seu estudo.

Maund e outros autores (2020), discorrem sobre essa motivação, alegando que se tem uma compreensão limitada sobre o envolvimento dos cientistas cidadãos em projetos, enfatizam que saber porque são motivadas a participar pode informar o *design* e a eficiência do projeto, aumentando a probabilidade de coletar dados precisos e de longo prazo para informar decisões de conservação robustas e melhorar o recrutamento/retenção de colaboradores. Entendem que os cientistas cidadãos de Ciência Cidadã, ao perceberem que suas motivações na participação estão contribuindo para o projeto e a tendência é de permanecerem comprometidas e

engajadas no mesmo. Ressaltam, ainda, que na literatura, somente 3% dos artigos de Ciência Cidadã, levam em consideração essas motivações.

Morais (2016), em sua explanação, versa sobre o papel dos cidadãos em projetos de Ciência Cidadã, como sendo de caráter tradicional e anterior ao advento da *internet*. No entanto, com o uso dessa ferramenta, se estabeleceram os pilares tecnológico e motivacional. Objetivou, assim, analisar o modelo de interação humana com a tecnologia. Ressaltou que compreender as motivações dos cientistas cidadãos é fundamental para planejar, projetar e gerenciar tais projetos. A motivação dos cidadãos é analisada por meio de entrevistas e de sua produção, na tentativa de caracterizar os seus perfis.

Os cientistas cidadãos passaram a desempenhar um papel fundamental em alguns projetos de pesquisa científica, pois sem sua colaboração, seria quase impossível sua viabilidade, principalmente projetos de grande robustez, onde seria necessário grande demanda de investimentos em recursos humanos e ferramentas, que geralmente em pesquisas científicas são limitados (MAUND *et al.*, 2020).

Portanto, nesses casos, a contribuição dos cidadãos é fundamental, visto que o custo financeiro é praticamente inexistente em projetos científicos de Ciência Cidadã - no que se refere a mão de obra voluntária -, e os resultados de observações, monitoramento, transcrições, registros, armazenamento de dados, entre outros, tornam as pesquisas mais rápidas, dinâmicas, visíveis e constantemente atualizadas, dependendo do tipo e objetivo de cada projeto os resultados que levariam anos de pesquisa tem seu tempo reduzido.

McKinley e outros autores (2017), discorrem sobre como os cientistas cidadãos contribuirão nos esforços de conservação, gestão de recursos naturais e proteção ambiental, por meio de dois caminhos: em primeiro lugar entendendo os desafios de conservação e projetando estratégias, principalmente quando se relacionam as dimensões do espaço geográfico e em segundo lugar relaciona-se ao envolvimento dos cidadãos no processo de conservação, isto é, como sua contribuição pode refletir de forma indireta na proteção ambiental.

As contribuições de cientistas cidadãos em projetos de Ciência Cidadã, diferem do que chamamos de voluntariado social/atividade voluntária ou participação em associações/instituições ou ações voluntárias em prol da sociedade. Nesse caso o voluntariado social disponibiliza seu tempo a uma instituição ou grupo social desenvolvedor do projeto, sem receber valores financeiros em troca e sua contribuição

não gera nenhuma pesquisa científica, diferente do cientista cidadão que contribui para a ciência/pesquisa científica (ROMAR, 2018).

A Ciência Cidadã não contempla o sujeito de pesquisa¹³. O voluntário é o foco do estudo, aquele que permite a coleta de seus dados, como por exemplo em pesquisa na área da saúde (NPCRS, *online*).

As atividades desenvolvidas pelos cidadãos na Ciência Cidadã, não precisam ser efetuadas/desenvolvidas em todas as etapas de um projeto, visto que cada um tem seu espaço temporal (período estimado e local de ação) e peculiaridades distintas, mas todos seguem a metodologia científica de cada proposta e projeto, nada é efetuado de forma aleatória. Sua colaboração agrega resultados em pesquisas científicas, melhoria das condições sociais e ambientais do planeta e, dependendo dos resultados, até mesmo na elaboração de novas políticas públicas (LEAL, 2019).

A participação dos cientistas cidadãos nos projetos é limitada (MARTINS, 2017). Projetos que permitem ao cidadão participar em um grau maior, indo além da coleta, observação e registro de dados, como por exemplo, propor perguntas de pesquisas ou ter acesso ao uso de ferramentas mais sofisticadas na coleta de dados são raros, assim como a oportunidade de ele acompanhar os impactos e os resultados das descobertas gerados por sua colaboração, atuam como mero sensores na hora de coletar dados, e agem como algoritmos quando vão interpretá-los (OLIVEIRA, 2015).

No entanto, para Ries e Oberhause (2015), o uso da *internet* e das novas tecnologias móveis podem aumentar a capacidade de atuação dos cidadãos, podendo inclusive levar a um aumento no uso de dados de Ciência Cidadã em publicações revisadas por pares.

Haklay (2013), discorre sobre o papel dos cientistas cidadãos nos projetos, como exigência da quantidade de tempo disponibilizada, tanto ao ar livre, observação e coleta de dados, como em ambientes fechados, podendo ser no envio ou na análise desses dados. Ressalta que os participantes muitas vezes necessitam de treinamento ou de auxílio para obter as habilidades necessárias na conclusão de tarefas na coleta de dados.

Albagli (2014) identifica na literatura os diferentes tipos de iniciativa de Ciência Cidadã:

¹³ Sujeito de pesquisa, é aquele que fornece amostras para estudos científicos.

a) **Computação compartilhada (*volunteer computing*)** - a contribuição dos cidadãos nos projetos é a disponibilização de seus recursos computacionais na ampliação da capacidade de processamento de pesquisas.

b) **Inteligência distribuída (*volunteer thinking*)** - doação de seu tempo de cognição e inteligência, desenham um papel ativo em pesquisas massivas de análise de dados.

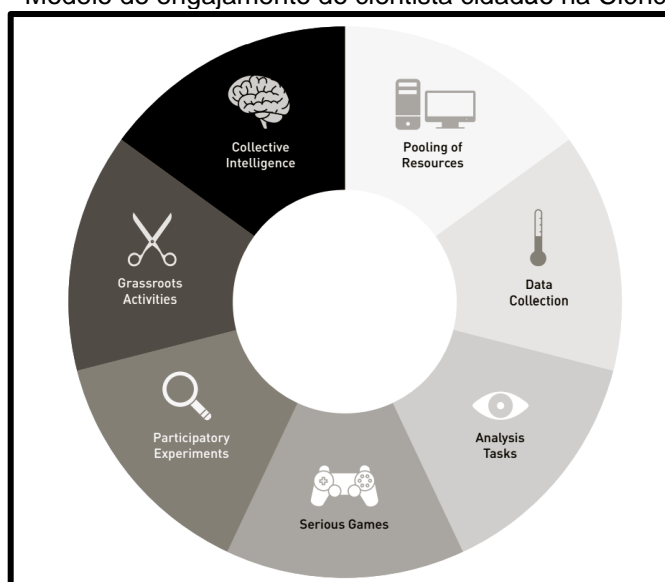
c) **Sensoriamento voluntário (*volunteer sensing*)** - a participação do cidadão é a coleta de dados para as investigações científicas.

d) **Diálogo com a sociedade** - interação entre os cientistas profissionais e os não cientistas, por meio das plataformas na internet.

e) **Pesquisa direta** - criação de espaços comunitários com infraestrutura para o desenvolvimento e a colaboração nos projetos de pesquisa, experimentos e aprendizado.

O importante é o engajamento dos não-cientistas em projetos que contribuem para pesquisas científicas, seja por meio de observações, identificação, coleta, contagem, registros, compartilhamento de dados, que possam gerar não só conhecimento, mas também resultados na melhoria da qualidade de vida dos seres humanos e da biodiversidade do planeta, conforme demonstrado na Figura 4.

Figura 4 - Modelo de engajamento do cientista cidadão na Ciência Cidadã



Fonte: Livro Branco sobre Ciência Cidadã

No Brasil, com o crescente número de projetos que estão surgindo, os cientistas cidadãos estão ganhando cada vez mais espaço, visto que existem diferentes plataformas colaborativas de coleta e análise de dados que armazenam

diferentes projetos de Ciência Cidadã, assim possibilitando ao cidadão escolher aqueles de seu interesse para participarem.

Para fins deste trabalho, tendo em vista que existe uma diversidade de conceitos aplicados, adota-se os termos: ‘Cidadão” e “Cientista Cidadão” para categorizar as pessoas que participam como voluntários, corrobora Leal (2019).

2.2.2 Projetos de Ciência Cidadã

Os projetos de Ciência Cidadã têm tradicionalmente uma base científica e não comercial (SOARES e SANTOS, 2011). Portanto, eles têm o objetivo de tornar os resultados da pesquisa científica mais democráticos e abrangentes, integrando diversas perspectivas, novas bases de conhecimento e outras formas de produzir ciência, com foco na inovação.

Conforme McKinley e outros autores (2017), os projetos de Ciência Cidadã podem ser de pesquisas básicas ou aplicadas, observar linhas de orientação ecológica ou ambientais, atender crise e apontar ações de gestão, tratar de questões em escala local, regional ou global.

Neste contexto, Fritz, See e Grey (2022), complementam a importância do uso potencial de seus dados ambientais, gerado por projetos de Ciência Cidadã por governos, autoridades locais, entre outros, assim como elaboração de projetos menores de baixo custo.

Ao longo dos últimos anos, vem ocorrendo a proliferação de inúmeros projetos, principalmente nas áreas de biodiversidade e saúde, atribuindo-se ao aumento do conhecimento do público em geral (CONRAD, 2011), bem como a motivação da valorização dos projetos pela comunidade científica. A facilidade de utilização da tecnologia e de recursos móveis viabiliza o acesso e a participação de cidadãos não cientistas na observação e coleta de dados, assim como no registro, de forma rápida, nas plataformas de projetos escolhidos em qualquer lugar do planeta.

Para Cohen e outros autores (2015), os projetos de Ciência Cidadã podem ser classificados em categorias ativa e passiva. Projetos de categoria ativa são os que incluem os cientistas cidadãos ativamente na coleta de dados, como por exemplo observação e coleta de informações que serão analisados posteriormente por pesquisadores qualificados. A categoria de projetos passivos, os cidadãos oferecem espaços físicos (que seja possível, por exemplo colocação de armadilhas) ou

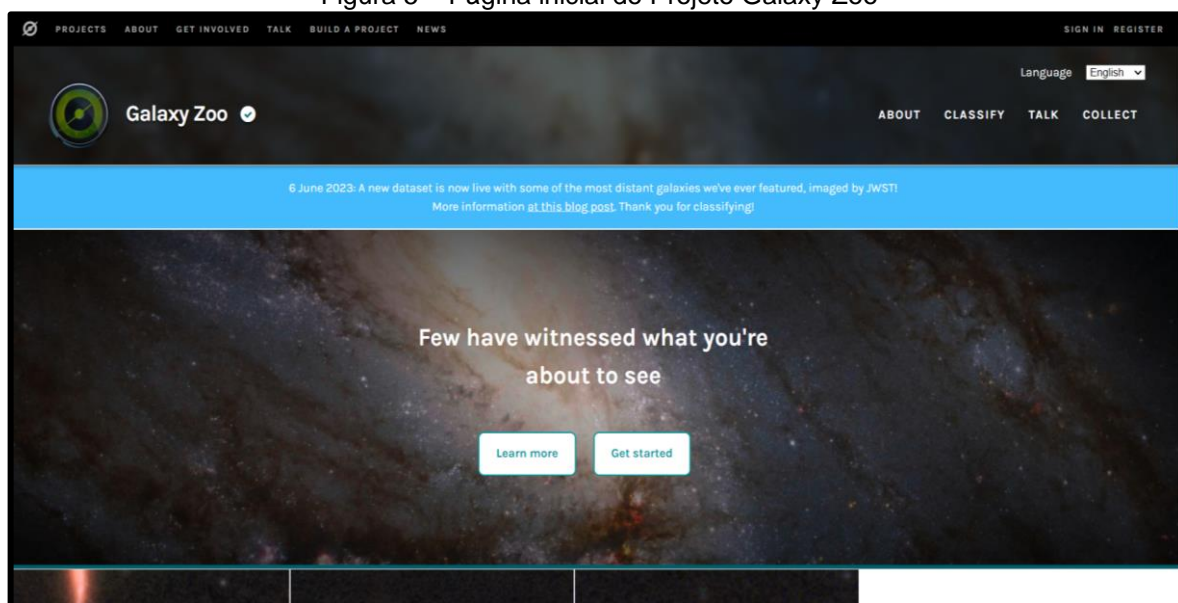
computacional (período em que o computador pessoal não está sendo usado) para as investigações, eles não atuam diretamente no processo de observação e coleta de dados.

Cohen e outros autores (2015), ainda, abordam que devido à grande variabilidade de pesquisas em Ciência Cidadã e a inviabilidade da presença de cientista cidadão presencialmente em todos os lugares, os projetos de Ciência Cidadã podem ser coletados em ambientes internos e externos. Nos espaços externos o cientista cidadão pode locomover-se livremente de um lugar para outro, participando ativamente das coletas de dados. Nos ambientes internos os cidadãos poderão incluir tarefas nas interfaces da *web*, por meio de seus dispositivos móveis, o que vai proporcionar maior visibilidade e alcance para os pesquisadores. Enfatizam que projetos baseados na *web*, podem ser denominados de "ciber-ciência cidadã" (termo utilizado por François Grey, em 2009) e divididos em três categorias: computação voluntária, pensamento voluntário, e sensoriamento participativo.

Para Grey (2009), os projetos de cunho científico *online* permitem aos cidadãos classificarem imagens, como por exemplo as do projeto Galaxy Zoo¹⁴, onde as imagens são classificadas como elíptica ou espirais. Em um período de poucos meses do início do projeto, cerca de 100 mil voluntários classificaram mais de 1 milhão de galáxias, com mais perfeição que qualquer algoritmo de computador, isso ocorre porque a mesma imagem pode ser classificada por diferentes cidadãos, mostrando que a média estatística é mais precisa que a de um astrônomo profissional. A Figura 5 apresenta a página inicial do projeto.

¹⁴ <https://www.zooniverse.org/projects/zookeeper/galaxy-zoo/>

Figura 5 – Página inicial do Projeto Galaxy Zoo



Fonte: Zooniverse (2023)

Já, Bonney e outros autores (2009), descrevem que existe pelo menos três modelos de cooperação em projetos de Ciência Cidadã e os descrevem:

1 - **Projetos de contribuição:** os cidadãos participam apenas na coleta dos dados para pesquisa, não apresentam resultados.

2 - **Projetos de colaboração:** é onde os cientistas cidadãos, além da coleta de dados, se envolvem em outros aspectos do processo científico, como por exemplo no desenvolvimento de objetivos ou questões relacionadas à pesquisa, na análise de dados ou mesmo na divulgação das descobertas.

3 - **Projetos co-criados:** são os projetos elaborados por cientistas e pelos cidadãos, na definição do problema, no desenho da pesquisa e também em todas as fases do processo científico.

Conforme Haklay (2013), a maioria dos projetos desenvolvidos pela Ciência Cidadã estão vinculados à biodiversidade e aos desafios do desenvolvimento sustentável, mas essa apropriação pode ser feita por diferentes campos acadêmicos, como:

[...] antropologia, geografia, ciência da computação, interação humano-computador, ciência da informação geográfica e engenharia – que estão engajados na realização de pesquisas que possibilitem a participação das comunidades em todos os estágios de investigação científica, desde a definição do problema, até a coleta e da análise de dados (HAKLAY, 2013, *online*).

Tal como, o da BBC como o projeto *WW2 People's War*¹⁵ (Guerra Popular da Segunda Guerra Mundial), que foi realizado no período de 2003 a 2006, onde a participação do cidadão era que contribuísse com as memórias e artefatos daquele período histórico da guerra, em seu *site*. Tinha como objetivo formar a base de um arquivo digital que forneceria um recurso de aprendizagem para gerações futuras. Como se percebe, projetos também têm período de validade, no caso deste, que foi encerrado e a página na *web* consta como arquivada, mas seus resultados são de dados abertos, portanto é possível visualizar todo o desenvolvimento do projeto e seus resultados.

O Programa Horizonte 2020 (H2020)¹⁶, apresenta-se como o maior programa de investigação e inovação da Europa, vem se destacando ao financiar diferentes associações/instituições voltadas para projetos de Ciência Cidadã, em diferentes áreas do conhecimento, busca contribuir para um crescimento inteligente, sustentável e inclusivo, onde se possa produzir ciência e tecnologia de ponta. Essa associação procura estabelecer uma cooperação eficaz entre a ciência e a sociedade. Apoia e financia projetos que envolvam os cidadãos nos processos de investigação científica.

Como exemplo da atuação e eficiência do Programa H2020, tem-se a plataforma da *EU-Citizen.Science*¹⁷, que é financiada pelo programa e está empenhada na busca da consolidação dos conhecimentos oriundos da Ciência Cidadã em suas boas práticas, assim como incentivar a participação popular e acadêmica em pesquisas e projetos, pois entende que sejam novos caminhos para uma governança participativa, possibilitando que seja desenvolvida políticas públicas que beneficiem a todos. Ressalta que a Ciência Cidadã, por ter conceito flexível, pode ser aplicada em diferentes contextos, desde que os cientistas cidadãos possam estar envolvidos, permitindo assim a democratização da ciência.

O Zooniverse¹⁸ é considerada a maior plataforma do mundo em pesquisas de Ciência Cidadã. O portal possibilita que milhões de voluntários, no mundo inteiro, participem de projetos científicos, ajudando cientistas e profissionais na classificação de galáxias. Desse modo, torna executáveis pesquisas que não seriam possíveis sem esse contingente, seja ele individual ou não, pois o objetivo é a produção de dados

¹⁵ <http://www.bbc.co.uk/history/ww2peopleswar/>

¹⁶ <https://ec.europa.eu/programmes/horizon2020/>

¹⁷ <https://www.plataforma.edu.pt/eu-citizen-science>

¹⁸ <https://www.zooniverse.org/>

confiáveis e precisos, de forma a economizar tempo e recursos, e por consequência levando a novas descobertas.

Projetos como o *Transcribing Torrey*¹⁹ (Transcrevendo Torrey), na área de humanidades, foi desenvolvido pela Biblioteca LuEsther²⁰ T. Mentz, em Nova Iorque, com objetivo na preservação, transcrição e digitalização das anotações, correspondências, obras e notas e ilustrações, escritas e recebidas por John Torrey, datadas do século XIX. Torrey foi um dos maiores botânicos dos EUA do século XIX e sua documentação é de caráter histórico, de acordo com a Figura 6. Para ser voluntário no projeto, bastava saber ler, escrever e entender a caligrafia da época. Projeto já encerrado.

Figura 6 – Página inicial do Projeto de transcrição dos documentos de John Torrey

The screenshot shows the NYBG website page for the John Torrey Papers project. The page includes the NYBG logo, navigation links (Mertz Library Homepage, Research Guides, Archives & Library Special Collections, John Torrey Papers, Transcribing the John Torrey Papers), a search bar, and a main section titled 'TRANSCRIBING TORREY'. The main section contains a description of the project, a list of resources, and a call to action to become a volunteer. The resources listed include: Guidelines for Transcribing the John Torrey Papers, Names Frequently Appearing in the John Torrey Papers, Transcribing Torrey Tip Sheet, and NYBG Archives Finding Guide for the John Torrey Papers. The page also features a portrait of John Torrey and a sample of his handwritten letters.

Fonte: Biblioteca LuEsther T Mentz

O Brasil, em relação a outros países, ainda tem muito a desenvolver na área de Ciência Cidadã, mas já é possível localizar plataformas que promovem a Ciência Cidadã e suposto encontrar projetos interessantes ativos, independentemente do tempo de seu desenvolvimento ou do número de cientista cidadão envolvidos.

Um dos exemplos é o Sistema de Informação da Biodiversidade Brasileira (SIBBR)²¹, desenvolvido pelo Ministério da Ciência e Tecnologia. Inovações e Comunicações (MCTIC), com apoio da ONU *environment programme* (programa de meio ambiente), a primeira estrutura brasileira de dados de informação em

¹⁹ <https://libguides.nybg.org/JohnTorreyPapers/transcribe>.

²⁰ <https://www.nybg.org/learn/mertz-library/>

²¹ <https://www.sibbr.gov.br/>

biodiversidade e seus ecossistemas, com o objetivo agrupar e proporcionar o acesso a dados e informações para, contribuir com políticas públicas, no apoio às práticas de conservação e de uso sustentável assim como promover o conhecimento da diversidade brasileira. Atualmente armazena 53 projetos, como o SOS Mata Atlântica, BioGeoMar, Arara ECOS, entre outros.

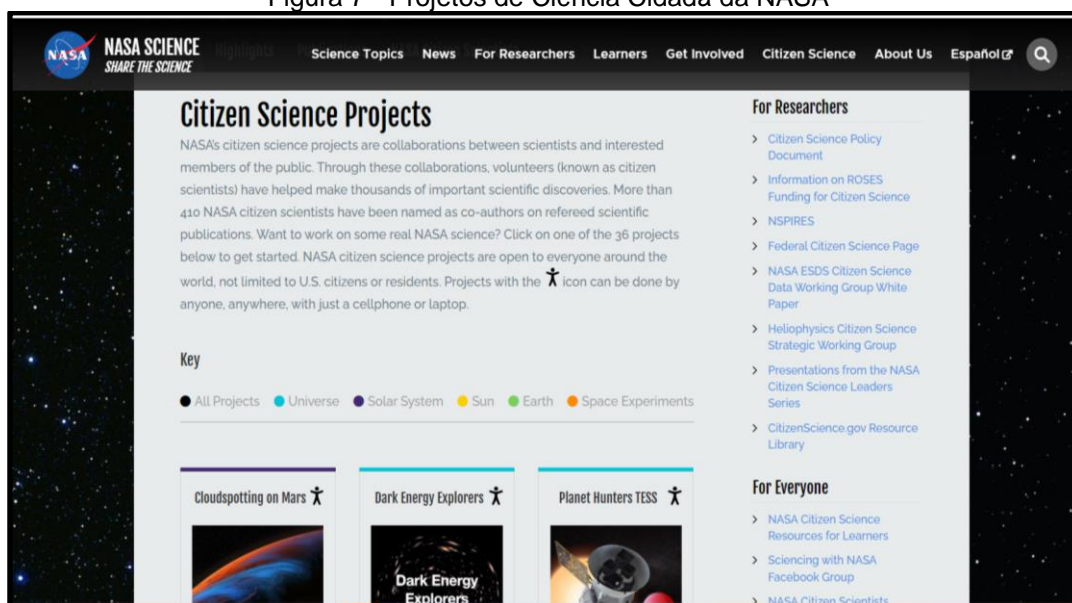
Ressalta-se que, assim como em outros países, o "fazer" Ciência Cidadã também é uma prática antiga no Brasil. Tem-se como pioneiro em projetos de Ciência Cidadã, o médico Vital Brazil. Em 1911, com os estudos de Vital sobre o soro antiofídico, o Instituto Butantan passou a receber serpentes. O projeto se dava por meio da participação da população nas permutas de cobras por soro antiofídico. Os cidadãos ao coletarem as serpentes e entregarem ao Instituto tornavam-se aliados na prática científica, possibilitando que pesquisas e resultados fossem divulgados (TEIXEIRA; COSTA e ZAHER, 2014).

Outro projeto brasileiro que ganha destaque é o Atlas de Registros de Aves Brasileira²² (ARA), um portal interativo que busca reunir dados de ocorrência de aves silvestres no país. Sua proposta é estruturar parte da informação já disponível da classificação das aves brasileiras, facilitando assim o acesso aos dados, bem como permitir a colaboração de pesquisadores e observadores amadores.

De acordo com Soares e Santos (2011), os projetos de Ciência Cidadã são diferentes de estudos aparentemente semelhantes, como o *Seti@home*, que analisa os sinais de rádio em busca de indícios de vida extraterrestre inteligente. Já o *Einstein@home* procura sinais astrofísicos fracos de estrelas giratórias de nêutrons. Ambos utilizam o tempo ocioso de computadores pessoais no envolvimento de análises automáticas. Assim como o *Clickworkers*, de iniciativa da NASA, onde os cidadãos contribuem na identificação e estabelecimento da idade das crateras de Marte (MARTINS, 2017). Atualmente, são 36 projetos da NASA, que estão abertos para cientistas cidadãos de todo o mundo participarem, de acordo com a Figura 7. São projetos de *crowdsourcing*, administrados por instituições científicas e apresentam uma característica em comum que é a colaboração dos cidadãos no fornecimento de dados (MARTINS, 2017).

²² <http://ara.cemave.gov.br/>

Figura 7 - Projetos de Ciência Cidadã da NASA



Fonte: NASA²³

McKinley e outros autores (2017), reportam que projetos de Ciência Cidadã desenvolvidos em circunstâncias corretas, podem trazer muitos benefícios, visto que pode envolver pessoas nos processos de tomada de decisões, incentivar os participantes a se tornarem responsivos e a participarem de reuniões públicas contribuindo com sugestões nas ações propostas e recomendarem novos projetos de Ciência Cidadã.

Conforme Leal, Silva e Gabriel Junior (2021) os projetos de Ciência Cidadã, possibilitam, aos cientistas cidadãos, conhecimento científico sobre conservação, preservação, cidadania, entre outros, e em contrapartida proporciona aos cientistas informações que não teriam acesso rápido se não existisse o compartilhamento, por exemplo, do monitoramento e coleta dados que são efetuadas por esse grande grupo. Os autores ponderam que os projetos podem ser desenvolvidos por um tempo determinado, com data de início e de fim, ou então por um tempo indeterminado, sem data prevista para ser encerrado.

Os projetos podem gerar uma quantidade significativa de dados úteis para a sociedade, além de subsidiar outras pesquisas, entretanto, é fundamental que esses dados estejam disponíveis para uso. Uma das vertentes digital da Ciência Cidadã é conhecida como *Crowdsourcing*.

²³ <https://science.nasa.gov/citizenscience>

2.2.3 Crowdsourcing

Com o uso das Tecnologias da Informação e Comunicação, algumas aplicações destinam-se ao uso *online*. No caso da Ciência Cidadã, esse contexto se aplica ao uso em ferramenta *online*, denominada de *crowdsourcing*.

Jeff Howe e Mark Robinson (2006a) foram os primeiros a forjarem o termo *crowdsourcing*, em 2006, ao fundirem as palavras *crowd* (multidão) e *outsourcing* (terceirização), para descreverem novos modelos de negócios voltados para a *web*.

Para eles *crowdsourcing*:

[...] representa o ato de uma empresa ou instituição assumir uma função uma vez executado por funcionários e terceirizando-o para uma rede indefinida (e geralmente grande) de pessoas na forma de uma chamada aberta. Isso pode assumir a forma de produção entre pares (quando o trabalho é executado colaborativamente), mas também é frequentemente realizado por indivíduos individuais. O pré-requisito crucial é o uso do formato de chamada aberta e da grande rede de potenciais trabalhadores (2006a).

A partir de então, muitos pesquisadores vêm relacionando a Ciência Cidadã com a definição de *crowdsourcing*, voltado especialmente para as plataformas *online*.

O site *CrowdSourcing Week*²⁴ (2019), define como sendo:

[...] prática de envolver uma “multidão” ou grupo para um objetivo comum – geralmente inovação, solução de problemas ou eficiência. É alimentado por novas tecnologias, mídias sociais e web 2.0. O *crowdsourcing* pode ocorrer em muitos níveis diferentes e em vários setores. Graças à nossa crescente conectividade, agora é mais fácil do que nunca as pessoas contribuírem coletivamente – seja com ideias, tempo, experiência ou fundos – para um projeto ou causa. Essa mobilização coletiva é *crowdsourcing*.

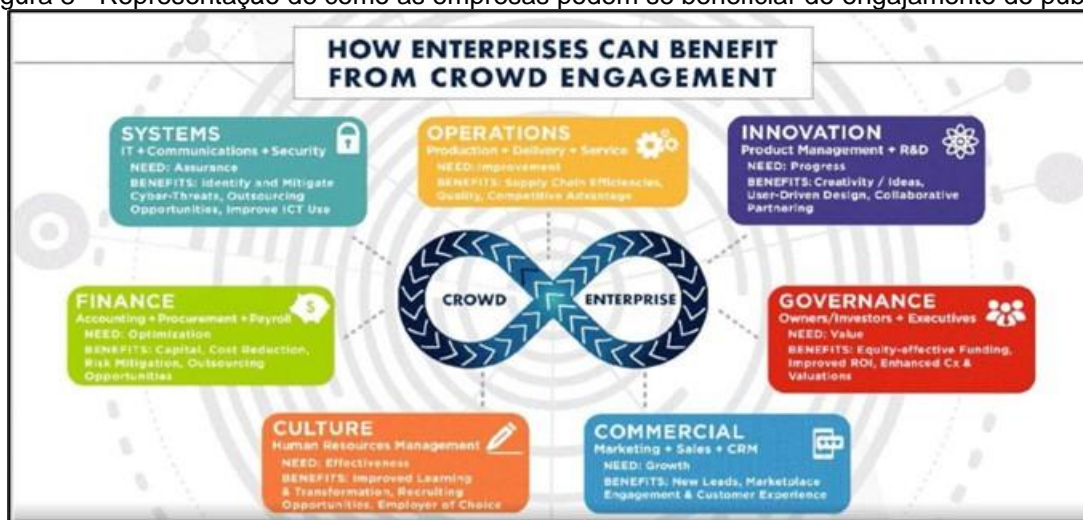
Estelles-Arolas e Gonzáles-Ladron-de-Guevara (2012), apresentam o *crowdsourcing* como uma atividade *online* participativa na qual uma pessoa ou uma empresa ou uma instituição ou ainda uma organização sem fins lucrativo, submete à apreciação de uma tarefa a um grupo indivíduos, com diferentes características e de números variados, por meio de uma chamada aberta flexível a se voluntariaram para efetuarem uma determinada tarefa. A complexidade e estratégia de construção do processo para a realização das atividades podem ser variadas: a multidão pode participar com seu trabalho, dinheiro, conhecimento e/ou experiência, tendo como resultado o benefício mútuo. O cidadão receberá algum tipo de benefício, seja financeiro, satisfação pessoal, reconhecimento social ou até mesmo o

²⁴ <https://crowdsourcingweek.com/>

desenvolvimento de uma habilidade individual e o *crowdsourcer* se beneficia e vai utilizar-se dos resultados que os cientistas cidadãos trouxeram para o projeto/empreendimento, que vai depender do tipo de tarefa que foi efetuada.

A Figura 8, mostra a representação das diferentes áreas de atuação, na qual o cientista cidadão pode contribuir e com seu comprometimento trazer benefícios para as empresas.

Figura 8 - Representação de como as empresas podem se beneficiar do engajamento do público



Fonte: Retirado do site: <https://crowdsourcingweek.com/what-is-crowdsourcing/>

Eklund, Stam e Liebermann (2019, *online*), salientam que o conceito de *crowdsourcing* vem sendo utilizado em alguns sites no aproveitamento do tempo, da energia e do talento das pessoas, em atividades que os computadores não podem fazer. Definem “[...] *crowdsourcing* como um processo digital empregado para obter informações, ideias e solicitar contribuições de trabalho, criatividade e assim por diante, de grandes multidões *online*”. Acrescentam ainda que é a versão digital da Ciência Cidadã, de etimologia híbrida: multidão mais terceirizada, visto que os cientistas envolvem as multidões para participar de experimentos *online*, como criação de livros infantis, tradução de manuscritos antigos, entre outros.

Da mesma natureza do *crowdsourcing* temos o *crowdcrafting*, que pode ser visto como um desdobramento da Ciência Cidadã, mas voltado para a *WEB*²⁵:

Crowdcrafting é um serviço baseado na *Web* que convida voluntários a contribuir para projetos científicos desenvolvidos por cidadãos, profissionais ou instituições que precisam de ajuda para resolver problemas, analisar dados ou concluir tarefas desafiadoras que não podem ser feitas apenas por

²⁵ *World Wide Web* - designa um sistema de documentos em hipermídia que são interligados e executados na internet

máquinas, mas requerem inteligência humana (VERHULST, 2016, tradução nossa, *online*).

Gide (2019), reporta que os conceitos de ciência colaborativa, Ciência Cidadã e *crowdsourcing*, apresentam poucas distinções uns dos outros e na forma como ocorre a colaboração entre os diferentes conceitos e no nível de participação do cientista cidadão, por meio de etapas.

A Ciência Cidadã reflete, o reconhecimento dos dados e conhecimentos gerados pelos cidadãos que participam dos projetos desta área, como dados científicos, contribuindo para a produção científica em suas diferentes etapas. Assim inclui responsabilidade social e as atribuições científicas, considerando o envolvimento do público não acadêmico no processo.

2.3 REPOSITÓRIOS DE DADOS DE PESQUISA

Com um número gigantesco de produção científica disponibilizada na *web*, as universidades, instituições, associações, órgãos governamentais e privados viram a necessidade de terem um dispositivo que reunisse e armazenasse os dados procedentes das pesquisas científicas. Os repositórios permitem o compartilhamento de dados brutos, que facilita a reutilização e a reprodução em outros estudos, permitindo a confiabilidade do estudo original.

Para a RNP (2021), repositório de dados de pesquisa são bases de dados digitais que buscam atender as necessidades de armazenar, organizar e disponibilizar documentos digitais, isto é, albergam os dados coletados durante pesquisas científicas. Os dados de pesquisa que são depositados em repositórios são passíveis de citação, caracterizam-se como produção científica legítima e são reconhecidas, assim, contribuem para o aumento da visibilidade da produção científica.

Segundo Silva Júnior e Borges (2014), os repositórios digitais foram as primeiras plataformas virtuais de acesso aberto. Nesses espaços era possível reunir informações e dados de diferentes campos do conhecimento, tendo como principal função fomentar a sua visibilidade. Os autores reportam que os repositórios podem ser de diferentes tipos:

- a) **Temáticos ou disciplinares**- reúne conteúdo de disciplinas ou assuntos específicos.
- b) **Nacionais** - alberga a produção científica de um determinado local/país.

- c) **Institucional** - criados por instituições de investigação científica (universidades, centros de investigação, laboratórios) para reunir, preservar e disseminar a produção intelectual de uma instituição.
- d) **Pesquisa ou de dados científicos** - criados por diferentes tipos de organizações para recolher, preservar e compartilhar os conjuntos de dados de pesquisa resultantes das atividades investigadas.

Criado na década de 1990, na Universidade Harvard, o *Dataverse*²⁶ é um dos pioneiros e principais sistemas de abertura e de gestão de dados científicos do mundo. Surgiu na área das Ciências Humanas e Sociais para atender a necessidade de compartilhamento de dados quantitativos. É um repositório de dados de código aberto que compartilha, preserva, cita, explora e analisa dados de pesquisa, de forma gratuita. Atualmente, é adotado por inúmeras instituições de pesquisas e centros de dados de CA, é usado para armazenamento e gerenciamento de dados de todos os domínios do conhecimento (LAENDER, *et al.*, 2020).

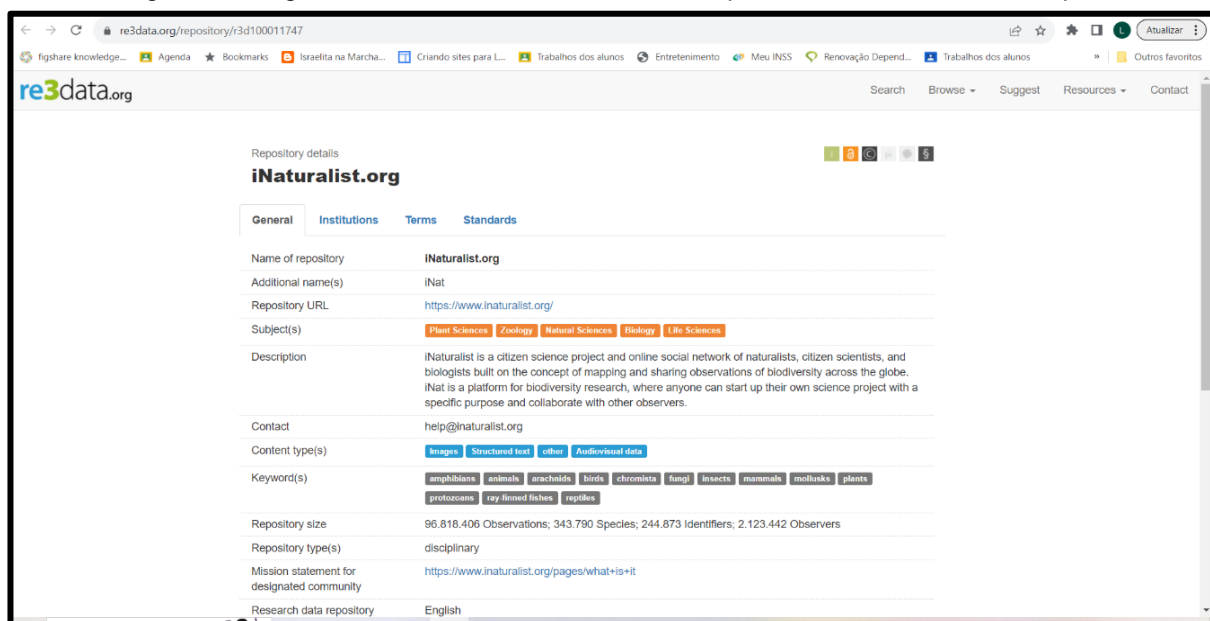
Segundo Rice e Southal (2016), instituições, editores, autores, pesquisadores, entre outros, que armazenam dados no *Dataverse*, recebem os créditos devido, por meio de citação de dados com um identificador persistente, como por exemplo o DOI e o ORCID.

Conforme Gabriel Júnior e outros autores (2019), em 2012 foi lançado um registro global de repositórios de dados de pesquisa, o *Re3data*²⁷, que tem por objetivo fomentar o compartilhamento, maior acesso e visibilidades dos dados de pesquisa, em diferentes áreas do conhecimento. Os repositórios de dados para efetuarem o registro no *Re3data*, devem seguir determinadas regras como: ser administrado por uma entidade legal, esclarecer condições de acesso aos dados e ao repositório, termos de uso e foco em dados de pesquisa. Figura 9 apresenta a página inicial do iNaturalist no *Re3data*.

²⁶ Harvard *Dataverse*. <https://dataverse.harvard.edu>

²⁷ *Re3data* é uma ferramenta de ciência aberta que oferece aos pesquisadores, organizações de financiamento, bibliotecas e editores uma visão geral dos repositórios internacionais existentes para dados de pesquisa. https://en.wikipedia.org/wiki/Registry_of_Research_Data_Repositories

Figura 9 – Página do iNaturalist no Cadastro de Repositórios de Dados de Pesquisa



Fonte: Re3data²⁸

Outro Diretório de destaque é o *OpenDoar*²⁹, lançado em 2005, com parceria entre as Universidade de Nottingham e a Universidade de Lund. É um Diretório global de repositórios de acesso aberto, onde é possível pesquisar repositórios de acesso gratuito e aberto registrados, com base em diferentes recursos e ferramentas acadêmicas. Para manter a qualidade e confiabilidade, cada novo registro de repositório é revisado e processado por membros da equipe editorial, mantendo, assim, um serviço confiável para a comunidade.

Os repositórios de dados contribuem para disseminação de informações científicas. Conforme Sayão e Sales (2015, p.53):

Os repositórios de dados de pesquisa são banco de dados que recebem, gerenciam e disseminam dados e conjuntos de dados de pesquisa. Eles constituem o lugar mais apropriado para que seus dados sejam preservados e possam ser recuperados, acessados e citados por outros pesquisadores, ou seja, tenham visibilidade em escala mundial.

No Brasil, tem-se como exemplo de repositórios de dados aberto de pesquisa, na área de saúde, o da FIOCRUZ. O Repositório Institucional da FIOCRUZ– ARCA³⁰, foi criado em 2007, com objetivo de preservar a produção intelectual em acesso aberto do Instituto de Comunicação e Informação Científica e Tecnológica em Saúde (ICT).

²⁸ <https://www.re3data.org/>

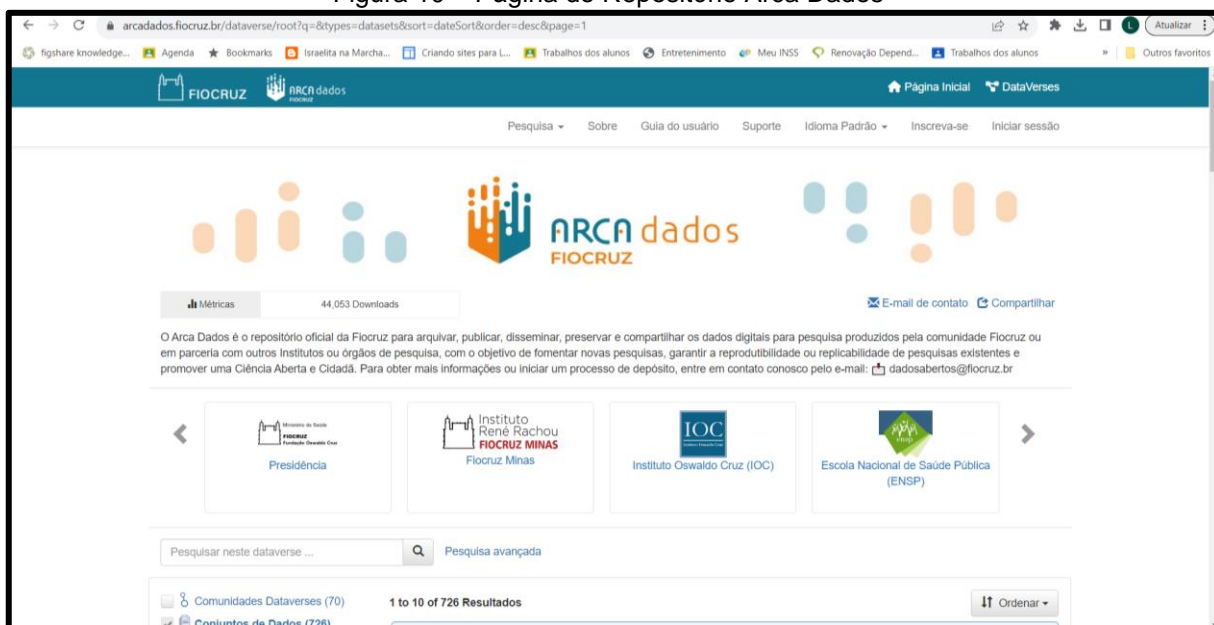
²⁹ <https://v2.sherpa.ac.uk/opensoar/about.html>

³⁰ www.arca.fiocruz.br

Em 2011, O ARCA é oficializada como repositório institucional pela FIOCRUZ e mantido pelo ICICT. Com os princípios de CA aliados as diretrizes da Instituição, a FIOCRUZ busca envolver a sua comunidade científica e públicos interessados, ao disponibilizar cursos à distância sobre Ciência Aberta, que aborda em seu conteúdo a Ciência Cidadã e Dados Aberto.

Como repositório oficial de dados de pesquisa da FIOCRUZ, tem-se o Arca Dados³¹, que objetiva arquivar, publicar, disseminar, preservar e compartilhar os dados digitais produzidos pelos pesquisadores da instituição ou em parceria com outros órgãos de pesquisa, fomentando novas pesquisas e assim garantindo a reprodutividade e replicabilidade de pesquisas existentes. A Figura 10, apresenta página do Arca dados.

Figura 10 – Página do Repositório Arca Dados



Fonte: FIOCRUZ

O Instituto Brasileira de Informação e Tecnologia - IBICT, fomenta a abertura de dados ou acesso aberto à informação científica. Alberga em seu *site* o Manifesto de Acesso Aberto a Dados da Pesquisa Brasileira para Ciência Cidadã e uma relação dos repositórios brasileiros de acesso aberto a publicações científicas. Define repositórios digitais (RDs) como “bases de dados *online* que reúnem de maneira organizada a produção científica de uma instituição ou área temática” (IBICT, 2016). Os RDs armazenam arquivos de diversos formatos.

³¹ <https://arcadados.fiocruz.br/>

A EMBRAPA³², por meio de seus repositórios de Acesso Aberto busca compartilhar informações de seus dados espaciais com a sociedade. São repositórios de acesso livre e gratuito de publicações produzidas em décadas de pesquisas científicas. Destacam-se os repositórios: *Redape*³³ – Repositório de Dados de Pesquisa da Embrapa, objetiva preservar e facilitar a busca dos dados de pesquisas elaborados pela Embrapa e está dividido por temas; o Infoteca-e³⁴ – Repositório de Informação tecnológica da Embrapa; o Alice³⁵ – Repositório Acesso Livre à Informação Científica da Embrapa.

Dentre as plataformas brasileiras, destaca-se a Plataforma Brasileira de Ciência Cidadã (Civis), desenvolvida pelo IBICT, em código aberto, a partir da *Eu-Citizen.Science*, da União Europeia, com o objetivo de oferecer infraestrutura e conteúdo que possa contribuir, disseminar e ampliar o conhecimento sobre a Ciência Cidadã, bem como dar assistência para iniciativas desenvolvidas na área. Reúne projetos de Ciência Cidadã em âmbito nacional.

Prince-Jones e outros autores (2022), abordam que os dados coletados em projetos de Ciência Cidadã sobre espécies exóticas, tem pouco menos de um terço de seus dados abertos disponibilizados de forma gratuita em repositórios nacionais e institucionais e, apesar de muito projetos terem gerado artigos revisados por pares, nem todos eram de acesso aberto. Os autores apontam que as razões da não abertura de dados podem ser várias, como por exemplo: questões de licenciamento, limitação de financiamento, empecilhos técnicos ou até mesmo por privacidade dos participantes. No entanto, entendem que esses problemas podem ser resolvidos e superados pelo grupo gestor, mas que é fundamental para a ciência e a sociedade que esses dados estejam armazenados em algum repositório.

2.4 DADOS ABERTOS DE PESQUISA EM CIÊNCIA CIDADÃ

Esta é a era da desinformação e alguns elementos vêm destacando cada vez mais a importância da Ciência Cidadã, na construção do saber, na aproximação com a sociedade e com a ciência, tem-se como exemplo o compartilhamento e transferência de dados de pesquisa, o que vem a apresentar outro caráter para a

³² Empresa Brasileira de pesquisa Agropecuária - <https://www.embrapa.br/>

³³ <https://www.redape.dados.embrapa.br/>

³⁴ <https://www.infoteca.cnptia.embrapa.br/>

³⁵ <https://www.alice.cnptia.embrapa.br/>

ciência, que é a percepção pública (o tornar público), é o entendimento e o conhecimento da sociedade no que se está desenvolvendo na área científica.

Conforme Ellenburg e outros autores (2019), com a invenção da *internet* e novas tecnologias diferenciadas de monitoramento, em projetos de Ciência Cidadã, se tornou possível a coleta e o compartilhamento de dados acessíveis a todos. Destacam, ainda, que o compartilhamento de dados, a nível mundial, quebra as fronteiras geográficas e culturais, pois estamos todos em um mundo conectado virtualmente, por meio da *internet*.

A ciência cidadã determina a abertura e a transparência no que se refere a coleta e compartilhamento de dados, visto que formam os 10 princípios da Ciência Cidadã em projetos definidos pela ECSA. Leva-se em conta que os dados coletados são de domínio público, portanto, os dados e seus resultados devem ser compartilhados publicamente (WOLKORTE, HEESINK; MICHELLE, 2022).

Os dados gerados pela Ciência Cidadã, estão trazendo contribuições significativas para pesquisas científicas, em diferentes áreas de pesquisa. Atualmente, as instituições acadêmicas e associações governamentais e privadas, têm aberto seus dados, devido aos resultados positivos que vem apresentado em seus projetos e no compartilhamento dos mesmos (FRITZ; SEE; GREY, 2022).

A importância da abertura de dados, em alguns casos pode ser fundamental para o desenvolvimento científico, tem-se como exemplo na área da saúde o compartilhamento da abertura dos dados de pesquisa, no que se refere a crise na saúde mundial³⁶, que assolou o planeta, em 2020, devido a propagação da pandemia do Covid-19, onde a troca de informações entre os pesquisadores de diferentes locais foi necessária, visto que possibilitou o desenvolvimento de vacinas que pudessem conter o vírus.

Para melhor compreensão do que é abertura de dados ou dados abertos, se faz necessário a definição do que é dados. Segundo a *National Science Foundation*³⁷, o conceito de dados é:

[...] qualquer informação que pode ser armazenada na forma digital, incluindo texto, números, imagens, vídeo ou filmes, áudio, *software*, algoritmos, equações, animações, modelos, simulações etc. Tais dados podem ser gerados por vários meios, incluindo observação, computação e experimento (2005, p.13).

³⁶ <https://www.oecd.org/coronavirus/policy-responses/por-que-a-ciencia-aberta-e-fundamental-no-combate-a-covid-19-ca4bdcf9/#section-d1e423>

³⁷ National Academies of Sciences, Engineering and Medicine. Open Science by Design - Realizing a Vision for 21st Century Research. <http://nap.edu/25116>.

A *Open Knowledge International*³⁸ (OKF, 2004, *online*) entende que “[...] dados abertos são informações públicas ou privadas, disponíveis para serem acessadas ou reutilizadas por qualquer pessoa, para qualquer fim [...]”. Destaca que para serem considerados dados abertos, conforme o Manual dos dados aberto: governo³⁹ (traduzido e adaptado de opengovdata.org/⁴⁰), elaborado por um grupo de trabalho da *Open Government Data*, em 2007, deve-se seguir os oito princípios balizadores, conforme Quadro 1. Enfatiza-se que a Fundação OKF busca trabalhar com projetos que envolvam a abertura de dados de pesquisa.

Quadro 1 - Princípios de Dados Governamentais Abertos

| Princípios | Descrição |
|---------------------------------|---|
| Completo | Todos os dados públicos são disponibilizados. |
| Primários | Dados coletados na fonte, como maior refinamento possível, sem modificações. |
| Oportuna | Dados atuais, disponibilizados de forma rápida, para preservar, quando necessário, o valor do dado. |
| Acessível | Dados disponíveis para o maior número de usuários e a maior variedade de finalidade. |
| Processáveis por máquina | Dados razoavelmente estruturados para permitir o processamento automatizado. |
| Não discriminatório | Os dados são disponibilizados a todos, sem que seja necessária identificação ou registro. |
| Não proprietário | Dados disponíveis em um formato sobre o qual nenhuma entidade tem controle exclusivo. |
| Sem licença | Dados que não estão sujeitos a nenhuma regulamentação de direitos autorais, patente, marca registrada ou segredo comercial. Restrições razoáveis de privacidade, segurança e privilégio podem ser permitidas. |

Fonte: Da Autora (Open Government Data)

Portanto, se entende que dispositivos que controlem ou impeçam o acesso a dados publicados, *online* devem ser evitados, mesmo que governamentais, visto que

³⁸ *Open Knowledge*, antes conhecida como *Open Knowledge Foundation*, é uma organização sem fins lucrativos que promove conhecimento livre: <https://okfn.org/>.

³⁹ Manual dos dados aberto: governo

[//www.w3c.br/pub/Materiais/PublicacoesW3C/Manual_Dados_Abertos_WEB.pdf](http://www.w3c.br/pub/Materiais/PublicacoesW3C/Manual_Dados_Abertos_WEB.pdf)

⁴⁰ [https://www.w3c.br/pub/Materiais/PublicacoesW3C/Manual_Dados_Abertos_WEB.pdf](http://www.w3c.br/pub/Materiais/PublicacoesW3C/Manual_Dados_Abertos_WEB.pdf)

o objetivo da abertura de dados é que qualquer pessoa, tanto física quanto jurídica, tenham liberdade de acessar, explorar, modificar e compartilhar, para qualquer finalidade, desde que respeitem e deem crédito a fonte e a licença original dos dados compartilhados.

David Eaves, em 2009, desenvolveu as três leis para dados abertos governamentais, que podem ser aplicáveis a dados abertos de modo geral:

1. **Se o dado não pode ser encontrado ou indexado na *web*** - ele não existe.
2. **Se o dado não estiver disponível em formato aberto e legível por máquina** - ele não pode ser reutilizado.
3. **Se algum dispositivo legal não permitir seu compartilhamento** - ele não é útil.

Conforme Silveira (2021), os dados de pesquisa começam a ser valorizados a partir do movimento de abertura de dados e da Ciência Aberta, onde passam a ser vistos como informações importantes na contribuição de projetos de pesquisa e conseqüentemente para a ciência, visto que seu compartilhamento permite o seu uso e reuso por outros pesquisadores.

Jorge e Albagli (2018), complementam que o compartilhamento de dados deve ser elencado por acordo de acesso que permitam as condições de seu uso ou regras que restrinjam o seu compartilhamento.

Ainda a *Open Knowledge International* (OKF, 2004, *online*), aborda os dados abertos por meio de três pontos principais:

- a) Disponibilidade e acesso:** os dados devem ser convenientes e modificáveis e estarem disponíveis sem custo ou custo razoável, fornecidos para reprodução e que possam ser baixados da internet.
- b) Reutilização e redistribuição:** dados que contenham termos que consintam a sua redistribuição e seu reuso, bem como a combinação/interoperabilidade com outros conjuntos de dados.
- c) Participação universal:** o uso, o reuso e a redistribuição dos dados sem restrição de área de atuação, pessoas e grupos, tudo aquilo que possa impedir o uso dos dados para determinados fins.

Parra (2015) destaca a participação de cientistas e não cientistas nos desafios da mediação digital no processo investigativo, assim como as diferentes definições para o que é a Ciência Cidadã, priorizando o conceito de *crowdsourcing*, face ao uso da comunicação digital, bem como as diferentes formas do cidadão participar das

pesquisas. Salieta que um dos desafios enfrentados pela Ciência Cidadã são as condições de acesso e o uso da informação. Ressalta que:

[...] as possibilidades de compartilhamento inauguradas pelas tecnologias digitais colidem com as tendências de mercantilização e privatização dos conhecimentos produzidos. Os novos *enclosures* (para usar a feliz expressão de James Boyle) sobre a produção do comum (informação, conhecimento e cultura) realizam-se tanto através da expansão do regime jurídico de propriedade intelectual como através da adoção tecnológica de *software* e *hardware* proprietários que limitam as possibilidades de uso, modificação e reapropriação das informações (PARRA, 2015, p. 135-136).

Conforme Laender e outros autores (2020), as iniciativas de dados abertos devem ser vistas por dois ângulos: as práticas adotadas pelos pesquisadores e as políticas preconizadas ou incentivadas por agências de fomento e instituições acadêmicas públicas e privadas. Ressaltam que os pesquisadores, normalmente seguem as determinações ou recomendações de suas áreas de pesquisas, como por exemplo na escolha do uso de um determinado repositório, no tipo de armazenamento ou de metadados. Enfatizam, ainda, que a nível internacional os dados sigam os princípios FAIR⁴¹, visto que não basta que sejam abertos, eles precisam ser encontráveis e acessíveis para reuso.

Os dados não precisam estar totalmente abertos para atender os princípios FAIR (Mons e outros autores, 2016). Para atender essa demanda os repositórios disponibilizam diferentes níveis de abertura, totalmente abertos ou disponibilizar os dados sob embargo, com acesso de solicitação razoável para publicação dos metadados (implica somente na descrição do conjunto de dados sem que os dados sejam compartilhados) (WOLKORTE; HEESINK; MICHELLE, 2022).

Como a qualidade dos agrupamentos de dados de pesquisas são importantes, se faz necessário definir critérios para avaliá-los, assim como é essencial ter uma boa gestão de dados, visto que facilita as descobertas, inovações e reutilização dos mesmos. Na busca para atender essa demanda em 2014, o termo FAIR é lançado em um dos *Workshop* de Lorentz, da Universidade de Leiden, na Holanda. Eram descrito como sendo um conjunto de princípios orientadores para tornar os dados Localizáveis, Acessíveis, Interoperáveis e Reutilizáveis⁴², onde tanto máquinas, quanto seres humanos “[...] poderiam descobrir, acessar, interoperar e reutilizar com mais facilidade, com a devida citação, as vastas quantidades de informações geradas

⁴¹ Findable, Accessible, Interoperable e Reusable.

⁴² <https://force11.org/info/the-fair-data-principles/>

pela ciência intensiva de dados contemporânea (FORCE 11, *online*)”. Os princípios FAIR foram publicados em 2016.

Conforme Sayão e Sales (2016), o armazenamento e a disponibilidade dos dados na *web*, atualmente requerem uma gestão mais ativa, para que o reuso de dados atinja sua real potencialidade, como por exemplo em uma adoção de práticas melhores de gestão, na estruturação dos dados para interoperabilidade, na absorção de metadados de qualidade, no licenciamento adequado e na acessibilidade.

A OECDiLibrary em seu documento “*Principles and Guidelines for Access to Research Data from Public Funding*”, orienta a forma de como disponibilizar os dados de pesquisa e o define como sendo “[...] registros factuais (pontuações numéricas, registros textuais, imagens e sons) usados como fontes primárias para pesquisa científica, e que são comumente aceito na comunidade científica como necessário para validar resultados da pesquisa” (OCDE⁴³, 2007, p. 13, tradução nossa).

Ao mesmo tempo que Haklay e outros autores (2021, *online*) ressaltam a importância da transparência de dados em projetos de Ciência Cidadã, enfatizam que nem todos os dados da Ciência Cidadã podem ser de acesso aberto, devido a questões de privacidade, segurança, sensibilidade ou soberania.

Borgman (2010), aborda como desvantagem do compartilhamento de dados os fatos de recompensas e premiações pelas publicações e não pela gestão de dados, visto que organizar dados para a interpretação e uso de outras pessoas demanda muito mais tempo do que organizá-los para si mesmo, além de que, os pesquisadores se preocupam com a competição no estabelecimento dos créditos de pesquisa.

Para Haklay (2020) a cooperação no desenvolvimento da ciência e da pesquisa é fundamental para resolver desafios globais onde os dados de pesquisas devem ser compartilhados. Enfatiza que a comunidade científica se beneficia com a redução do ciclo de inovação. Afirma ainda que a Ciência Cidadã tem um enorme potencial – permitindo coletar dados no terreno e abordar questões globais. A abertura de dados de pesquisa na Ciência Cidadã é essencial pois permite efetuar uma avaliação do desempenho dos projetos e analisar seus resultados.

Crall e outros autores (2010), no artigo “*Improving and integrating data on invasive species collected by citizen scientists*”, apresentam as dificuldades de alguns projetos de Ciência Cidadã, devido à falta de recursos adequados, como o de

⁴³ <https://www.usmd.edu/institutions/profile?Inst=UMCP>

ferramentas para gerenciar e analisar os dados de pesquisas coletados, assim como a resistência de alguns pesquisadores no compartilhamento de seus dados de pesquisa. Alegam que esses elementos são necessários para o compartilhamento e gerenciamento dos dados de pesquisa apurados, visto que a falta dos mesmos dificulta o processo de pesquisa, assim como o seu uso e o reuso, bem como a qualidade dos dados apurados.

Wu, Washbourne e Haklay (2022, *online*), complementam afirmando que em “muitos casos, a Ciência Cidadã permite uma ampla cobertura geográfica e velocidade na atualização das informações e, portanto, ela pode diminuir a lacuna de dados” por meio de compartilhamento dos mesmos.

Em 2016, o IBICT⁴⁴ lançou em seu *site* o Manifesto de Acesso Aberto a Dados da Pesquisa Brasileira para Ciência Cidadã, com intuito de dar continuidade e ampliar sua política de apoio ao acesso aberto livre à informação científica no Brasil, reforçando o manifesto de 2005, ao qual formalizou sua adesão ao movimento iniciado na Europa no início do século XXI. Objetiva apoiar o movimento de Ciência Aberta no Brasil, caracterizado no acesso amplo e irrestrito das fontes primárias de pesquisas, bem como seu compartilhamento, reprodutividade, reuso, redistribuição, entre outros, em diferentes contextos e em pesquisas colaborativas e interdisciplinares (IBICT, 2016).

Nesse sentido Souza, Araújo e Rocha (2018), entendem que a partir da abertura total de todo o processo científico é possível compreender a importância e a contribuição da Ciência Cidadã e sua visibilidade, não só no meio científico, mas também na sociedade, bem como a participação desta, nesse processo, visto que a participação do cidadão é primordial na evolução coletiva das investigações, mas, ressalta que isso só pode ocorrer se forem consideradas as bases da Ciência Cidadã como acesso aberto, plataformas abertas, agenda de pesquisas abertas, além da informação de qualidade e literacia científica⁴⁵.

⁴⁴ Instituto Brasileiro de Informação em Ciência e Tecnologia

⁴⁵ Literacia científica significa que se é capaz de ler com compreensão artigos sobre a ciência na imprensa popular e encetar conversas sobre a validade das suas conclusões.

3 METODOLOGIA

O estudo apresentado caracteriza-se como uma pesquisa básica do ponto de vista de sua natureza, conforme Gerhardt e Silveira (2009), o estudo de natureza básica tem por objetivo criar novos conhecimentos que sejam úteis para o avanço da ciência, sem aplicação prática prevista da pesquisa.

Para alcançar os objetivos propostos foi realizado um estudo de caráter descritivo, onde efetuou-se o levantamento dos dados, por meio de pesquisa bibliográfica, nos textos que versam sobre a temática. A pesquisa descritiva, tem como intuito identificar na literatura estudos já efetuados sobre o tema apresentado, assim sendo possível descrever e interpretar os dados levantados no decorrer da pesquisa.

Conforme Triviños (1987), a pesquisa descritiva solicita ao pesquisador uma série de informações sobre o que deseja pesquisar, visto que esse tipo de pesquisas pretende descrever os fatos e fenômenos de uma determinada realidade.

Para Marconi e Lakatos (2017), a coleta de dados é a etapa da pesquisa em que se inicia a aplicação dos instrumentos elaborados e das técnicas selecionadas, a fim de efetuar a coleta dos dados previstos. Consiste na parte de demanda com maior tempo do pesquisador, exigindo além da perseverança e esforço pessoal, uma atenção redobrada do registro dos dados. Neves (2007, p. 58) entende que “[...] a escolha do instrumento de coleta de dados depende dos objetivos que pretendemos alcançar com a pesquisa e do universo a ser investigado”. Portanto, para atender os objetivos propostos utilizou-se a revisão bibliográfica.

O processo de execução da pesquisa utilizada para a coleta de dados foi por meio de pesquisa bibliográfica, que conforme Köche (2011, p. 123) destaca sua importância, que é “[...] conhecer e analisar as principais contribuições teóricas existentes sobre um determinado tema ou problema, tornando-se instrumento indispensável a qualquer tipo de pesquisa”. Enfatiza que a revisão bibliográfica é efetuada em fontes primárias e secundárias, com relatos do que se produziu e apresenta importantes resultados de pesquisas já realizadas.

Para identificar na literatura, as relações da Ciência Cidadã com Dados de Pesquisa na literatura internacional, realizou-se uma pesquisa bibliográfica, com estratégia de busca pré-definida para atingir os resultados dessas relações, utilizou-se a revisão bibliográfica, para sistematizar e facilitar a análise dos trabalhos

recuperados, utilizou-se o *software StArt*, que será descrito mais adiante. Nos trabalhos selecionados será realizado uma análise de conteúdo.

Para descrever e interpretar os dados coletados sobre Ciência Cidadã e os seus dados abertos, o estudo utilizou a metodologia e técnica de abordagem qualitativa, que segundo Apollinário (2004, p. 151), “[...] trata dos fenômenos prevendo a análise hermenêutica dos dados coletados”. Minayo (2007), complementa que não há necessidade de métodos e técnicas estatísticas, já que o ambiente natural é a fonte da coleta de dados e o investigador/pesquisador é a ferramenta chave, que tende a analisar os dados indutivamente. Portanto, nessa abordagem a interpretação do investigador é imprescindível, porque vai considerar a natureza do que se está pesquisando, nesse caso, a análise de conteúdo, que contempla as etapas definidas na análise de literatura, com a categorização cognitiva dos dados classificados.

Para contemplar os resultados dos dados qualitativo, utilizou-se a análise de conteúdo, que segundo Freitas, Cunha e Moscarola (1997), a análise de conteúdo nada mais é que uma técnica refinada, que exige paciência, rigor, perseverança, disciplina, dedicação e tempo do pesquisador, que utiliza de sua intuição, imaginação e criatividade, na definição de categorias de análise.

No entanto, para Bardin (2016) a análise do conteúdo é um conjunto de instrumentos de cunho metodológico que busca obter, por meio de procedimentos, sistemático a descrição das mensagens e suas inferências nos dados coletados. Enfatiza que a análise de conteúdo se caracteriza como um conjunto de métodos investigativos de análise da comunicação com a utilização de procedimentos sistemáticos e objetivos de descrição do conteúdo dos documentos analisados, devido ao seu rigor metodológico é o que vai validar a pesquisa.

Para tanto, a análise de conteúdo empregada, ocorre por meio de três etapas sugeridas por Bardin (2016):

- 1) **Pré-análise**: organização e leitura e análise dos documentos coletados, - foram utilizados os critérios elegibilidade.
- 2) **Exploração do material** - classificação e categorização dos documentos selecionados na pré-análise - tendo como critério a leitura na íntegra a análise dos artigos elegíveis.
- 3) **Tratamento dos resultados, inferência e interpretação** - condensação e evidência das informações, culminando na análise reflexiva e crítica - relatos e recomendações, pós-qualificação.

Portanto, para delinear a pesquisa bibliográfica, foi necessário estabelecer a estratégia de busca que abrangesse a temática da Ciência Cidadã com as palavras-chave “*citizen science*” (ciência cidadã), “*community science*” (comunidade científica), “*crowdsourcing*”, e com a temática de Dados Aberto com: “*data sharing*” (compartilhamento de dados), “*open data repository*” (repositório de dados abertos) e “*research data*” (dados de pesquisa), por utilizar a base *Scopus*, somente a versão em inglês dos termos foi empregada. O Quadro 2 apresenta a estratégia de busca:

Quadro 2 - Estratégia de Busca para Revisão Bibliográfica

```
TITLE-ABS-KEY(("citizen science" OR "community science" OR
"crowdsourcing") AND ("research data" OR "open data repository"
OR "data sharing")) AND (( PUBYEAR >= 2011) ND (PUBYEAR <=
2022))
```

Fonte: Elaborado pela Autora

A escolha da base de dados *Scopus*, para a revisão bibliográfica, decorreu de seu vasto número de documentos indexados e os mesmos serem frutos de intensa pesquisa e terem passado pela avaliação de uma comissão científica, além do que, a Ciência Cidadã ser realizada dentro da academia, dessa forma, pressupõe-se que a comunicação deve ocorrer em trabalhos acadêmicos.

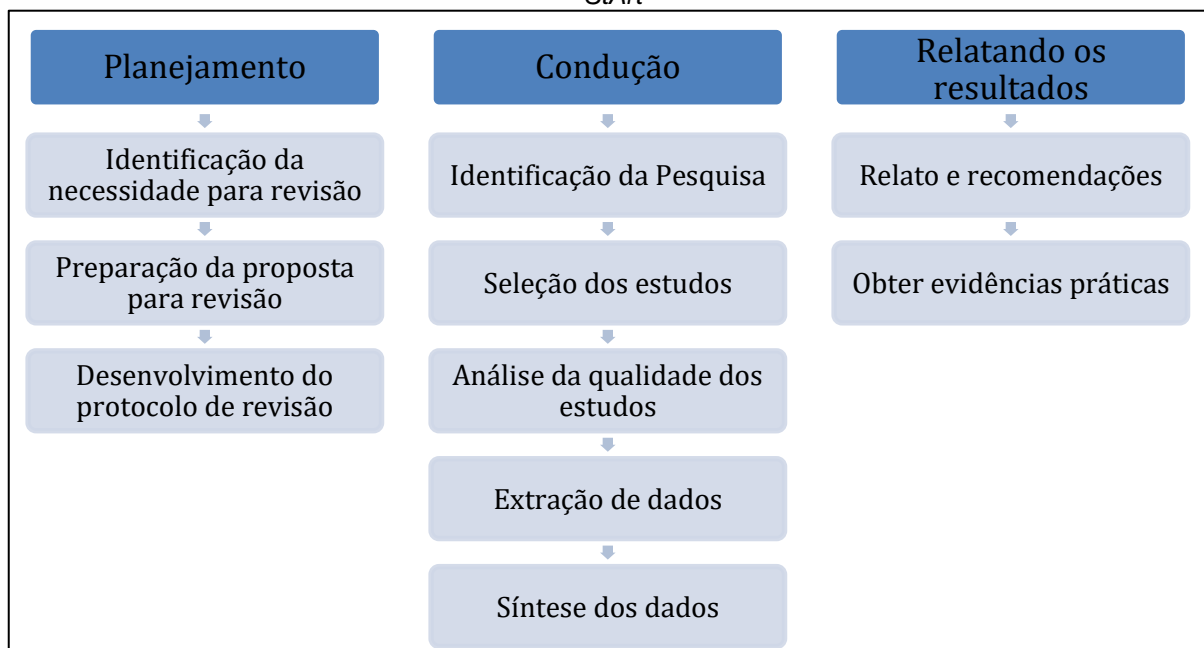
Os metadados completos do resultado da busca na base *Scopus* foram exportados para a ferramenta *StArt*⁴⁶ (*State of the Art through Systematic Review*), no formato BibTeX⁴⁷. Gerando um *corpus* de estudo contendo os documentos científicos, e a descrição dos autores, título, palavras-chave, resumo, ano e comentários visíveis na ferramenta *StArt*.

Utilizou-se como base o modelo da técnica de revisão sistemática de literatura do *software Start*, seguindo as seguintes etapas demonstradas no Quadro 3:

⁴⁶ *State of the Art through Systematic Review* desenvolvida pelo Laboratório de Pesquisa em Engenharia de Software (LaPES) da Universidade Federal de São Carlos (UFSCar).

⁴⁷ BibTeX é uma ferramenta de formatação usada em documentos LaTeX, foi criada para facilitar a separação da bibliografia com a apresentação do texto.

Quadro 3 - Modelo base dos procedimentos para Análise Sistemática de Literatura na ferramenta *StArt*



Fonte: Elaborado pela Autora (2022)

Na etapa planejamento, além da parte onde são preenchidos os dados da descrição do título, autor e o projeto é composta de uma subdivisão, que é o protocolo, que vai atender as necessidades da pesquisa e coleta de dados. Para essa subdivisão, desenvolveu-se os critérios de interesse e foram preenchidos os dados de estratégias necessárias para a coleta de dados, que são: problema de pesquisa, objetivos, palavras-chave, definição de critérios de seleção, idioma, método de pesquisa (*web*), base de dado, seleção dos critérios de inclusão e exclusão, definição do tipo de pesquisa, pesquisa de seleção inicial, avaliação da qualidade da pesquisa e campos do formulário de extração de dados e resumo dos resultados (Galvão, Pansani e Harrad, 2015). O Quadro 4, apresenta os critérios de inclusão e exclusão utilizados no modelo base de Revisão Sistemática de Literatura.

Quadro 4 - Critérios de inclusão e exclusão utilizados no modelo base da Revisão Sistemática de Literatura

| Critérios de inclusão | Critério de exclusão |
|---|--|
| <ul style="list-style-type: none"> ● Publicações nas línguas portuguesa inglesa e espanhola ● Que contenham resumo ● Artigos publicado entre 2012 e 2021 ● Texto completo | <ul style="list-style-type: none"> ● Falta de aderência ao tema e termos selecionados ● Publicações em <i>Workshop</i>, Congressos e Eventos |

Fonte: Elaborado pela autora

Optou-se em não excluir trabalhos sem as palavras-chave, pois em uma observação identificaram-se alguns trabalhos pertinentes e que os metadados não estavam completos.

Como limitador temporal utilizou-se, inicialmente, o período de 10 anos, de 2012 a 2021, por se tratar de um período mais recente da literatura. Limitar o espaço temporal é importante, é ele que vai delimitar o período da amostra do fenômeno a ser estudado. (GIL, 2010).

A partir da definição dos critérios de planejamento, parte-se para a etapa de execução, no subitem estudo de identificação, efetuou-se a leitura do título, palavras-chaves e resumo, buscando selecionar nos artigos indexados os que atendessem aos critérios de inclusão e exclusão.

Para o desenvolvimento da etapa de síntese foi utilizada a classificação e análise da produção anual, autores e palavras-chave, correlacionados com o número de documentos produzidos nos metadados extraídos da base de dados *Scopus*. Para compor resultados mais precisos da revisão, efetuou-se um recorte e analisou-se somente os autores, análise diacrônica da evolução dos trabalhos e análise das palavras-chave, relacionadas a outros temas. Para tanto, utilizou-se o Excel e o Notepad++⁴⁸, para efetuar a organização e limpeza dos dados, bem como agrupar elementos semelhantes como o nome dos autores, palavras-chaves, entre outros.

Outro dado interessante no uso da ferramenta *StArt* é que permite a organização dos dados de forma gráfica, com objetivo de entender visualmente a estrutura elaborada a partir dos termos achados.

Com o propósito de alcançar o terceiro e quarto objetivo, após a organização e preservação dos dados para a análise, efetuou-se a leitura completa de cada um dos artigos, com intuito de obter uma visão geral das informações e refletir sobre os aspectos relevantes para essa pesquisa.

Optou-se por elaborar duas tabelas contendo os dados dos artigos selecionados. Na Tabela 2, presente no Apêndice 1, foram incluídos os dados fornecidos pelo *software* da *StArt*, listando apenas o ID e o título dos artigos. Essa decisão foi tomada para simplificar e reduzir o tamanho das outras tabelas, pois muitos

⁴⁸ Notepad++ é um editor de texto e de código fonte de código aberto sob a licença GPL. Suporta várias linguagens de programação rodando sob o sistema Microsoft Windows, é distribuído como um software livre

dos títulos são extensos. Essa abordagem visa facilitar a elaboração das tabelas subsequentes.

No desenvolvimento da Tabela 3 (Apêndice 2), foram selecionados e organizados os seguintes dados: ID *StArt*, *crowdsourcing* (abordagem dos artigos relacionada a essa tipologia, considerando que o tema central é a Ciência Cidadã), projetos (indicação se os projetos estão presentes ou apenas mencionados nos artigos), armazenamento (informação sobre o armazenamento dos dados gerados dos projetos se estão presentes ou não) e por fim, observações. As observações abrangem detalhes sobre o conteúdo do artigo, se apresenta ou cita repositórios ou similares, e quais são eles. Essa estruturação tem o intuito de facilitar a seleção e identificação dos repositórios que contêm os resultados dos projetos mencionados nos artigos analisados.

Após a seleção dos dados identificados na Tabela 3, com o propósito de verificar a disponibilidade dos dados de projetos de Ciência Cidadã mencionados nos artigos, foram desenvolvidos procedimentos que seguiram algumas etapas bem definidas, com finalidade de responder os objetivos propostos.

1. Identificação dos artigos que mencionam projetos de Ciência Cidadã;
2. Identificação dos repositórios: verificar a disponibilidade dos dados e sua acessibilidade (gratuita ou mediante registro);
3. Categorização dos tipos de dados disponibilizados abrangendo dados brutos, dados processados, metadados, imagens, vídeos, entre outros;
4. Descrição dos tipos de dados: natureza dos dados, como foram coletados, a periodicidade da coleta, entre outras características;
5. Formato dos dados: em que formato estão disponíveis, como CSV, planilha Excel, texto estruturado, entre outros;
6. Acesso aos dados, se estavam disponíveis de forma livre ou restrita;
7. Detalhamento dos dados, verificação se os dados permitem replicação dos resultados e validação das conclusões;
8. Verificar se a documentação e referências foram fornecidas, permitindo que outros pesquisadores pudessem citar os dados corretamente;
9. Resultado da análise dos projetos de Ciência Cidadã armazenados em repositórios e similares, com apresentação das descobertas obtidas durante o processo de verificação.

Entende-se que no armazenamento dos dados de pesquisa originários de artigos sobre Ciência Cidadã, eles podem variar dependendo do projeto e da instituição responsável pela coleta e análise dos dados. Geralmente os dados são armazenados de forma organizada e segura para garantir sua integridade e acessibilidade. Tem-se como exemplo os bancos de dados, repositórios *online*, arquivos digitais, entre outros.

A análise dos tipos de dados disponibilizados pelos repositórios é valiosa para auxiliar outros pesquisadores na busca por fontes confiáveis de informações e contribuir para a reprodutibilidade e transparência na pesquisa científica (SILVA JUNIOR; SANTOS, 2019).

4 RESULTADOS E ANÁLISES

Para atender a proposta de pesquisa e os objetivos propostos, são apresentados os resultados e as análises dos dados encontrados, na base de dados *Scopus*. Salienta-se que o presente estudo se desenvolveu em duas etapas. Sendo que a primeira etapa foi concluída em agosto de 2022.

Os resultados da primeira etapa decorreram de uma revisão bibliográfica na identificação e análise da produção anual, das palavras-chave, dos autores e suas filiações, correlacionadas com o número de produção e países de publicação.

Para a segunda etapa, a análise ocorreu a partir dos documentos selecionados na primeira etapa, onde buscou-se identificar nas evidências (qualitativas) como os dados de pesquisa resultantes dos projetos de Ciência Cidadã estão/são disponibilizados pelos repositórios de dados, quando apresentados.

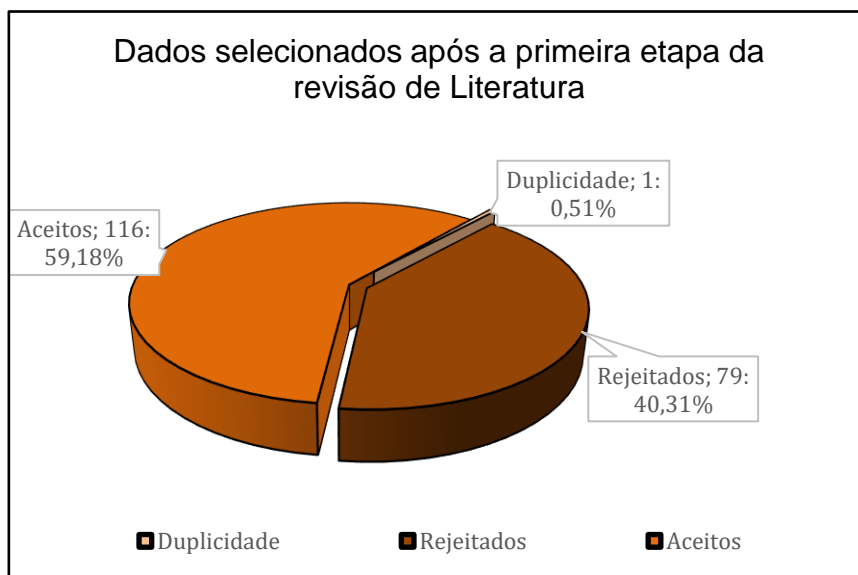
4.1 ANÁLISE DOS DADOS DE PESQUISA NO CONTEXTO DA CIÊNCIA CIDADÃ NA BASE DE DADOS SCOPUS

Como estratégia de busca, foi usada a base de dados *Scopus*, onde inicialmente recuperou-se 196 documentos que foram exportados para a ferramenta *StArt*, onde efetuou-se os procedimentos iniciais da revisão.

Nos procedimentos de coleta de dados, efetuou-se a leitura dos títulos, palavras-chave e resumos dos documentos selecionados, buscando identificar estudos ligados à Ciência Cidadã e Dados Abertos de Pesquisa, respeitando os critérios selecionados de inclusão e exclusão.

Para o processo de execução foram selecionados os 196 documentos do modelo da técnica de revisão sistemática da *StArt*, sendo que na aplicação dos critérios de inclusão e exclusão, resultou em um *corpus* de 116 documentos aptos para análise e foram descartados os 79 rejeitados e 1 (uma) duplicidade. Salienta-se que todas as exclusões foram justificadas na ferramenta *StArt* para futura verificação e foram revisados pelo professor orientador. O Gráfico 1, representa os dados selecionados na primeira etapa da revisão, em numeral e em percentagem.

Gráfico 1 - Representação dos dados selecionados após primeira etapa da Revisão bibliográfica.



Os resultados apresentados para as palavras-chave recuperadas nos 116 documentos, retornou 1637 resultados, dos quais 138 se aderem ao tema de pesquisa, sendo que dessas palavras 110 foram selecionadas para o estudo. A diferença de 28 resultados são complementos, variantes e ou não estarem ligadas diretamente ao tema, portanto, não compondo o grupo de palavras-chave de interesse.

No Quadro 5 e no Quadro 6, são apresentados os termos e campos de buscas utilizados e o número de documentos recuperados no período de 2012 a 2021, com frequência superior a duas incidências sobre a Ciência Cidadã e Dados de Pesquisa, a pesquisa limitou-se aos termos na língua inglesa.

Quadro 5: Base de dados *Scopus*: termos e campos de buscas utilizados e documentos recuperados (2012 até dez. 2021) com frequência superior a duas incidências sobre Ciência Cidadã

| Expressão de busca e suas variantes | Documentos recuperados |
|--|-------------------------------|
| <i>Citizen Science</i> (Ciência Cidadã) | 29 |
| <i>Crowdsourcing</i> | 48 |
| <i>Citizen participation</i> | 5 |
| <i>Scientific community</i> | 3 |
| <i>Crowdsourcing platforms</i> | 3 |

Fonte: Elaborado pela Autora (*Scopus* 2021)

Quadro 6 - Base de dados *Scopus*: termos e campos de buscas utilizados e documentos recuperados (2012 até 2021) com frequência superior a duas incidências sobre Dados de pesquisa

| Expressão de busca e suas variantes | Documentos recuperados |
|--|-------------------------------|
| <i>Data sharing</i> | 29 |
| <i>Research data</i> | 3 |
| <i>Open data</i> | 3 |

Fonte: Elaborado pela autora (*Scopus*, 2021)

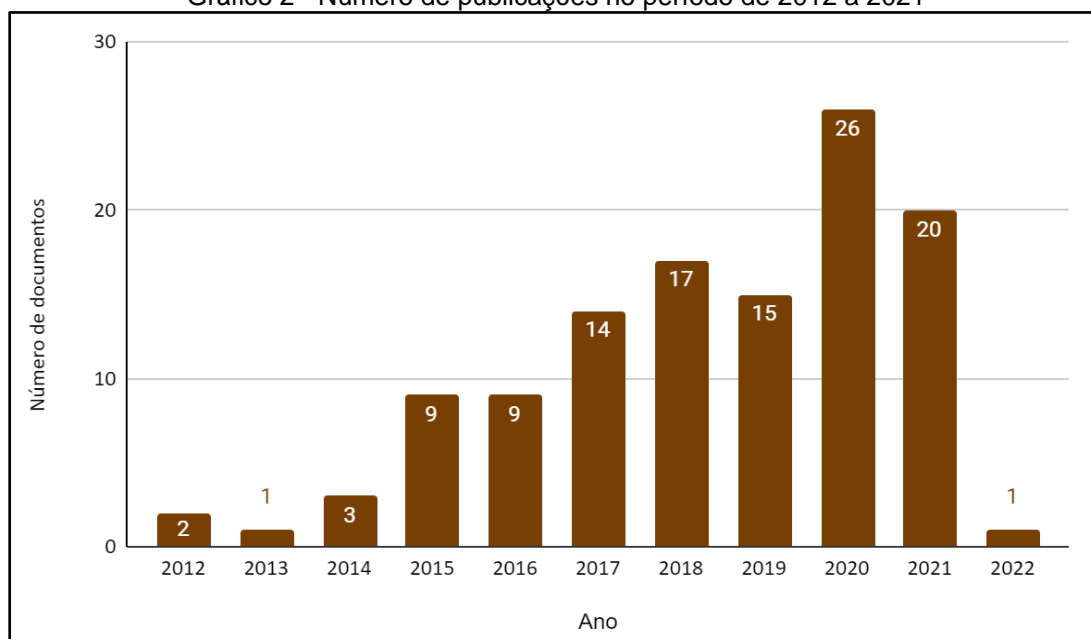
Na análise das seis palavras-chave escolhidas nas publicações que atenderam aos critérios, selecionou-se as que constavam no título ou no resumo, a que se destacou em maior número foi "*crowdsourcing*" com 48 resultados, seguida de "*data sharing*" e "*Citizen Science*" com 29 resultados cada uma, na sequência temos "*research data*" com 3 resultados, "*community science*" com um (1) resultado e por fim "*open data repository*" que não apresentou nenhum resultado nas buscas.

Observou-se que o número de recuperação do termo "*crowdsourcing*", deve-se ao fato de compor várias áreas do conhecimento, além de ser considerado uma "versão digital da Ciência Cidadã" (EKLUND; STAM; LIEBERMANN, 2019, *online*), de caráter híbrido, visto que os pesquisadores abarcam multidões na colaboração em experimentos *online*.

A análise diacrônica foi efetuada no período de 2012 até dezembro de 2021, observou-se que o número de documentos publicados apresentou diferentes variações. Sendo que entre os anos de 2012 e 2016, o número de publicações relacionadas aos cinco anos seguintes não foi expressivo. Observa-se que as publicações sobre Ciência Cidadã e abertura de dados de pesquisa tiveram um aumento considerável nos cinco anos seguintes, apresentando um aumento de 80,17% de artigos publicados, conforme demonstrado no Gráfico 2.

Justifica-se pelo fato da Ciência Cidadã, estar sendo considerada uma ciência contemporânea, pouco explorada e apesar de ainda não estar consolidada, vem ganhando confiabilidade e visibilidade, estimulada pelo envolvimento de um número maior de cientistas cidadãos, assim como a aderência dos projetos ao acesso aberto, bem como o desenvolvimento de dispositivos móveis específicos para a coleta e compartilhamento de dados, tornou-se inevitável no processo da produção científica.

Gráfico 2 - Número de publicações no período de 2012 a 2021



Fonte: Elaborado pela autora

Ainda na análise da produção anual, constatou-se uma anomalia nos dados diacrônicos do artigo “*Challenges of evidence synthesis during the 2020 COVID pandemic: a scoping review*” que em dezembro de 2021, constava na base com o ano de 2022, como *online first*, explica-se pelo fato de ser disponibilizado antes da publicação integral do fascículo. Devido a sua aderência a essa proposta de trabalho, optou-se por manter como parte do estudo.

Na Tabela 1 listam-se os autores com o número de documentos superior a duas (2) publicação sobre a temática Ciência Cidadã, abertura e compartilhamento de dados de pesquisa. Para a identificação dos autores e o número de documentos que produziram, utilizou-se a planilha *Excel* e com isso obteve-se como resultado uma lista com o nome dos autores e sua produção.

Foram identificados 599 autores, um número elevado de pesquisadores em relação ao número de documentos produzidos. Isso ocorre devido ao processo de pesquisa e produção serem compartilhados ou resultados de projetos de Ciência Cidadã que compunham inúmeros pesquisadores. Alguns artigos chegam a contar com mais de 40 autores. Tem-se como exemplo o artigo “*Crowdsourcing the General Public for Large Scale Molecular Pathology Studies in Cancer*” como 46 autores, dentre eles o brasileiro Francisco José Cândido dos Reis, da Universidade de São Paulo, que atua na área de saúde. O artigo é o resultado do projeto *Citizen Science*

Cell Slider, no compartilhamento de imagens de tumores com o público, para que identificassem marcadores tumorais de forma independente, por meio de uma interface baseada na *internet*.

Tabela 1 - Lista de autores mais produtivos

| # | Nome do autor | N. de documentos |
|----|----------------------|------------------|
| 1 | Wiggins A. | 3 |
| 2 | Preece J. | 3 |
| 3 | Zhang Y. | 2 |
| 4 | Yang H. | 2 |
| 5 | Wang Q. | 2 |
| 6 | Shaw D.L. | 2 |
| 7 | Pradhan S. | 2 |
| 8 | Ma J. | 2 |
| 9 | Lukyanenko R. | 2 |
| 10 | Li A. | 2 |
| 11 | Khan S. A. | 2 |
| 12 | Ignat T. | 2 |
| 13 | He Y. | 2 |
| 14 | Hammock J. | 2 |
| 15 | Ganzevoort W. | 2 |
| 16 | Estrada-Galiñanes V. | 2 |
| 17 | Chen Y. | 2 |
| 18 | Blumesberger S. | 2 |
| 19 | Ayris P. | 2 |

Fonte: Elaborado pela Autora (Scopus, 2021)

Das autoras, Wiggins e Preece lideraram o *ranking* com maior número de artigos publicados, cada uma com três. A autora Andrea Wiggins faz parte do quadro docente da *University of Nebraska*, em Omaha, nos Estados Unidos. Conforme a base de dados, a autora tem 50 documentos publicados, com 4.418 citações e índice-*h* de 25, identificou-se que no ano de 2021 foi citada 981 vezes; seus assuntos de pesquisa são multidisciplinares, destacando-se *Citizen Scientists* e *Crowdsourcing*.

A autora Jennifer J. Preece, é professora na *University Of Maryland*, em College Park (USA), possui doutorado pela *Open University*, do Reino Unido, sua área de interesse de pesquisa é comunicação mediada por computador, interação humano-computador, comunidades *online*, motivação para participação em mídias, comunicação móvel, entre outras. Publicou 119 documentos, com 5.939 de citações e índice-*h* de 36 e o ano em que foi mais citada foi 2017 com 453 citações.

4.2 IDENTIFICAÇÃO DE ARMAZENAMENTO DE DADOS DE PESQUISA EM PROJETOS EM CIÊNCIA CIDADÃ EM REPOSITÓRIOS

Com os resultados da primeira etapa concluídos, partiu-se para a segunda etapa de revisão com intuito de atender aos objetivos restantes. Para tanto, efetuou-se a leitura dos documentos na íntegra, assim, podendo responder ao questionamento sobre a forma de como está ocorrendo o processo de compartilhamento de dados abertos em repositórios de projetos que têm vertentes na Ciência Cidadã? Assim como o objetivo de identificar repositórios de dados nos trabalhos identificados dentro dos textos.

Após verificação dos resultados da etapa inicial, optou-se, por sugestão da banca de qualificação, por uma ampliação nos critérios de inclusão no período de publicação do ano de 2022 e nos critérios de exclusão acrescentou-se os artigos sem acesso ao texto na íntegra ou disponível *online* e em acesso aberto. A ampliação se deve a possibilidade de que se pudesse mapear e caracterizar mais dados.

Constatou-se que alguns documentos, a partir do ano de 2016, ainda não estavam anexados à base de dados *Scopus*. Houve a inclusão de 38 artigos, e desses 18, atenderam aos critérios de inclusão e 20 foram descartados, por meio dos critérios de seleção.

Com a inclusão de novos artigos o número total de documentos analisados ficou em 134, dos quais foi efetuada a leitura dos documentos na íntegra e após a aplicação dos critérios de seleção resultaram em 73 aceitos e 61 foram descartados.

Na revisão bibliográfica dos artigos selecionados após o processo de seleção, identificou-se 34 documentos que apresentam estudos e/ou projetos em Ciência Cidadã. Constatou-se que o maior enfoque era em biodiversidade ou assemelhados, com 18 trabalhos, na temática sobre saúde foram sete e nas diferentes áreas do conhecimento ou relacionado ao tema foram nove documentos e o restante dos artigos não abordam nada sobre projetos, mas apresentam estudos teóricos sobre Ciência Cidadã e/ou *crowdsourcing*.

O que se observou nos artigos, oriundos dos projetos sobre Ciência Cidadã (os projetos que divulgam que houve armazenamento), é que os dados gerados no decorrer da pesquisa são armazenados em algum tipo de repositório ou dispositivos semelhantes (na maioria dos documentos aparece somente as sugestões de onde os dados gerados poderiam ser armazenados). Os dados apresentados indicam que

ainda não está bem definida a escolha dos pesquisadores pelo local onde vão compartilhar e armazenar os dados de pesquisa gerados pelos projetos e nem o tipo de abertura que terão esses dados.

A partir da leitura dos trabalhos, identificou-se 26 artigos que abordam questões relacionadas com a Ciência Cidadã e ao *crowdsourcing*, enfatizando a importância e a necessidade de partilhar e armazenar dados resultantes de estudos teóricos em repositórios ou similares. No entanto, esses artigos não fornecem informações detalhadas sobre os tipos de armazenamento em questão, suas características específicas, para que tipos de dados sejam adequados e não especificam se o acesso a esses dados seria de forma aberta ou restrita.

Ainda neste grupo de documentos, constata-se uma preocupação com a extração e qualidade dos dados, assim como o desafio em projetos de *crowdsourcing*, no uso de diferentes tecnologias, como aplicativos móveis e o uso de inteligência artificial, além de também enfatizarem a necessidade armazenamento em bancos de dados ou no desenvolvimento de plataformas de inovação ou *software* de código aberto, adaptado às necessidades dos usuários.

Dos 73 artigos, somente 14 mencionam que seus dados são de acesso aberto, nos documentos restantes essas informações não estão explícitas.

O número de documentos que teorizam sobre projetos de Ciência Cidadã e que não apresentam projetos é de 40 trabalhos, somente citam as plataformas, repositórios, bancos de dados, *software*, entre outros, abordando a sua importância no compartilhamento e armazenamento de dados, mas não informam se armazenam, no entanto, mencionam exemplos dos mesmos, explicando para qual função poderiam utilizá-los, dentro de seus respectivos estudos. A Tabela 2, apresenta os exemplos de locais sugeridos para essa finalidade.

A Tabela 4 apresenta os depósitos de armazenamento, indicados pelo ID do artigo disponibilizado no *software* da *StArt*, a descrição do depósito e o URL. As descrições dos artigos estão referenciadas por ordem numérica nos Apêndices 1 e 2, nas Tabelas 3 e 4.

Tabela 2 - Dispositivos que armazenam dados de pesquisa de projetos em Ciência Cidadã, mencionados nos artigos analisados.

| Depósito de armazenamento | Descrição | Endereço (URL) |
|--|---|---|
| PostgreSQL ID: 19 | Sistema de banco de dados objeto-relacional de código aberto. | https://www.postgresql.org/ |
| GeoServer ID:19 | Servidor de código aberto para compartilhamento de dados geoespaciais. | https://geoserver.org/ |
| GS Cloud ID:27 | Serviço de dados de sensoriamento remoto. | www.gscloud.cn |
| GSA ID:27 | Repositório de dados para coletar, arquivar, gerenciar e compartilhar dados brutos de sequência. | https://ngdc.cncb.ac.cn/gsa/ |
| GCdataPR ID. 27 | Plataforma original de publicação de dados de pesquisa em ciências da terra. | https://www.geodoi.ac.cn/WebEn/Default.aspx . |
| Mobistudy ID. 29 | Plataforma aberta para pesquisa clínica baseada em dispositivos móveis. | https://mobistudy.org/ |
| DataSHIELD ID: 37 | Software de código aberto para a análise de preservação da privacidade de dados biomédicos. | https://www.datashield.pt/ |
| INaturalist ID: 38; 147 | Rede social <i>online</i> de pessoas que compartilham informações sobre biodiversidade para ajudar uns aos outros a aprender sobre a natureza. | https://www.inaturalist.org/ |
| Amazon Mechanical Turk ID: 39; 133, 21370 | Serviço de <i>crowdsourcing</i> . Seus usuários, como empresas e pesquisadores, contratam <i>crowdworkers</i> , trabalhadores remotos terceirizados para realizar tarefas que computadores ainda não são capazes de realizar. | https://www.mturk.com/ |
| Citsci.org ID: 42, 44 | Plataforma global de apoio à Ciência Cidadã. | www.citsci.org |
| Open Humans.org ID: 42, 75 | Comunidade <i>online</i> e infraestrutura digital que fornece suporte para auto pesquisa, ferramentas para acesso a dados pessoais e agregação de dados, entre outros. | https://www.openhumans.org/ |
| iSpotNature.org ID:52 | Plataforma de Ciência Cidadã para a biodiversidade. | https://www.ispotnature.org/ |
| AirBnB ID: 57 | Serviço <i>online</i> comunitário para as pessoas anunciarem, descobrirem e reservarem acomodações e meios de hospedagem | https://www.airbnb.com |
| TripAdvisor ID: 57 | Sítio de viagens que fornece informações e opiniões de conteúdos relacionados ao turismo. | https://www.tripadvisor.com.br/ |
| PubMed ID:59 | Base de dados <i>PubMed</i> de citações e resumos da literatura biomédica. | https://pubmed.ncbi.nlm.nih.gov/about/ |
| Scopus ID: 59 | Banco de dados de resumos e citações de artigos para jornais/revistas acadêmicos. | https://www.scopus.com/ . |
| ACM Digital Library ID: 59 | Biblioteca digital com objetivo de preservar as tradições culturais literárias do mundo. | https://asianlegacylibrary.org/ |

| Depósito de armazenamento | Descrição | Endereço (URL) |
|-----------------------------|---|---|
| Web of Science ID: 59 | <i>Site</i> de acesso baseado em assinatura a vários bancos de dados que fornecem dados abrangentes de citações para diferentes disciplinas acadêmicas. | https://www.webofscience.com |
| IEEE Xplore ID: 59 | Banco de dados de pesquisa para descoberta e acesso a artigos de periódicos, anais de conferências, normas técnicas e materiais relacionados à ciência da computação, engenharia elétrica e eletrônica e áreas afins. | https://ieeexplore.ieee.org/ |
| Hubzero ID: 59 | Plataforma de <i>software</i> de código aberto para construção de sites que suportam atividades científicas. | https://hubzero.org/ |
| Zooniverse ID: 59; 133 | Portal de Ciência Cidadã de propriedade e operado pela " <i>Citizen Science Alliance</i> ". É onde estão os mais populares e bem-sucedidos projetos de Ciência Cidadã da <i>internet</i> . | www.zooniverse.org |
| PatientsLikeme ID: 59 | Comunidade integrada, gerenciamento de saúde e plataforma de dados do mundo real. | https://www.patientslikeme.com/ . |
| MongoDB ID: 65 | <i>Software</i> de banco de dados orientado a documentos livre, de código aberto e multiplataforma, escrito na linguagem C++. | https://www.mongodb.com/ |
| Open SNP ID: 71 | <i>Site</i> de código aberto onde os usuários podem compartilhar suas informações genéticas. | https://opensnp.org/ |
| Naveed Aziz/ CGEn ID: 71 | Plataforma nacional financiada pelo governo federal para sequenciamento e análise do genoma. | https://www.cgen.ca/ |
| MySQL.com ID: 76; 107 | Sistema de gerenciamento de banco de dados, que utiliza a linguagem SQL como interface. | https://www.mysql.com/ |
| Carto.com ID: 93 | <i>Software</i> GIS e inteligência de localização nativa da nuvem para analistas, cientistas de dados e desenvolvedores que resolvem problemas geoespaciais com pilhas de análises modernas. | https://carto.com/ |
| OAPEN ID: 108 | Uma monografia de acesso aberto, acadêmica que é disponibilizada abertamente <i>online</i> , com licença aberta. | https://www.oapen.org/ |
| JSTOR ID: 108 | Sistema <i>online</i> de arquivamento de revistas acadêmicas com sede nos EUA. | https://www.jstor.org/ |
| Wordpress ID: 112 | Sistema livre e aberto de gestão de conteúdo para <i>internet</i> , baseado em PHP com banco de dados MySQL | https://wordpress.com/pt-br/ |
| Omeka ID: 112 | <i>Software</i> livre e de código aberto, para gerenciamento de conteúdo de coleções digitais. | https://omeka.org/ |
| Datahub ID: 118 | Datahub é a maneira mais rápida para indivíduos, equipes e organizações publicarem, implantarem e compartilharem seus dados. | https://datahub.io/ |
| CUAHSI HIS ID: 120 | Sistema de Informação Hidrológica CUAHSI (HIS) é um sistema baseado na Internet para compartilhar dados hidrológicos. | https://his.cuahsi.org/ |

| Depósito de armazenamento | Descrição | Endereço (URL) |
|---|--|--|
| LimeSurvey ID: 122 | <i>Software</i> livre para aplicação de questionários <i>online</i> escrito em PHP, podendo utilizar bancos de dados MySQL, PostgreSQL ou Microsoft SQL. | https://www.limesurvey.org/pt-br |
| Ushahidi, ID: 130 | <i>Software</i> livre e " <i>open-source</i> " para a coleta de informações, visualização e mapeamento interativo. | https://www.usahidi.com/ |
| GitHub ID: 130; 141, 21370, 21385 | Plataforma de desenvolvedor com tecnologia de IA para criar, dimensionar e fornecer <i>software</i> seguro. | https://github.com/ https://github.com/hctshra/ecpo-segment |
| LinkedIn ID: 136 | Plataforma de mídia social focada em negócios e emprego que funciona através de <i>sites</i> e aplicativos móveis. | https://www.linkedin.com/ |
| GenBank ID: 141 | Banco de dados de anotações de sequências de nucleotídeos publicamente disponíveis e suas traduções de proteínas. | https://www.ncbi.nlm.nih.gov/genbank/ |
| ArcGIS ID: 143 | <i>Software</i> de sistema de informação geográfica cliente, servidor e <i>online</i> desenvolvido e mantido pela Esri. | https://www.arcgis.com/ |
| OpenKnowledge Foundation ID: 149 | O principal sistema de gerenciamento de dados de código aberto do mundo. | https://okfn.org/ |
| Apple Researchkit. ID:152 | <i>Software</i> para aplicativos que permite que pesquisadores médicos coletem dados robustos e significativos. | https://www.apple.com/lae/researchkit |
| Sage Bionetworks ID: 152 | Plataforma de tecnologias escaláveis e portáteis que facilitam dados, integração, validação, harmonização, compartilhamento, exploração e análise distribuída em ambientes de computação em nuvem. | https://sagebionetworks.org/ |
| IPBES ID: 165 | Plataforma Intergovernamental de Políticas Científicas sobre Biodiversidade e Serviços Ecossistêmicos, criada para melhorar a interface entre ciência e política em questões de biodiversidade e serviços ecossistêmicos. | https://www.ipbes.net/ |
| Scholarometer. ID: 183 | Plataforma social que estende o <i>Google Scholar</i> para facilitar a análise de citações acadêmicas. | https://scholarometer.indiana.edu/ |
| Zenodo ID: 21364; 21366 | Repositório aberto de uso geral desenvolvido sob o programa europeu OpenAIRE e operado pelo CERN. Ele permite que os pesquisadores depositem trabalhos de pesquisa, conjuntos de dados, <i>software</i> de pesquisa, relatórios e quaisquer outros artefatos digitais relacionados à pesquisa. | https://zenodo.org/ |
| Spipoll ID: 21365; 21392 | Plataforma <i>online</i> de compartilhamento de dados. | https://www.spipoll.org |
| Global Biodiversity Information Facility GBIF ID: 21382 /21366 | Rede internacional e infraestrutura de dados financiada pelos governos do mundo e destinada a fornecer a qualquer pessoa, em qualquer lugar, acesso aberto a dados sobre todos os tipos de vida na Terra. | https://www.gbif.org/ |

| Depósito de armazenamento | Descrição | Endereço |
|---|---|---|
| Dryad ID: 21366 | Plataforma de publicação de dados abertos e uma comunidade comprometida com a disponibilidade aberta e a reutilização rotineira de todos os dados de pesquisa. | https://datadryad.org/stash |
| Open Science Framework ID: 21366 | Aplicativo da <i>Web</i> gratuito e de código aberto que conecta e oferece suporte ao fluxo de trabalho de pesquisa, permitindo que os cientistas aumentem a eficiência e a eficácia de suas pesquisas. | https://osf.io/ |
| Prolific ID: 21370 | <i>Site</i> onde se é pago por participar de pesquisas, geralmente acadêmicas ou de cunho comercial. | https://www.prolific.co/ |
| AMUNATCOLL ID: 21376 | Banco de dados contém dados digitais sobre espécimes, amostras biológicas, fontes bibliográficas e documentos de natureza multimídia. | https://amunatcoll.pl/ |
| BioCASE Provider Software ID: 21376 | Pacote de <i>software</i> a ser instalado no servidor <i>web</i> do provedor de dados. | https://www.biocase.org/products/provider_software/ |
| IRIS ID: 21377 | Maior instalação do mundo para arquivamento, curadoria e distribuição de dados e produtos de dados sísmológicos e de outras pesquisas geofísicas. | http://service.iris.edu/fdsnws/dataselect/1/ |
| European Alien Species Information Network ID: 21382 | Sistema de informação que facilita a implementação do Regulamento da UE sobre Espécies Exóticas Invasoras | https://easin.jrc.ec.europa.eu/easin/ |

Fonte: Da Autora (2023)

As informações apresentadas na Tabela 4 mostram a existência de uma variedade de opções de armazenamento dos resultados de dados de pesquisa gerado por estudo e projetos em Ciência Cidadã. Essa diversidade de possibilidade reflete o esforço contínuo da comunidade científica para promover a transparência e abertura na pesquisa. A escolha da melhor opção dependerá das políticas institucionais, das características dos dados e das questões legais envolvidas nos projetos de Ciência Cidadã.

4.2.1 Análises dos compartilhamentos de dados em repositórios

Na linha de projetos de Ciência Cidadã, foi possível identificar armazenamentos de dados de pesquisa, seja parcial ou na sua totalidade. Foram identificados sete projetos de Ciência Cidadã. Salienta-se que os projetos identificados, estão com seus dados armazenados em repositórios de dados, plataformas e banco de dados. Verificou-se, também, que havia um estudo de caso de Ciência Cidadã, que teve seus resultados de pesquisa armazenados e foi possível examinar os dados brutos.

4.2.1.1 RDR - Research Data Repository (ID 13⁴⁹)

Nos estudos “*Built to last! Embedding open science principles and practice into European universities*” de Ignat e Ayrís (2021), dissertam sobre a necessidade de mudanças culturais em universidades, por meio da ciência aberta e da ciência cidadã, pois entendem que na atualidade os pesquisadores estão sempre competindo por mais publicações visibilidade e ganho nas avaliações comparativas onde o sucesso é medido por outras pessoas.

Dentro desse contexto a UCL - *University College London*, percebeu a necessidade de terem um repositório de dados abertos, onde as informações fossem encontradas de formas acessíveis, inoperáveis e reutilizadas. Para tanto, lançaram em 2019, conforme Figura 11, o RDR - *Research Data Repository*⁵⁰, como maneira de estabelecer a colaboração em pesquisa, o armazenamento e o compartilhamento de seus dados de pesquisa com o restante do mundo.

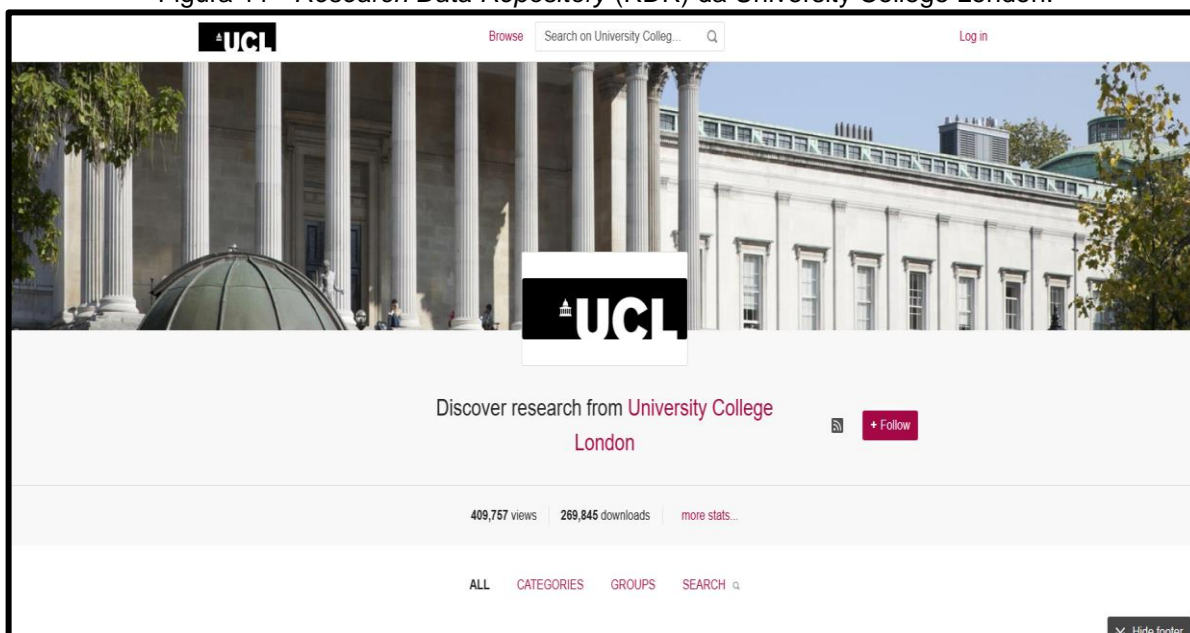
Para exemplificar a importância do armazenamento de dados de pesquisas gerados em projetos, artigo de Ignat e Ayrís (2021), apresenta um estudo de caso sobre o envolvimento de cidadãos e seu aprendizado em projetos de Ciência Cidadã, bem como a qualidade dos dados gerados pela coleta, são elementos importantes a serem gerenciados e que com os programas de treinamento correto, os dados de pesquisa coletados apresentarão o nível de qualidade esperado.

Embora o artigo não apresente seus dados armazenados no repositório criado, ainda assim, optou-se por efetuar uma análise de como estão depositados no repositório e também expor um projeto de escolha aleatória em Ciência Cidadã, para verificar quais os dados que são armazenados.

⁴⁹ *Built to last! Embedding open science principles and practice into European universities.*

⁵⁰ <https://rdr.ucl.ac.uk/>

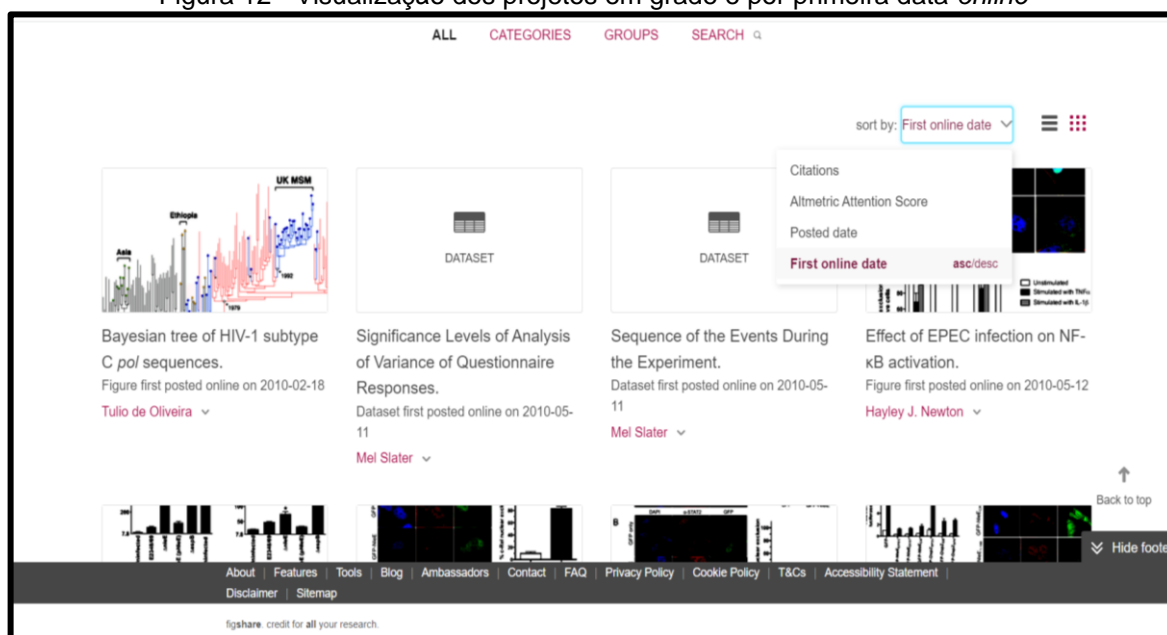
Figura 11 - *Research Data Repository (RDR)* da University College London.



Fonte: UCL - <https://rdr.ucl.ac.uk/>

Na página do RDR é possível, navegar e ver o número de visualizações e *downloads* e a possibilidade de verificar as métricas clicando em mais estatísticas, onde são apresentadas as visualizações por período (gráfico), sendo atualizadas diariamente e quando é utilizado o filtro se pode escolher o grupo de Instituições/Faculdade, o tipo de item (figuras, pôster, conjuntos de dados, entre outros), o período específico que se quer verificar e a atividade total por país, em percentual, podendo ser visto em mapa ou mesa.

Na Figura 12 é possível visualizar todos os projetos armazenados, com opção de visualização em grade ou lista, assim como acesso à opção de ordenação por data de envio, onde se tem a escolha de ver os projetos que tiveram citações, pontuação de atenção altimétrica e data do primeiro envio. Ao efetuar uma dessas escolhas, vai remeter aos títulos dos projetos conforme ordenamento e ao acessar o projeto é direcionado a um *hiperlink*, onde vão estar armazenados todos os dados do projeto.

Figura 12 - Visualização dos projetos em grade e por primeira data *online*

Fonte: UCL - <https://rdr.ucl.ac.uk/>

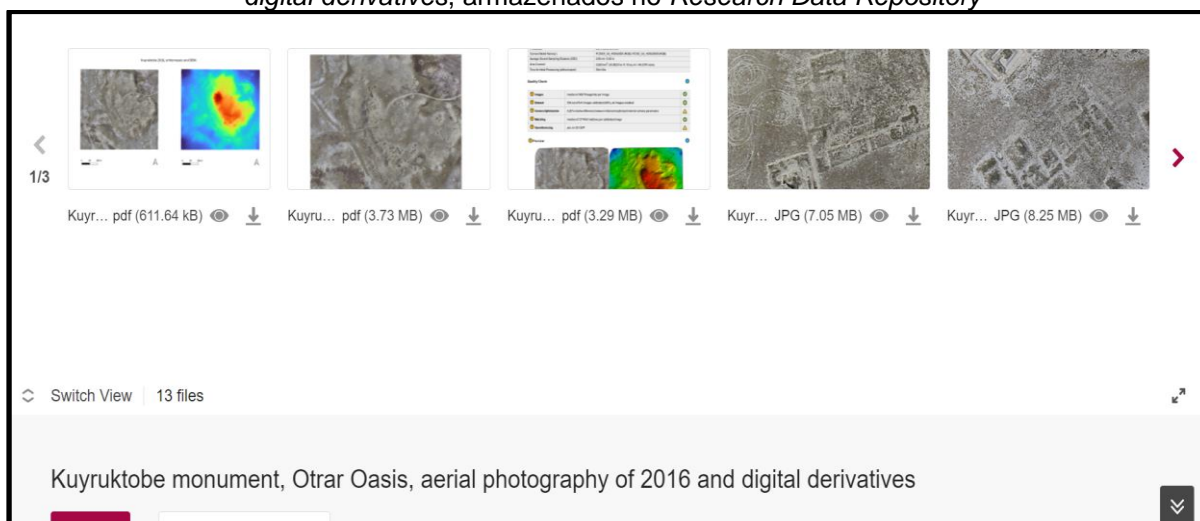
Outros dados que são apresentados na página são as categorias dos projetos, isto é, agrupamento por áreas de conhecimento (Arqueologia, Astronomia, Biologia, Bioinformática, entre outras), os grupos (Instituições, Faculdades, Departamentos e outros) e opção de pesquisa que também pode ser efetuada por data, grupo, categoria ou licença. O que se observa é que por meio da página inicial é permitido acessar os projetos disponibilizados de diferentes formas.

Todas as informações dos projetos e informações que constam no final da primeira página como sobre, características, ferramentas, blog, embaixadores, política de cookies, declaração de acessibilidade, isenção de responsabilidade, direciona para o *hiperlink figshare*⁵¹ e o contato, perguntas frequentes, políticas de privacidade, termos e condições e o mapa do *site* encaminha para o *site* da Universidade.

Para a análise dos dados distribuídos no repositório RDR escolheu-se um projeto da área de conhecimento da Arqueologia, armazenado em 27/04/2023, da autoria de Gai Jorayev, com o título de *Kuyruktobe monument, Otrar Oasis, aerial photography of 2016 and digital derivatives* -, os dados coletados foram fotografias aéreas do Sítio Arqueológico, datado entre os séculos I e X d.C. que ainda apresentam vestígios urbanos de Otrar, supõe-se que seja da cidade medieval de Keder. A Figura 13 apresenta imagens do sítio arqueológico em referência.

⁵¹ <https://knowledge.figshare.com/about>

Figura 13 - Dados do Projeto *Kuyruktobe monument, Otrar Oasis, aerial photography of 2016 and digital derivatives*, armazenados no *Research Data Repository*



Fonte: *Research Data Repository*

Conforme consta no repositório os tipos de dados apresentados são *Dataset*, de dados brutos, que vem em forma de planilha e estão acessíveis; constam 13 arquivos acessíveis, assim como fazer *download* (7,89GB). As informações que são apresentadas no repositório são: as coordenadas geográficas do sítio, o ID, o nome do responsável por tirar as fotos a data em que foram tiradas e os instituições parceiras, a data de publicação dos dados processados, a categoria da área de conhecimento, palavras-chave, tipo de licença, métricas de uso, tipos e exportação que é possível efetuar: RefWorks, BibTex, Ref. manager, Endnote, DataCite, NLM, DC.

O conjunto de dados que são disponibilizados⁵² são: 545 fotografias; dois relatórios de processamento em formato PDF explicando os detalhes do processamento e geração dos derivados em dois pacotes de *software* diferentes; dois Ortomosaicos Digitais e dois Modelos Espaciais Digitais em formato TIFF para visualização e análise em *software* SIG; um conjunto foi processado em 2021 e outro em 2023; arquivo PDF para visualização mostrando os resultados do processamento de 2021; malha 3D em formato OBJ e textura como arquivo Zip, para ser baixado e

52

https://rdr.ucl.ac.uk/articles/dataset/Kuyruktobe_monument_Otrar_Oasis_aerial_photography_of_2016_and_digital_derivatives/22701175

visualizado com *software* 3D como o MeshLab. A malha 3D também é disponibilizada em plataforma externa (Sketchfab⁵³).

4.2.1.2 MFN Data Repository (ID 16⁵⁴)

O artigo “*A self-assessment of european citizen science projects on their contribution to the UN sustainable development goals (SDGs)*,” de Moczek e outros (2021), disserta sobre pesquisa efetuada acerca das avaliações na contribuição da Ciência Cidadã para os Objetivos de Desenvolvimento Sustentável (ODS) por participantes do projeto europeu, tendo como foco três dos ODS: boa saúde e bem-estar, educação de qualidade e vida na Terra. O foco da avaliação dos projetos europeus do estudo está ligado ao monitoramento hidrológico, voltado para áreas montanhosas remotas como o Nepal e Peru e para a zona rural do Quênia, entre outros locais, com monitoramento participativo. O desafio apontado é na tomada de decisões e gestão nos resultados dos dados coletados pelos participantes nos projetos.

No estudo, destacou-se a importância na análise rigorosa da qualidade dos dados extraídos para comprovação da robustez científica. Para tanto, optaram por adotar protocolos rigorosos, como amostragem repetida em intervalos predeterminados, estratégias aprimoradas para reduzir vieses espaciais, uso de questionários e jogos para avaliar a habilidade do observador e ferramentas para inclusão de dados sobre a qualidade do observador no banco de dados. Os resultados dos dados brutos estão disponíveis no repositório do Museum für Naturkunde Berlin⁵⁵, como exposto na Figura 14.

Figura 14 - Imagem do Repositório do Museum für Naturkunde Berlin, já na página do artigo onde estão os resultados dos dados coletados.

⁵³ <https://skfb.ly/ops6F>

⁵⁴ *A self-assessment of european citizen science projects on their contribution to the UN sustainable development goals (SDGs)*.

⁵⁵ <https://doi.naturkundemuseum.berlin/data/10.7479/mzy9-wp90>

MfN Data Repository

DOI: 10.7479/mzy9-wp90

Collection: A self-assessment of European Citizen Science projects on their contribution to the UN Sustainable Development Goals (SDGs)

| | | |
|----------------------|---|-------------|
| doi: | 10.7479/mzy9-wp90 | How to cite |
| author(s): | Moczek, N. | Share: |
| date of publication: | 2020 | |
| abstract: | <p>We launched and distributed a survey across European citizen science networks in summer 2020 via the online survey tool www.soscisurvey.de/. We presented an English and a German version. The survey was developed in preparation for the Citizen Science SDG conference 'Knowledge for Change: A Decade of Citizen Science (2020-2030) in Support of the SDGs', organised by the Museum für Naturkunde Berlin (Germany) on 14-15 October 2020, as an official event of Germany's 2020 EU Council presidency. The questionnaire was open from July, 29 until October 12, 2020. In total, we received 195 responses (see raw dataset). Due to a lack of consent to data storage (n = 2) or incomplete records (n = 68), 125 responses could ultimately be included in the analysis. All survey data were analysed through simple descriptive statistics in order to summarise and combine the collected</p> | |

Fonte: Museum für Naturkunde Berlin

Na página constam as informações disponibilizadas no repositório como o título, o DOI, os autores, a data de publicação, a licença e um resumo descrevendo a ferramenta *SoSci Survey*⁵⁶ *online* utilizada para o questionário da pesquisa, em língua inglesa e alemã, para a Conferência *Citizen Science SDG 'Knowledge for Change: A Decade of Citizen Science (2020-2030)*, em 2020.

O conjunto de dados brutos, valores e variáveis são em formato CSV e os dados compilados e limpos em arquivo XLSX. Os dados disponibilizados para análise foram: dados brutos (CSV, N=195), folha Excel com avaliações (n=125), lista de variáveis, lista de valores e questionário. Da coleção de dados analisados é possível efetuar *download* de todos eles.

No que se refere ao questionário disponibilizado, o mesmo esteve aberto para respostas no período de 29/07/2020 até 12/10/2020 onde foram avaliados e foram recebidos 195 retornos. Os dados de pesquisa foram analisados por meio de estatística descritiva simples, como forma de resumir e combinar as informações coletadas, que constam na planilha *Excel* disponibilizada no repositório.

As respostas do questionário foram verificadas por meio de análise de conteúdo simples, onde foram categorizados e quantificados. Obtendo como resultados: duas (2) respostas não puderam ser armazenadas devido à falta de consentimento, assim como 68 registros estavam incompletos, restando somente 125 registros para a análise.

⁵⁶ <https://www.soscisurvey.de/>

4.2.1.3 Smithsonian Research Online⁵⁷ e Plataforma Dryad⁵⁸ (ID.17⁵⁹)

Takashina e Bursar (2020), no documento “*Deloping generalized sampling schemes with known error properties: the case of a moving observer*”, apresentam projeto *Forest Dynamics Plot da Ilha de Barro Colorado*, sobre o levantamento de informações na área de ecologia, onde os cientistas cidadãos efetuam observações de campo, com objetivo de apresentar um método para quantificar a capacidade de cada observador e apresentar um protocolo de levantamento do que foi utilizado nas amostragens, incluindo seu grau de precisão e de erro.

As pesquisas sobre a Ilha de Barro Colorado ocorrem desde 1923, pois trata-se de floresta tropical úmida decídua e sua flora é uma das mais conhecidas no mundo, portanto, tornando-se foco de inúmeras pesquisas. Dados referentes ao local, constaram que tem uma altitude de 120m, precipitação média anual de 2580,691 mm e um déficit de humidade acumulado médio anual de -513,99951 mm, o que pode impactar nos dados finais de pesquisa.

Na Página do *ForestGeo Network*⁶⁰ (observatório da *Smithsonian*), sobre a flora, é apresentado os dados de uma forma simples, onde é informado, o período estabelecido, rede associada e Zona Ecológica (tipo de vegetação), os números de espécies (325), de hastes (871,730), de árvores (423,472) e número de parcelas (1) e os dados históricos, descritos na Figura 15. Ao clicar na zona ecológica é remetido a um *link* e locais com o mesmo tipo de vegetação e ao clicar nos dados históricos é apresentado os dados das espécies, árvores e hastes por ano de 1981 até 2016.

⁵⁷ Dados do projeto também estão disponíveis no *site*.

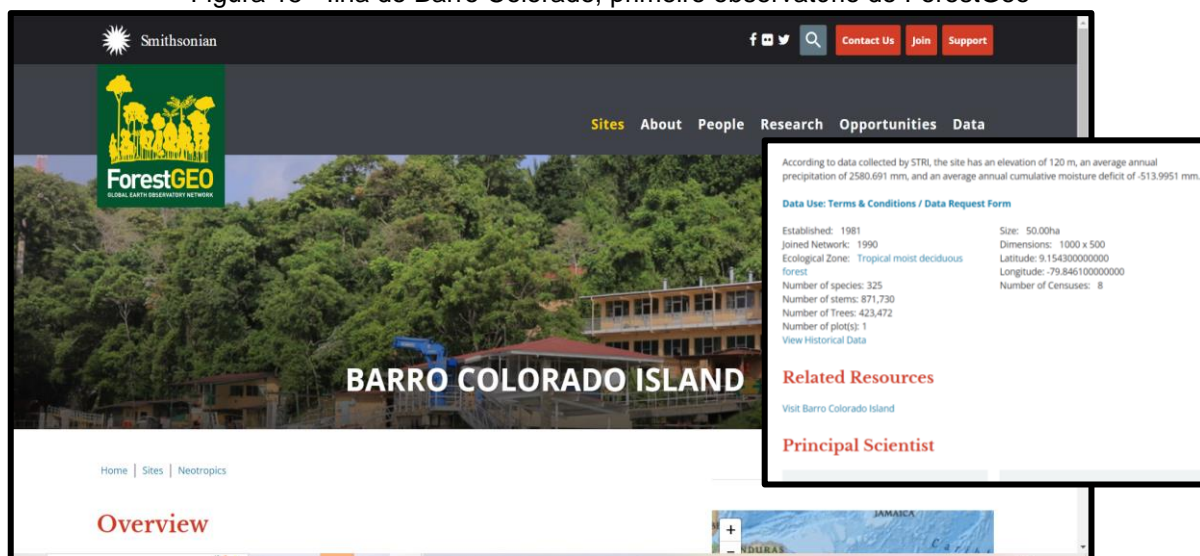
<https://datadryad.org/stash/dataset/doi:10.15146/5xcp-0d46>.

⁵⁸ <https://datadryad.org/stash/dataset/doi:10.15146/R3FH61>

⁵⁹ *Developing generalized sampling schemes with known error properties: the case of a moving observer*.

⁶⁰ <https://forestgeo.si.edu/sites/neotropics/barro-colorado-island>

Figura 15 - Ilha do Barro Colorado, primeiro observatório do ForestGeo



Fonte: ForestGeo Network

O conjunto de dados gerados no projeto estão armazenados no *Smithsonian Research Online*⁶¹, onde constam: título do projeto, nome dos autores, as datas de acesso e de disponibilidade, as coordenadas geográficas, o *link* do *site* onde está disponível e descrição da amostra, com os dados completos para os seis Censos concluídos: 1981-83, 1985, 1990, 1995, 2000 e 2005, do Lote BCI 50 ha na *Barro Colorado Island*, no Panamá, assim como o número de todas as espécies, que foram marcadas, medidas e mapeadas, que consistem em: seis arquivos de censo com as medições de localização e diâmetro, 6 arquivos de caule múltiplo com medições de caule múltiplo excluindo a maior medição que está nos arquivos de censo, a lista de espécies, dados de elevação, e um arquivo do MS *Word* explicando brevemente os campos nos arquivos do censo. Conforme observa-se na Figura 16.

Constam também os apoiadores do projeto: o US National Science Foundation, da John D. and Catherine T. MacArthur Foundation, o Smithsonian Tropical Research Institute e o BCI Forest Dynamics Plot, fazem parte do Centro de Ciências Florestais Tropicais. A saber, a estação de pesquisa de *Barro Colorado Island*, no Panamá, é administrada, desde 1946, pelo Smithsonian Tropical Research Institute. Trata-se de uma reserva natural, tomada em 1923, pelo governo dos Estados Unidos.

⁶¹ <https://repository.si.edu/handle/10088/784>

Figura 16 - *Smithsonian Research Online* - BCI 50 ha Plot 1982-2005 Census Data

The screenshot shows the Smithsonian Research Online interface. At the top, there is a navigation bar with the Smithsonian logo and the text "Smithsonian Research Online". Below this, a breadcrumb trail reads: "Página inicial → Smithsonian Tropical Research Institute → Center for Tropical Forest Science → CTFS BCI 2005 Census → Ver item".

The main content area is titled "BCI 50 ha Plot 1982-2005 Census Data" with authors "Condit, Richard; Hubbell, Stephen; Foster, Robin". The date is listed as "Data: 2006-09-21".

The "Resumo:" section contains the following text: "The 50 ha plot at Barro Colorado Island, Panama (utm: easting 625754, northing 1011569, zone 17), has 6 completed censuses: 1981-83, 1985, 1990, 1995, 2000, and 2005. All trees ≥ 10 mm dbh were tagged, measured, mapped and identified to species. The tab-delimited text files consist of the following datasets: 6 census files with the location and diameter measurements, 6 multiple-stem files with the multiple stem measurements excluding the largest measurement which is in the census files, the species list, elevation data, and an MS Word file briefly explaining the fields in the census files."

Below the summary, there is a link "Mostrar registro completo".

The "Arquivos deste item" section shows a file icon and the following details: "Nome: mult1.txt", "Tamanho: 333,8Kb", and "Formato: Text file". A link "Visualizar/Abriu" is also present.

On the right side, there is a "Buscar DSpace" search box with a search button and radio buttons for "Buscar DSpace" (selected) and "Esta coleção". Below this is a "Navegar" section with a tree view of the repository structure, including "Comunidades e Coleções", "Por data do documento", "Autores", "Títulos", "Assuntos", and "Esta coleção" with sub-items "Por data do documento", "Autores", "Títulos", and "Assuntos". At the bottom right, there is a "Minha conta" section with an "Entrar" link.

Fonte: *Smithsonian Research Online*

O conjunto de dados brutos, exibem a sua extensão em *bytes*, dados *Dataset*, dados dos textos em formato simples e em arquivo MsWord, na língua inglesa, localidade, nome do projeto, local de captura de dados, 18 arquivos, cada um como diferentes nomes, tamanho e formato. No documento MsWord é possível ler os registros de cada censo, como por exemplo que no ano inicial do Censo constava uma determinada espécie e no segundo censo essa espécie ainda se encontrava no local discriminado nos documentos ou não.

Na atualização de dados do projeto, apresentada pelos pesquisadores principais Condit e Hubbell, em 2019, na Plataforma Dryad, consta que o lote de 50Ha de Barro Colorado foi totalmente censado sete vezes a partir de 1981, a cada 15 anos. Inclui-se na pesquisa a medição de cada caule ao longo de 8 censos, com resultados em tabelas analíticas oito R na *árvore* formato (compactada em um único arquivo (bci,tree.zip)) e oito R de cada *tronco* (compactado em bci.stem.zip). Tabelas adicionais de todas as hastes individuais, medições em várias posições em algumas hastes e o banco de dados completo no formato MySQL original, compactado como condit_bci_23May2019.zip. A versão apresentada pelos pesquisadores em 2019, substitui o arquivo *Smithsonian* de 2012.

Consta ainda, das tabelas analíticas: 16 tabelas no formato R, oito tabelas no formato *árvore* e oito tabelas no formato *tronco*. Cada tabela *árvore* é incluído um registro para cada *árvore* do lote e para cada tabela de *tronco* uma linha para cada *tronco* já recenseado. Outras tabelas que constam nos resultados: uma tabela de texto

ASCII delimitada por tabulações, BCLelev.tsv, com topografia precisa do lote BCI de 50 há; uma tabela de texto delimitada por tabulações, FullMeasurementBCI.tsv, compactada como FullMeasurementBCI.zip.

Dados de extensão de arquivo: CSV (21), TXT (20), fecho éclair (9), xlsx (8), r (5), pdf (4), docx (3), md (3), tsv (3), terceiro (2), aba (2), xls (2), bash (1), mas (1), jejum (1), rápido (1), py (1), rar (1), rda (1), rda (1), dados (1), rtf (1), árvore (1) e wav (1).

Outros dados apresentados: título “*Complete data from the Varro Colorado 50-há plot: 423617 trees, 35 years*”, autores do estudo, DOI, data publicação: 2019, publicado por Dryad, resumo, notas de uso (descrição dos dados das tabelas analíticas, de espécies, de medição completa), códigos de *status* para análises, topografia do lote BCI, banco de dados MySQL completo, financiadores (National Science Foundation, John D. and Catherine T. MacArthur Foundation e Smithsonian tropical Research Institute), localização (latitude e longitude), arquivo de dados para baixar (de maio de 2019), as métricas (5411 visualizações, 1169 downloads e 17 citações), licença Universal (CC0 1.0) e *Public Domain Dedication* CC0 1.0.

4.2.1.4 Oceans 3.0 Data Portal (ID. 67⁶²)

No artigo de Wolf e outros autores (2019), “*Best practices in data management at ocean networks Canada: A citizen scientist case study*”, discorrem sobre a *Ocean Networks Canada* (ONC), que é uma instalação de observatórios oceânica, líder mundial e serviços de repositório de dados, de propriedade da *University of Victoria* (UVic). Em seu documento, apresenta os padrões da comunidade de dados de pesquisa e as melhores práticas em gerenciamento de dados, assim como a necessidade de atender aos rigorosos critérios de certificação de repositório de dados, enfatizando a importância dos dados FAIR (encontráveis, acessíveis, interoperáveis e reutilizáveis).

Para demonstrar as melhores práticas em gerenciamento de dados da ONC, criou o *Salish Sea Marine Survival Project*⁶³ da *Pacific Salmon Foundation*, em 2015, onde desenvolveram um canal de aquisição de dados e gerenciamento automatizado de dados para apoiar parcerias com programas de Ciência Cidadã. As embarcações

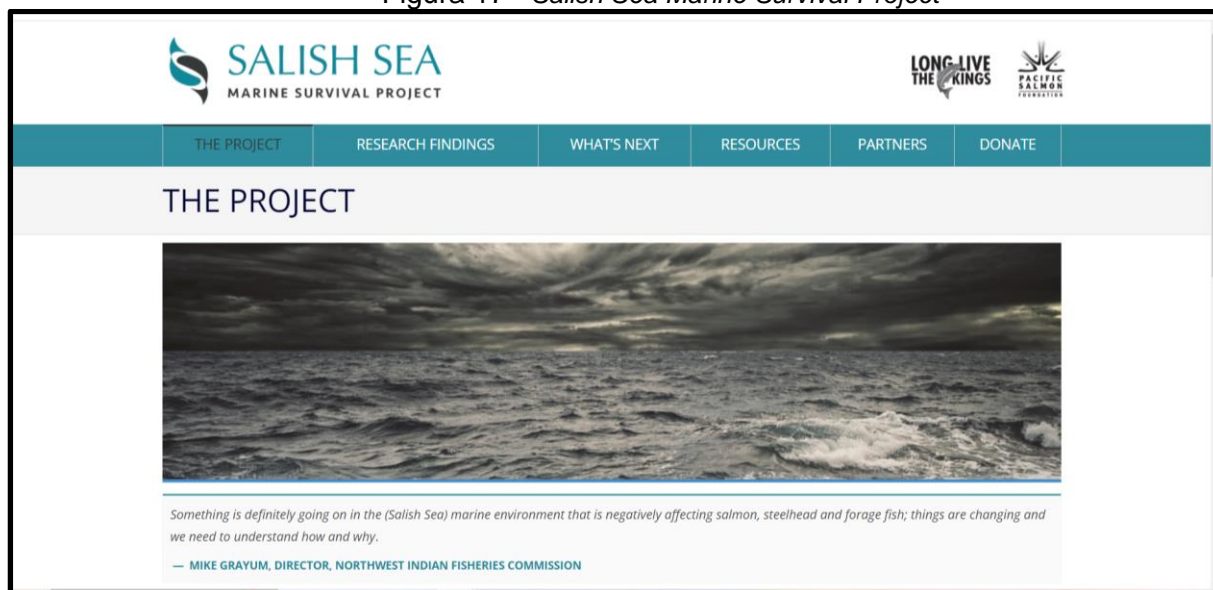
⁶² *Best practices in data management at ocean networks Canada: A citizen scientist case study.*

⁶³ <https://marinesurvivalproject.com/>

de pesca são pilotadas por aposentados ou cientistas cidadãos na coleta de dados oceanográficos durante a primavera.

A Figura 17, mostra o *site* do projeto, onde explicam os objetivos do projeto e a importância da pesquisa colaborativa, os relatórios dos resultados de pesquisas, que podem ser efetuados *download*, publicações, planos de pesquisa, pôsteres de apresentações e os parceiros.

Figura 17 - *Salish Sea Marine Survival Project*



Fonte: *Salish Sea Marine Survival Project*

Tem-se como as melhores práticas de gerenciamento de dados os oito componentes do ciclo de vida dos mesmos: planejar, coletar, garantir, descrever, preservar, descobrir, integrar e analisar. Para tanto, foi desenvolvido um aplicativo móvel que faz a transferência dos dados sem fio dos instrumentos dos cientistas cidadãos para a *Ocean*, garantindo assim a qualidade e consistência dos dados. Os dados brutos são transferidos para os servidores da ONC para serem processados, arquivados e distribuídos de forma automatizada e imediata.

Nos registros dos metadados estão as calibrações, as datas de implantação, as estações de fundição, o histórico do instrumento, as atribuições de colaboradores dos conjuntos de dados, o mapeamento de vocabulário controlado, entre outros.

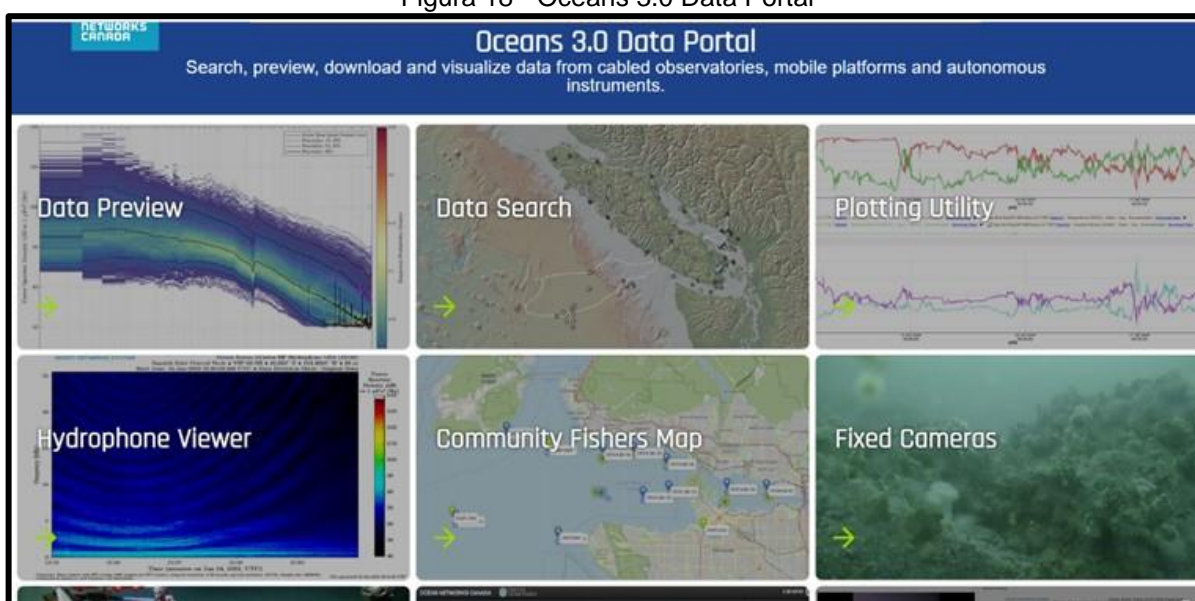
Os metadados são armazenados em um banco de dados PostgreSQL⁶⁴, que é acessível, por meio da interface *Oceans 2.0*. Ao chegar na informação disponibilizado

⁶⁴ <https://www.postgresql.org/>

da interface⁶⁵, é direcionada a um outro *hiperlink*, o Oceans 3.0 Data Portal, de acordo com a Figura 18.

No Portal de dados de pesquisa Oceans 3.0⁶⁶, é possível visualizar a grandeza e dimensão do projeto e as diferentes formas de atuação, nos oceanos, assim como os tipos de dados que são coletados. Existem 12 “janelas” de acesso: visualização de dados, dados de pesquisa, utilitário de plotagem, visualizador de hidrofone, mapa de pescadores comunitários, câmeras fixas, SeaTube V3 (dados de vídeos), pescadores digitais, painéis, API de serviços da web, OpeNDAP e o menu legado.

Figura 18 - Oceans 3.0 Data Portal



Fonte: Portal de dados de pesquisa PostgreSQL - <https://data.oceannetworks.ca/home>

Ao entrar em uma das 12 opções disponibilizadas percebe que os dados são alimentados *online*, em tempo quase real pelos cientistas cidadãos. Para cada uma das opções, é necessário fazer a seleção das fontes de dados, podendo ordenar por instrumentos por localização, instrumento de categoria e propriedade de localização, tendo como opção a optar pela localização de um dos oceanos (Ártico, Atlântico e Pacífico), após essa escolha, novamente deve-se escolher por qual baía, estreito, enseada golfo, etc., assim como a plataforma de onde os dados coletados serão analisados.

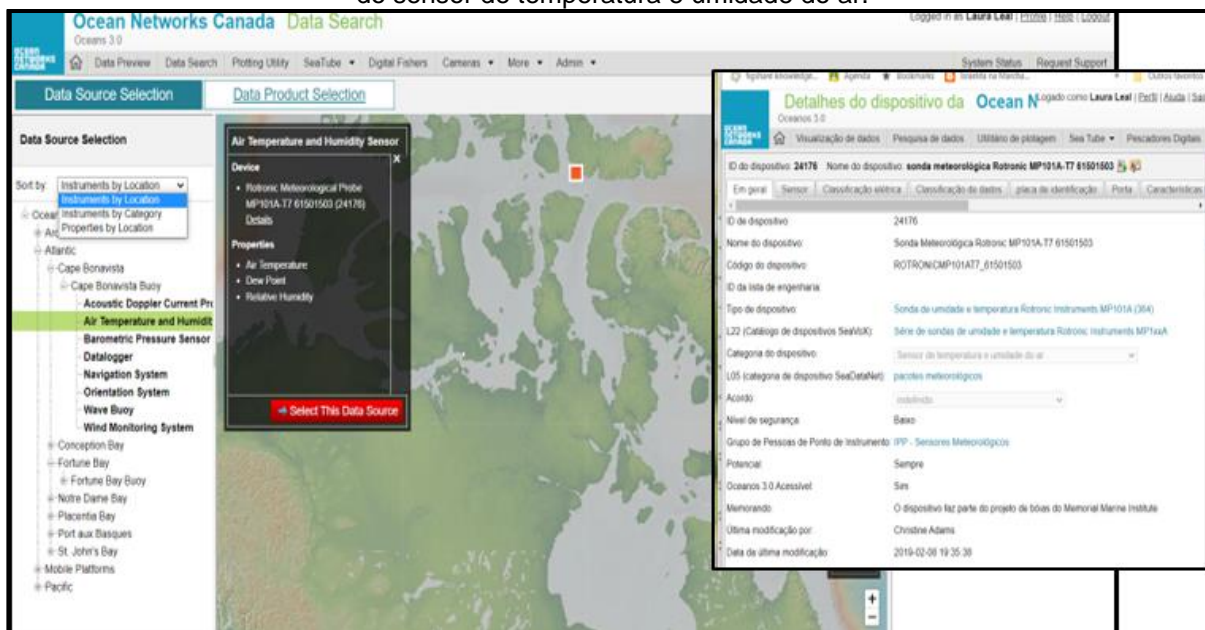
É possível visualizar os filtros de fontes de dados, filtros de intervalo de profundidade, do período e seleção de propriedade comum. Na seleção da opção de

⁶⁵ <https://ui.adsabs.harvard.edu/abs/2018AGUFMIN33E0891P/abstract>

⁶⁶ <https://data.oceannetworks.ca/>

filtro de visualização de dados, optou-se pelo Oceano Atlântico, na área *Cape Bonavista Buoy* e na verificação do sensor de temperatura e umidade do ar. Ao selecionar essa fonte de dados é remetido a um outro *link*, de acordo com a Figura 19, onde constam alguns dados como: quem fez e a data da última alteração, número de ID, o nome, código e tipo do dispositivo, entre outras informações técnicas.

Figura 19 - Visualização da localização do Oceano Atlântico, no *Cape Bonavista Buoy* e informação do sensor de temperatura e umidade do ar.



Fonte: Ocean Networks Canada: data Search

Como os dados divulgados são aparentemente *online*, torna-se complexo efetuar a análise, porque modificam-se de um dia para o outro, dependendo de onde estão ocorrendo os monitoramentos. Outro dificultador, nesse projeto, é a própria busca por informações dos dados brutos, para quem desconhece os procedimentos de como ocorre o armazenamento na base de dados, dificulta a busca e leitura dos mesmos.

4.2.1.5 Zooniverse e a Snapshot Wisconsi (ID.70⁶⁷)

O documento *“Tradeoffs and tools for data quality, privacy, transparency, and trust in citizen science”* de Anhalt-Depies e outros autores (2019), abordam as novas tecnologias, mais simples, na coleta de dados que são finos em escala, amplos em

⁶⁷ *Tradeoffs and tools for data quality, privacy, transparency, and trust in citizen Science.*

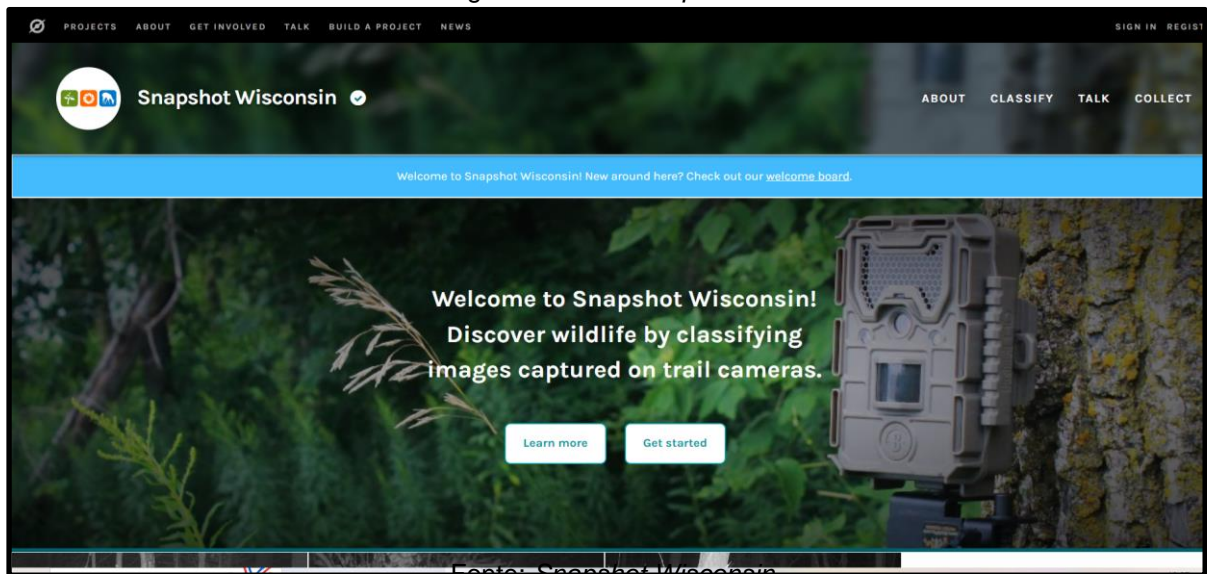
escopo e transparentes, com acesso aberto. Enfatizam que muitos dos conjuntos de dados resultantes podem conter informações confidenciais, como informações de um indivíduo ou colocar em risco os recursos naturais. No entanto, esses conflitos são relevantes para programas de Ciência Cidadã que envolvem cidadãos nas respostas de questões científicas, pois muitos dos projetos são desenvolvidos, justamente para alcançarem objetivos científicos (como dados de qualidade e compartilhamento aberto de dados) e sociais (maior transparência e confiança do público).

Apresentam um estudo de caso de um projeto de Ciência Cidadã, onde foi possível equilibrar a produção de compartilhamento de dados com a proteção de informações confidenciais.

O *Snapshot Wisconsin* foi lançado em 2016, pelo Departamento de Recursos Naturais de Wisconsin (WDNR), para monitorar a vida selvagem, a longo prazo, utilizando como recursos armadilhas fotográficas. O projeto tem como objetivos fornecer dados essenciais para tomada de decisões no manejo da vida selvagem e aumentar o envolvimento do cidadão com os recursos naturais.

A participação dos cientistas cidadãos pode ocorrer de várias formas. Primeiramente os participantes recebem treinamento do Wisconsin DNR e o equipamento necessário para participarem e são os responsáveis por enviarem as fotografias para o DNR várias vezes por ano. Na página do perfil *online*, cidadãos podem ver as fotos dos animais e ajudar na identificação das espécies. As fotos restantes, não identificadas, são transferidas para um *site* de *crowdsourcing*, onde pessoas do mundo inteiro podem ajudar na identificação dos animais. Figura 20, apresenta a página do *site*.

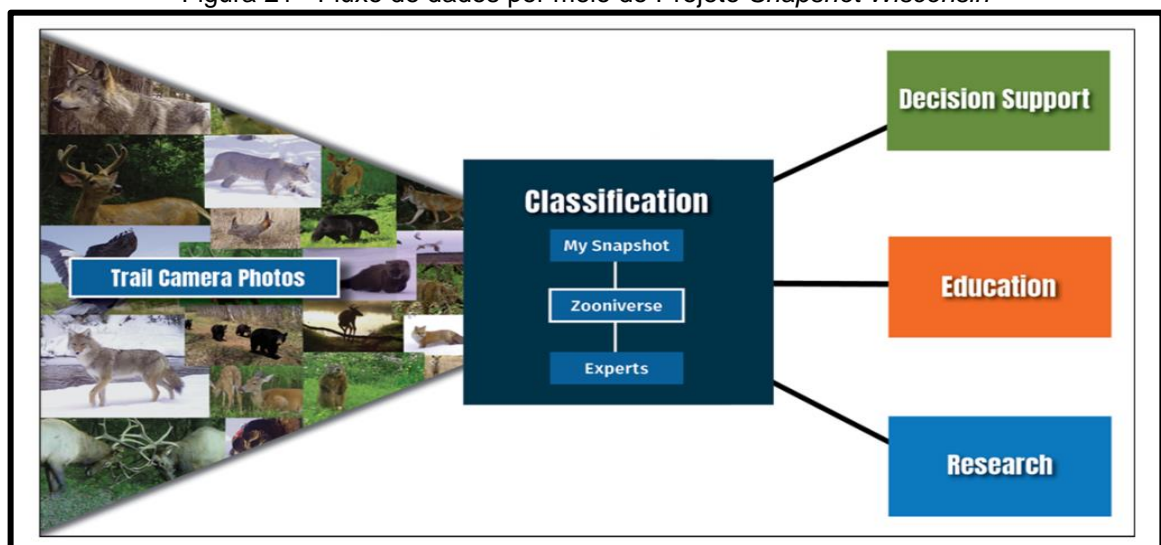
Figura 20 - Site Snapshot Wisconsin



Fonte: Snapshot Wisconsin

Apesar do projeto ter página própria na *web*, é na plataforma Zooniverse⁶⁸ que se encontram os dados mais relevantes, como por exemplo o número de cidadãos desde o início de 2023 como sendo mais 1800 como mais de 2000 câmeras de trilhas (*My Snapshot*), assim como um número em torno de 75 milhões de fotos. As fotos como difícil identificação ou não identificadas são enviadas para o Zooniverse e as fotos raras ou de pouca concordância no Zooniverse são retornadas para classificação especializadas pela equipe do Wisconsin DNR. A Figura 21, representa esse fluxo de dados.

Figura 21 - Fluxo de dados por meio do Projeto Snapshot Wisconsin



Fonte: Página do Snapshot Wisconsin na plataforma Zooniverse⁶⁹

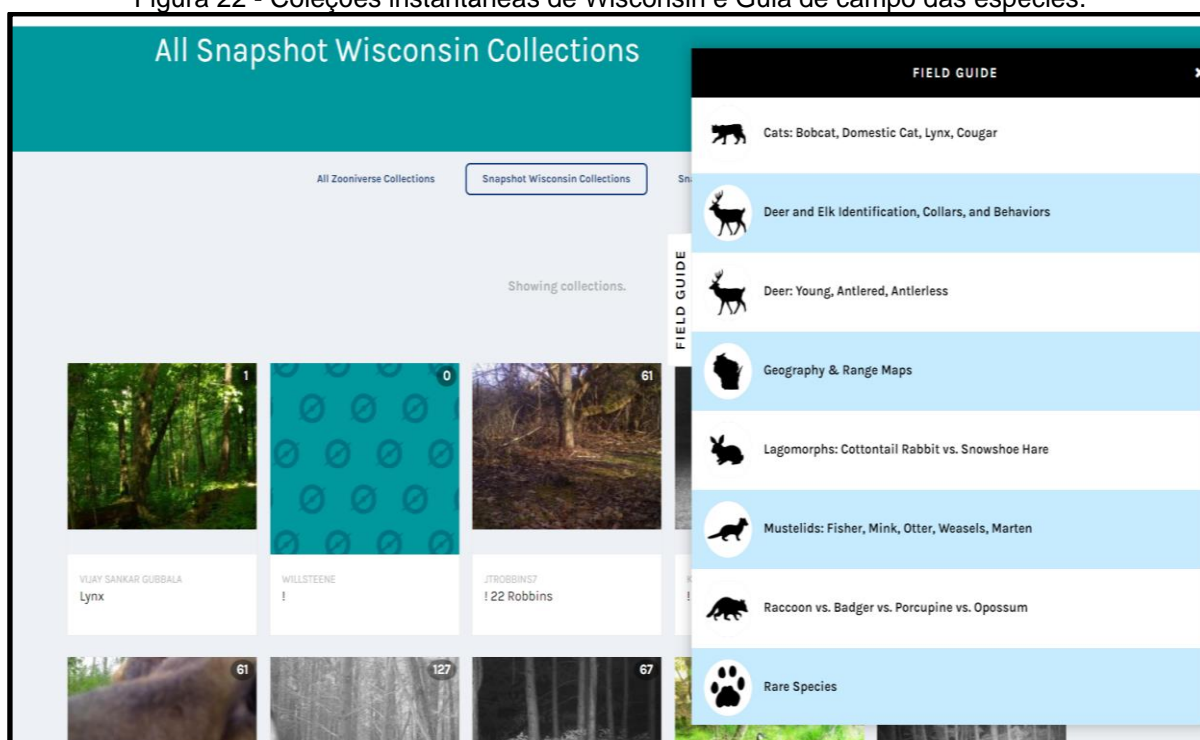
⁶⁸ Zooniverse é a maior e mais popular plataforma do mundo para pesquisas realizadas por pessoas.

⁶⁹ <https://www.zooniverse.org/projects/zooniverse/snapshot-wisconsin/about/results>

Consta ainda o grupo de trabalho com o nome dos responsáveis e a função de cada um, os moderadores, os colaboradores da pesquisa, os ex-moderadores, ex-membros da equipe, as organizações de apoio e perguntas frequentes.

Na coleta de dados estão as coleções instantâneas. Constam 137 páginas, distribuídas entre 20 participantes em cada página, com as imagens e número de registro de cada avistamento de cada espécie, de acordo com a Figura 22.

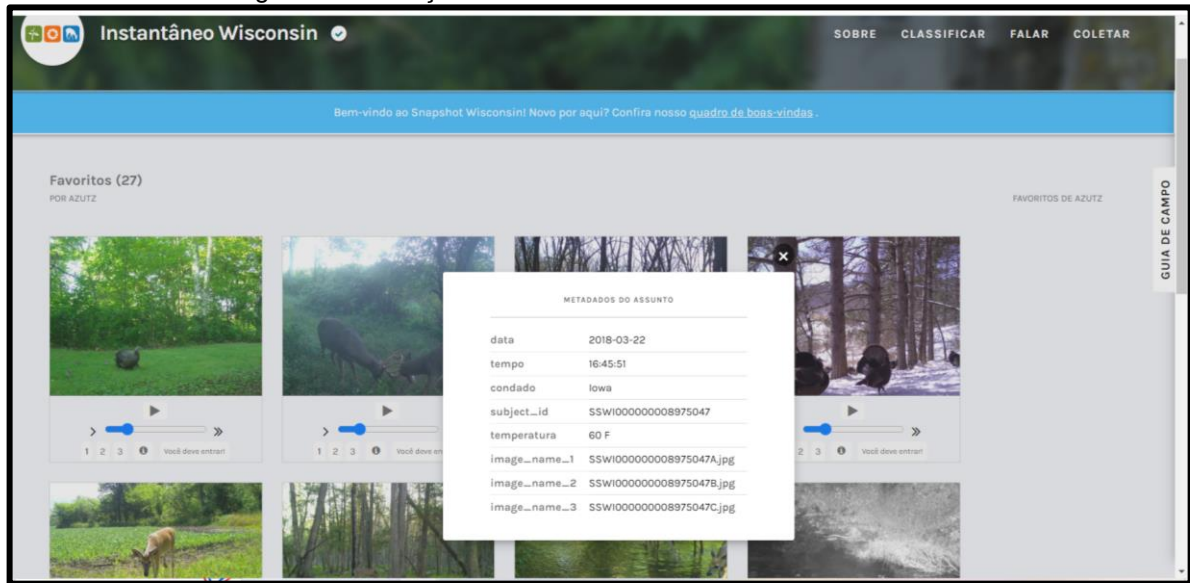
Figura 22 - Coleções instantâneas de Wisconsin e Guia de campo das espécies.



Fonte: Zooniverse, Projeto *Snapshot Wisconsin*

Ainda, nas coleções instantâneas, conforme a Figura 23, têm-se opções de visitar as favoritas que se encontram em três páginas, com um total de 27 grupos de avistadores. Ao clicar nas imagens de cada participante é possível analisar os três conjuntos de imagens postadas com os metadados do assunto, como a data em que a espécie foi filmada, tempo de duração, o local, identificação do assunto, nome de cada conjunto de imagem em .jpg, sendo que os dados de cada conjunto são diferentes, mas o animal é o mesmo.

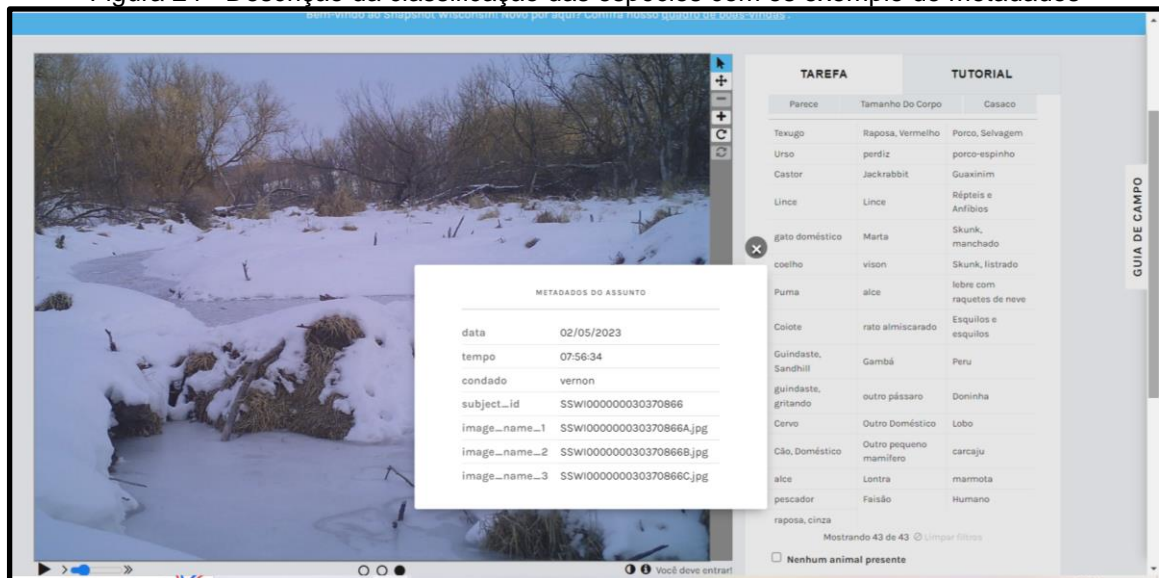
Figura 23 - Coleção instantânea dos favoritos e os metadados



Fonte: Zooniverse, Projeto *Snapshot Wisconsin*

Na parte da classificação estão disponíveis 43 espécies avistadas, clicando em cada uma tem-se a descrição detalhada da espécie (descrição do corpo, tipo de pelagem, manchas). No tutorial de classificação, identifica-se primeiro a espécie, onde pode identificar a espécie, se mais de uma espécie na imagem, o número de animais, as características, ajuda de identificação pelo guia de campo, em caso de erros de identificação, pode ser excluída a informação e por fim se no avistamento houve algum animal doente/ferido/moribundo, a classificação é feita normalmente, mas não podem ser compartilhadas no fórum de discussões do Zooniverse, em conformidade com a Figura 24.

Figura 24 - Descrição da classificação das espécies com os exemplo de metadados



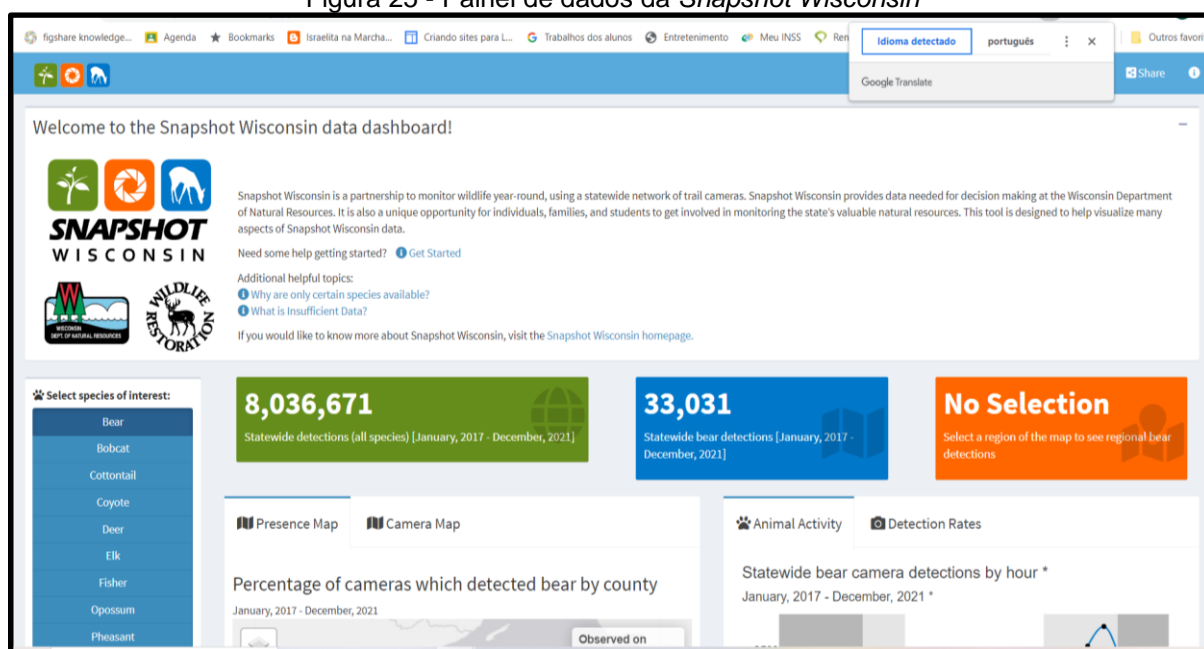
Fonte: Zooniverse, Projeto *Snapshot Wisconsin*

Dados instantâneos de estatísticas do *Snapshot Wisconsin*: o número de cientistas cidadãos é 18.187, imagens classificadas 179.592, assuntos abordados 58.972, assuntos/disciplinas completas 27.919 e consta que o projeto está concluído em 47%.

Na página do painel de dados da *Snapshot Wisconsin*⁷⁰ (Figura 25), a abordagem é o monitoramento da vida selvagem, durante todo o ano, para assim fornecer dados importantes para a tomada de decisões no Departamento de Recursos Naturais de Wisconsin. O envolvimento do cidadão é uma oportunidade na fiscalização dos recursos naturais do estado. O *site* foi projetado para ajudar na visualização dos dados gerados pelo projeto.

Consta na página seleção das espécies de interesse, seleção de mapa, onde é possível verificar qual espécie se encontra no local; o intervalo da data dos dados registrados (entre janeiro de 2017 e dezembro de 2021); percentual de câmeras que detectam as espécies.

Figura 25 - Painel de dados da *Snapshot Wisconsin*



Fonte: *Snapshot Wisconsin*

Projetos que abordam a biodiversidade, como o *Snapshot Wisconsin* disponibilizam uma variedade de dados de informações sobre o objeto pesquisado, neste caso, a vida selvagem e procuram dar visibilidade armazenando o projeto em

⁷⁰ <https://widnr-snapshotwisconsin.shinyapps.io/DataDashboard/>

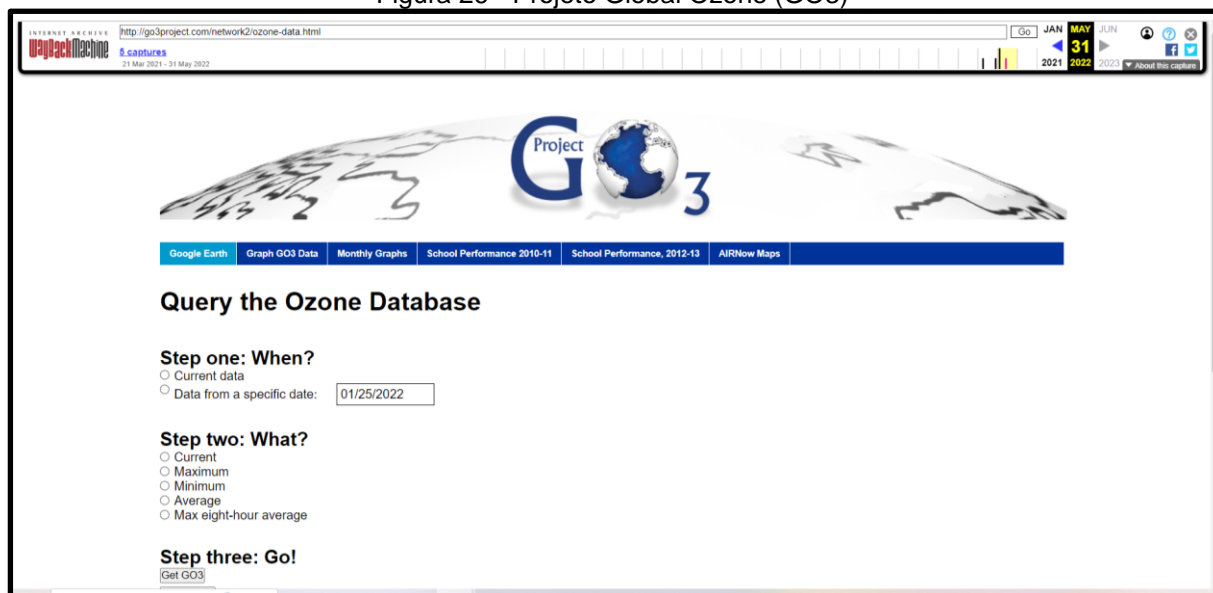
grandes plataformas como o Zooniverse, onde cidadãos de qualquer parte do mundo podem participar da identificação de forma *online*.

4.2.1.6 Internet Archive WayBack Machine (ID.69⁷¹)

No trabalho de Ellenburg e outros autores (2019), “*Global Ozone (GO3) Project and AQTreks: Use of evolving technologies by students and citizen scientists to monitor air pollutants*”, a abordagem é sobre as novas tecnologias em instrumentação de poluição do ar, tornando possível que o monitoramento seja efetuado por estudantes e cientistas cidadãos, assim como avanço na comunicação e transmissão de dados, tornando fácil o compartilhamento e a exibição de dados gráficos.

Para averiguação das informações utilizaram dois programas educacionais, o Projeto Global Ozone (GO3) e o AQTreks, onde o monitoramento da poluição do ar foi efetuado por estudantes de 100 escolas de todo o mundo, automatizando o compartilhamento dos dados. Nos projetos, o foco estava na qualidade dos dados e na seleção das tecnologias de monitoramento para uso científico. Na Figura 26 é possível visualizar a página inicial do projeto, o qual está desativado.

Figura 26 - Projeto Global Ozone (GO3)



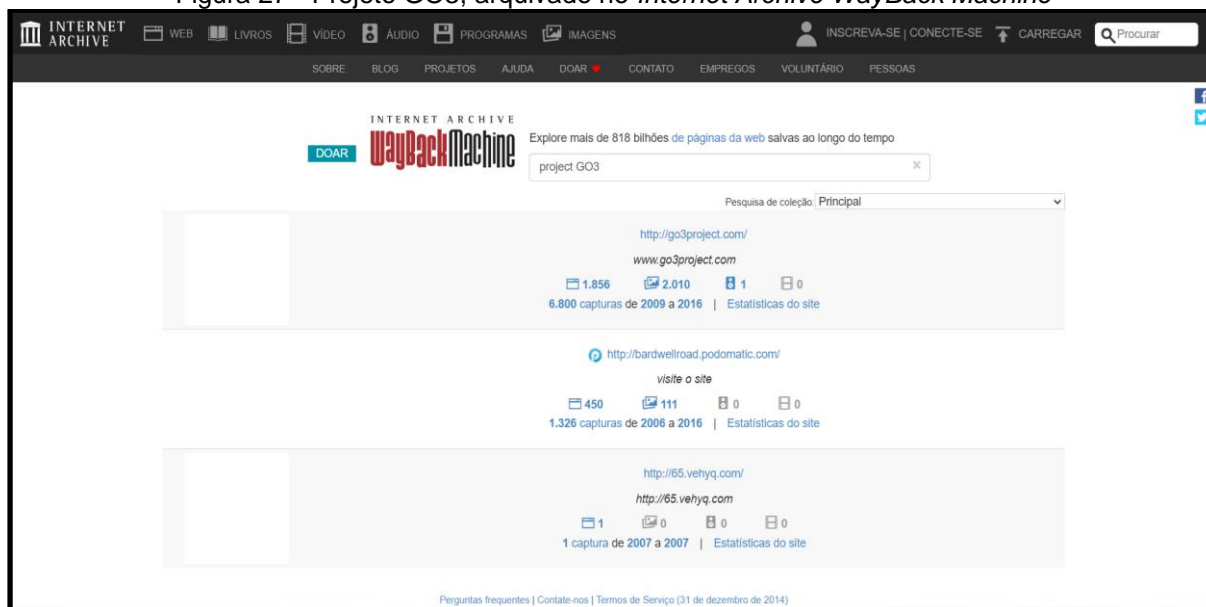
Fonte: Internet Archive WayBack Machine⁷²

⁷¹ Global Ozone (GO3) Project and AQTreks: Use of evolving technologies by students and citizen scientists to monitor air pollutants.

⁷² <https://web.archive.org/web/20220531160340/http://go3project.com/network2/ozone-data.html>

Para o Projeto GO3 é apresentado um *link* (www.go3project.com) onde deveriam constar os dados do projeto, infelizmente não está mais disponível. No entanto, após busca na internet, pelo projeto, foi possível resgatá-lo no *site Internet Archive WayBack Machine*⁷³, como arquivado, conforme Figura 27.

Figura 27 - Projeto GO3, arquivado no *Internet Archive WayBack Machine*

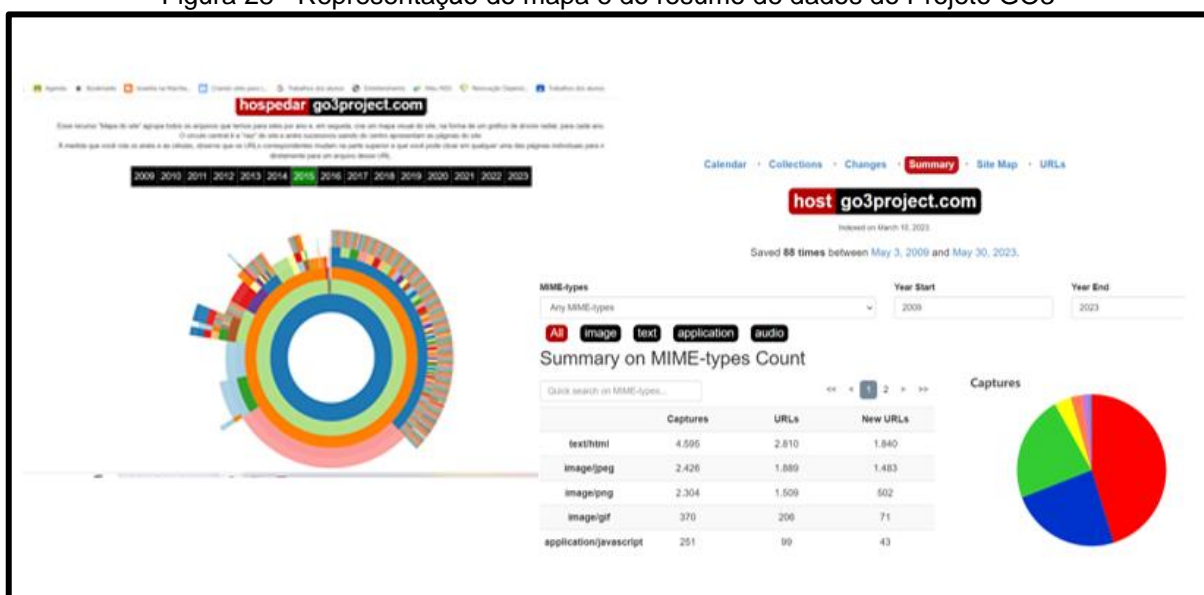


Fonte: *Internet Archive WayBack Machine*

Na página inicial, consta o *link* do *site* do projeto, o número de capturas de 6800 (entre 2009 e 2016), número de 2010 imagens do período e um áudio. Ao entrar na página das capturas, foi possível identificar 4310 URLs capturadas pelo prefixo de URL, podendo ser filtrado por URL ou tipo MIME; o resumo, onde consta o número de salvamento entre 03/05/2009 e 30/05/2023, de 88 vezes e contagem por tipos MIME, para mesmo período, em imagem, texto, aplicativo e áudio; das coleções obteve-se 5303 resultados por tipo de mídia, podendo ser ordenadas por visualizações, título, data de arquivo e o criador, desse total 2700 com certificados de transparência e dos rastreamentos das coleções identificaram-se 16282 resultados; do calendário, é exibido o número de vezes que o *site* do projeto foi rastreado pela *WayBack Machine*; captura *mapa* virtual do *site*, onde estão agrupados todos os arquivos gerados por ano, por as alterações (*changes*) onde é possível identificar e exibir mudanças no conteúdo de arquivos de URLs. É possível visualizar a imagem do calendário e o resumo na Figura 28.

⁷³ Wayback Machine é um banco de dados digital criado pela organização sem fins lucrativos *Internet Archive* - https://web.archive.org/web/20230000000000*/project%20GO3

Figura 28 - Representação do mapa e do resumo de dados do Projeto GO3



Fonte: *Internet Archive WayBack Machine*

Na leitura do artigo verificou-se que utilizam um *software* para carregarem os dados de ozônio e clima que são enviados para um banco de dados MySQL⁷⁴ e que a empresa *Sonoma Technologies, Inc.*, apoiadora do programa AirNow junto com a equipe do GO3 criaram um protocolo para transmitir todos os dados do GO3 para uma área da AirNow Tech⁷⁵, um repositório de dados *online* para agências aéreas dos EUA, a qual não foi possível o acesso para a leitura dos dados brutos.

Para o AQTreks foi criado um programa piloto como conceito de monitoramento móvel dos poluentes das atmosfera, no qual os usuários após o monitoramento tem opção de enviarem os dados para o *site CommunityAQ*⁷⁶. Para aumentar a capacidade de armazenamento de dados passaram a utilizar o banco de dados *PostgreSQL*, o qual pesquisou-se e não foram recuperadas as informações desejadas sobre os dados do Programa AQTreks.

No *site da CommunityAQ*⁷⁷, aparecem duas referências ao projeto piloto AQTreks, onde é apresentado um resumo dos Projetos GO3 e do AQTreks. Apontam o uso de sensores móveis emparelhados com um aplicativo para *smartphone* para o monitoramento da poluição do ar. A Figura 29, é um recorte do vídeo promocional

⁷⁴ <https://www.mysql.com/>

⁷⁵ <https://www.airnowtech.org/>

⁷⁶ <https://www.communityaq.com/search?query=Programa+AQTreks+>

⁷⁷ <https://www.communityaq.com/>

sobre os projetos.

Figura 29 - Imagens recortadas do vídeo do *YouTube* sobre os projetos



Fonte: YouTube -The GO3 Project and AQTreks: A decade of air pollution monitoring with students and the public ⁷⁸

Ainda sobre o AQTreks (AirQualityTrek), tem sua divulgação na plataforma SciStarter⁷⁹, voltada para projetos em Ciência Cidadã com participação dos cientistas cidadãos. Apresenta como projeto móvel de monitoramento dos níveis de poluição do ar e mapeamento da poluição do ar nos EUA, por meio da coleta e análise de dados da poluição do ar, efetuada por alunos e professores com o uso do dispositivo Personal Air Monitor (PAM), onde os poluentes do ar são medidos e os dados visualizados nos *smartphones*. Os dados coletados são carregados dos celulares via *wi-fi* ou 3G, compartilhados *online* para análise, discussão e comparação com dados Trek entre os diferentes grupos de alunos. A Figura 30, mostra a página do projeto na *SciStarter*.

⁷⁸ <https://www.youtube.com/watch?v=d1MjxGSo6Aw>

⁷⁹ <https://scistarter.org/>

Figura 30 - Dados do Projeto AQTreks (AirQualityTreks), apresentado na SciStarter

The screenshot shows the SciStarter website interface. At the top, there is a navigation bar with the SciStarter logo and the tagline "Science we can do together." Below the navigation bar, there are links for "SciStarter LIVE!", "Libraries", "Trainings", "@Girl Scouts", "Dashboard", "Projects", "Events", and "Partner Gateways". The main content area features a project card for "AQTreks (AirQualityTreks)". The card includes a photo of a group of people, a "Find more projects" link, and a "34750" count. The project details are as follows:

- PRESENTED BY:** 2B Technologies, Inc.
- GOAL:** Monitor air pollution levels and map air pollution around the US
- TASK:** Collect & analyze air pollution data using Personal Air Monitor
- WHERE:** View map...
- DESCRIPTION:** AQTreks (AirQualityTreks) is a mobile air pollution monitoring program for students, organizations, scout troops, and citizen scientists. In AQTreks, with the assistance of their teachers, students hypothesize where air pollution levels may be high or low and use a Personal Air Monitor (PAM) on Treks of their own design. Air pollutants are measured and data are viewed on their smartphones. The data are automatically uploaded from the phone via WiFi or 3G and shared on [See more](#)
- HOW TO GET STARTED:** Please email jessae@aqtreks.com if you are interested in participating. A three-week rental costs \$360 (shipping included) and includes the Personal Air Monitor, charger, program materials, technical assistance, and data storage.
- SOCIAL MEDIA:** [Twitter](#) [Facebook](#)
- HOW-TO VIDEO:** Check out the how-to video

Below the project card, there are buttons for "Off-site Project", "Visit", "Save to Review Later", and "Share". A section titled "Projects You May Like" features "TreeSnap".

Fonte: SciStarter

Projetos como o AQTreks, que está ligado a escolas - visto que as discussões sobre dados ocorrem entre alunos e professores, gerando atividades educacionais sobre os resultados -; e conectado com um dos Objetivos de Desenvolvimento Sustentável da ONU (ODS), é possível verificar o retorno para a sociedade, por meio das Metas de Desenvolvimento Sustentável aplicadas em cidades e comunidades sustentáveis, educação de qualidade e ação climática.

4.2.1.7 Repositório de citações *Open Dachs* (ID 21385⁸⁰)

O trabalho de Arnold, Lecher e Vogt (2020), "*Multilingual Research Projects: Non-Latin Script Challenges for Making Use of Standards, Authority Files, and Character Recognition*", trata-se de pesquisa acadêmica na busca por textos digitais que não estejam no alfabeto latino (NLS) e seus desafios, referem-se a jornais chineses da era republicana (1912-1949).

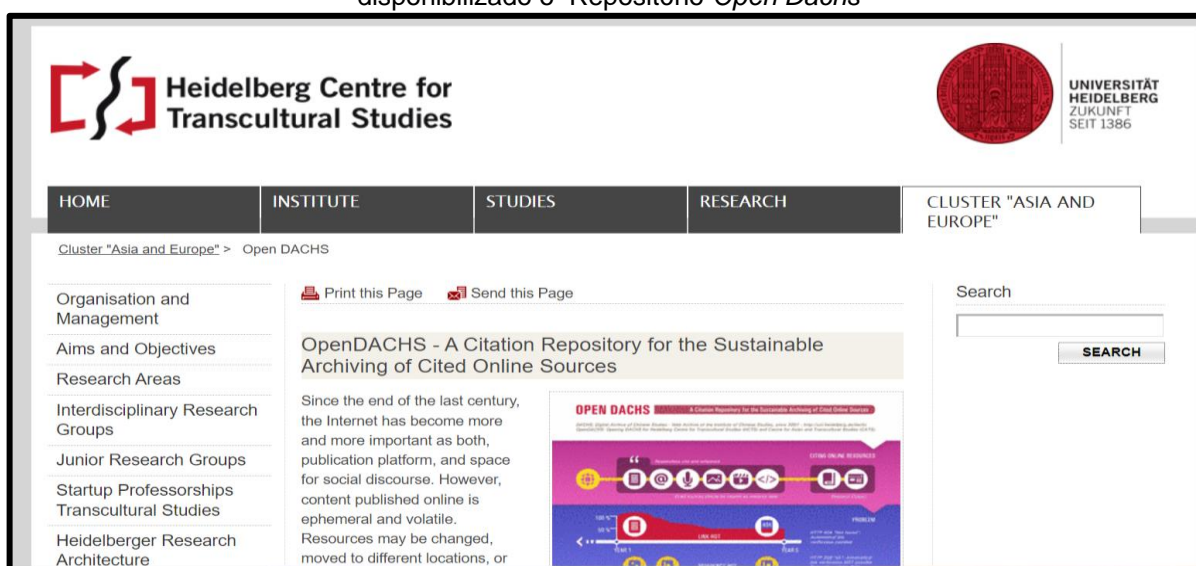
São apresentados dois estudos de caso na demonstração dos diferentes aspectos de tarefas que podem estar vinculadas ao *hardware* NLS. O primeiro se concentra nos padrões de metadados usados para descrever o *hardware* NLS e o segundo caso é a análise dos próprios dados de pesquisa. Para a interpelação foi combinado os métodos manuais e computacionais, como o *crowdsourcing*.

⁸⁰ *Multilingual Research Projects: Non-Latin Script Challenges for Making Use of Standards, Authority Files, and Character Recognition*.

Conforme o artigo, a seleção de referências é fornecida das páginas da *Web* por meio do repositório de citações *Open Dachs*, em parceria entre o Instituto de Sinologia do Centro de Estudos do Leste Asiático e a Arquitetura de Pesquisa de Heidelberg do Centro de Estudos Transculturais de Heidelberg (HCTS).

O *Open Dachs* disponibiliza serviços e fluxo de trabalho para pesquisadores que querem ter suas pesquisas arquivadas, com geração de identificadores DOI, podendo ser usados em publicações e criação de registros de catálogos de biblioteca. Ao efetuar a busca pelo repositório *Open Dachs*, automaticamente é remetido para a página mostrada na Figura 31.

Figura 31 - Página do Centro de Estudos Transculturais de Heidelberg (HCTS) onde está disponibilizado o Repositório *Open Dachs*



Fonte: Universität Heidelberg⁸¹

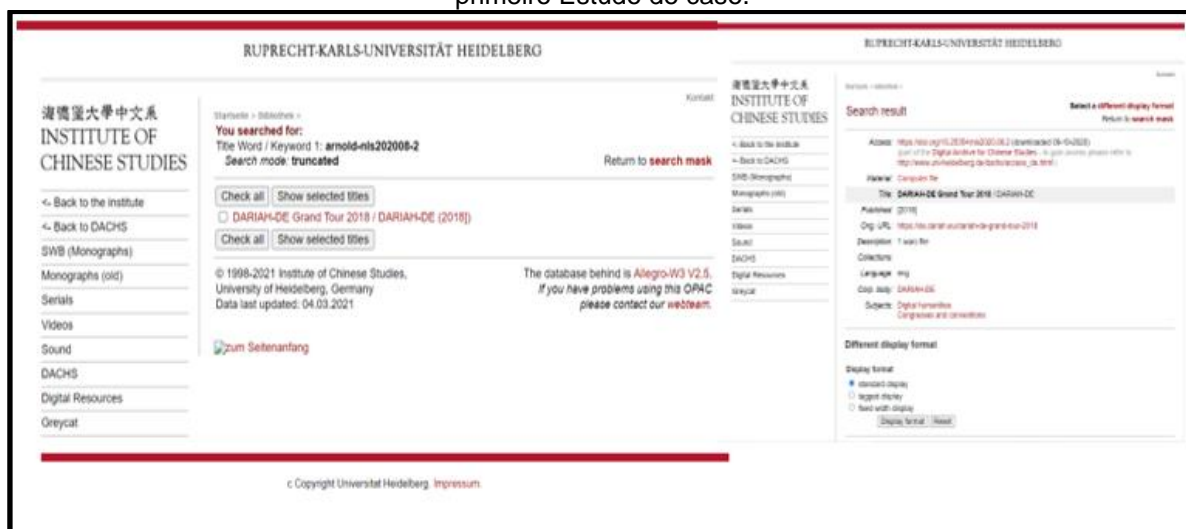
Na página há um pequeno histórico sobre os recursos utilizados pelo *Open DACHS* nas publicações, usando o formato de dados WARC, publicado como padrão ISO 28500:2017. O *software* de código aberto *Heritrix* é utilizado como rastreador na *web*. O *Internet Archive* é usado para arquivar arquivos, podendo ser visualizado por meio de ferramentas como o *wayback machine*. Os registros individuais arquivados, também podem ser visualizados com ferramentas como o *wayback machine*. Enfatizam que o *Open DACHS* oferece a opção de criar um *site* inteiro ou arquivar partes dele, configurar arquivos recorrentes, páginas da “*deep web*”, entre outros.

Os dados de pesquisa de cada etapa dos projetos são armazenados em diferentes repositórios ou bases de dados de dados, como exemplo os arquivos de

⁸¹ <https://www.asia-europe.uni-heidelberg.de/index.php?id=4425>

esquema expandido, são publicados no Repositório de documentos *GitHub*⁸², ao entrar no *link* indicado e definir os dados de interesse, remete para um outro *link*, como mais informações e orientado que seja efetuado o acesso por meio da universidade. Embora seja de acesso aberto, existem restrições de privacidade. A Figura 32 apresenta os dados gerados pelo artigo.

Figura 32 - Dados gerados do artigo de Arnold e outros (2020), no repositório *GitHub*, relacionado ao primeiro Estudo de caso.



Fonte: Universidade de Heidelberg⁸³

Na busca dos resultados finais, como o *site* é de acesso restrito, buscou-se diferentes forma de acesso, mas tendo como resultado de pesquisa poucos elementos como os acesso pelo DOI (baixado em 09/10/2020 parte do Arquivo Digital para estudos Chineses e a URL que precisa de senha para acesso; o tipo de material, listado como arquivo de computador; o título do estudo; a data de publicação; a URL original; a descrição de 1 arquivo *ward*; coleções; linguagem; corporação e assuntos, podendo ser exibido em diferentes formatos.

O primeiro estudos de caso utilizou obras culturais do guia *Catalogando Objetos Culturais: Um Guia para Descrever Obras Culturais e Suas Imagens (CCO)*, que abrange diferentes tipos de obras culturais como pinturas, manuscritos, gravuras, esculturas, sítios arqueológico, entre outros, visto fornecer padrões para conteúdo de dados, que indicam várias formas de combinar estruturas e dados padronizados, para

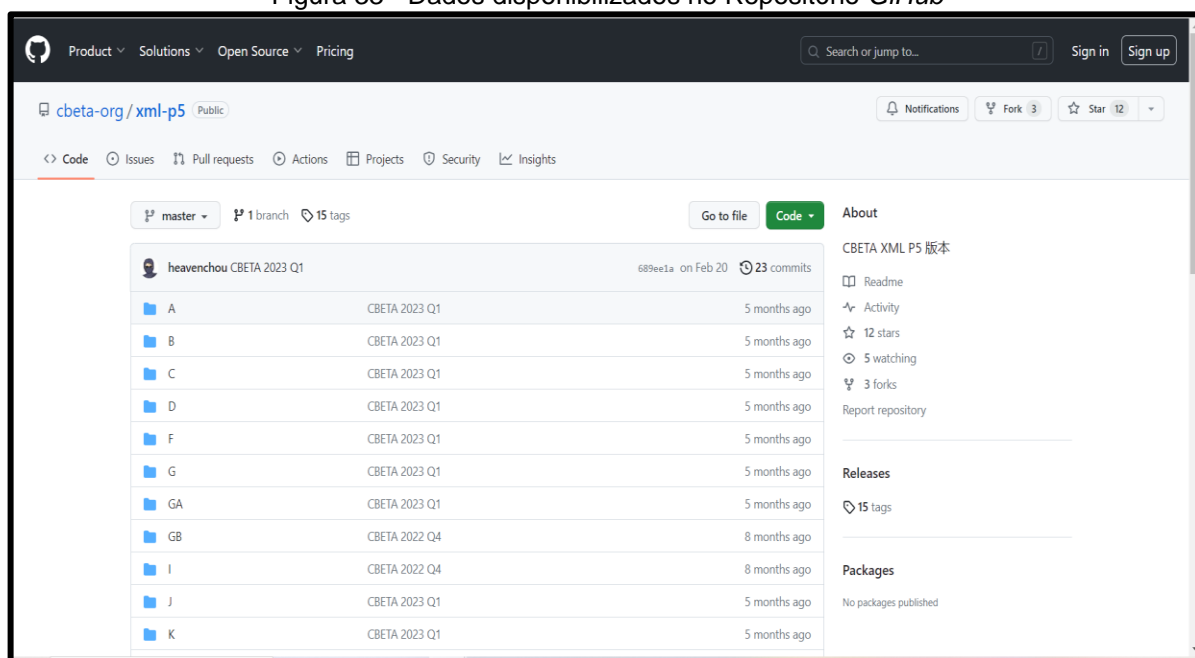
⁸² https://www.zo.uni-heidelberg.de/boa/digital_resources/dachs/access_de.html

⁸³ http://www.sino.uni-heidelberg.de/cgi-bin/webkat_imperia/allegro.pl?db=dachs&expand=yes&maxkrztitel=80&T11=arnold-nls202008-4

categoria de descrição de obra de artes (CDWA) e VRA Core, versão 4.0, com padrões para valores de dados, como vocabulário de autoridade.

No segundo estudo de caso, tratou-se da digitalização e texto completo. O material foi coletado de um subproduto de um projeto de pesquisa que continha mais de 300 publicações de era republicana, que podem ser lidas *online* em acesso aberto. De acordo com a Figura 33, os dados estão publicados em banco de dados textual no repositório público CBETA XML P5 GitHub, no formato TEI XML, atualizado regularmente.

Figura 33 - Dados disponibilizados no Repositório *GitHub*



Fonte: Repositório *GitHub*⁸⁴

4.3 INFORMAÇÕES RELEVANTES AO ESTUDO

Optou-se por efetuar mais uma subseção após a leitura de alguns documentos que apresentavam informações de abordagem consideradas interessantes na Ciência Cidadã.

Dos dados, sobre a definição de Ciência Cidadã e *crowdsourcing*, após leitura dos artigos, constatou-se que nos últimos dois anos, o número de artigos que apresentavam essas definições era raro, mesmo aqueles que não se reportavam a projetos. Supõe-se que para os pesquisadores, dos artigos analisados, esses

⁸⁴ <https://github.com/hcts-hra/ecpo-segment>

conceitos já estão claros e bem definidos tanto para a área científica, quanto para a sociedade, não havendo mais a necessidade de estudos exaustivos sobre essa terminologia, mostrando assim a credibilidade dos estudos sobre Ciência Cidadã.

No artigo de Ganzevoort e outros autores (2017), no trabalho “*Sharing biodiversity data: citizen scientists’ concerns and motivations*” apresenta uma nova expressão utilizada para se referir ao cientista cidadão, como instrumento de pesquisa, chamou a atenção por ser peculiar: *drones* de dados automatizados. Ainda fazem o alerta sobre a expressão ser diferente e poder causar conflitos, pois os cientistas, como produtores de conhecimento, não gostariam de ser chamados de *drones* de pesquisa. Nesse contexto, os autores, entendem que o cientista amador tem ideias próprias sobre os dados que coletou e podem se recusar a compartilhá-los, mostrando assim que em muitos projetos eles não são coletores dóceis, como esperado, podendo gerar problemas na qualidade dos dados finais.

Os autores acima mencionados, abordam também as questões sobre qual é o perfil desses cidadãos. Discorrem que em projetos sobre pássaros a maioria são de homens brancos, mais velhos com renda e nível de educação elevada. Enfatizam que em muitos projetos devem buscar os cidadãos certos, que tenham motivação sobre o tema escolhido, analisando por exemplo a faixa etária, gênero, escolaridade e a área de trabalho ou estudos deles, pois essas características podem determinar se estão motivados a permanecerem durante todo o desenvolvimento do projeto.

No que se refere a abordagem sobre e os cientistas cidadãos, em alguns projetos de *crowdsourcing*, é possível perceber o compartilhamento dos dados, mas não existe a participação do cidadão, em nenhuma das etapas desenvolvidas, seja na criação de ferramentas ou no desenvolvimento de projetos. Conforme Parson e outros autores (2022) a própria União Europeia (UE), na busca por soluções para cumprir metas e propostas relacionadas a sustentabilidade do planeta, utilizam-se da criação de compartilhamento de *big data*, por meio da “ferramenta *crowdsourcing*”, sem priorizar trabalho do cientista amador na coleta de dados e sim o uso de um algoritmo, onde a informação desejada é momentânea e compartilhada, concluindo o objetivo desejado.

No artigo “*Research data management in health and biomedical citizen science: Practices and prospects*”, de Borda, Gray e Fu (2020), as plataformas são desenvolvidas com objetivo de dar suporte às necessidades básicas de gerenciamento de dados na análise na *web* em como ela está organizada. A

operacionalidade das plataformas pode variar, algumas oportunizam a coleta e análise de dados ou agrupamento de dados pessoais, do tipo de projetos ou fornecimento de acesso de dados de diferentes áreas do conhecimento ou compartilhamento de dados com outros pesquisadores ou até mesmo a venda de dados de pesquisa para terceiros.

Existem plataformas que coletam dados padrão de usuários em páginas da internet, por exemplo, em visitas a *sites*, onde armazenam as informações de navegação dos visitantes ou por meio dos *cookies* para verificação de quanto tempo um usuário fica visitando determinado sítio eletrônico ou até mesmo mapeamento dos interesses que o internauta tem por determinado assunto.

As plataformas podem divergir entre si, apresentarem-se com *software* de código de acesso aberto e outras não. Alguns permitem que os usuários decidam como os dados que coletaram vão ser usados e com quem essas informações serão compartilhadas, caso os internautas optem por não fazerem o compartilhamento dos dados, eles serão descartados da pesquisa.

Na abordagem de Kaur e outros autores (2012), “*Scholarometer: A Social Framework for Analyzing Impact across Disciplines*”, referenciou que a tipologia usada para *crowdsourcing*, como sendo um novo modelo de plataforma de compartilhamento de dados e informações, semelhante ao modelo *Scholarometer*⁸⁵ ou uma espécie de *site online* de armazenamento, onde estão armazenados elementos de uso público, como fotos para identificação, por exemplo, de um determinado animal.

No que se refere a plataformas em projetos de Ciências Cidadã, em biodiversidade, como Spipoll⁸⁶, que na realidade é um projeto de Ciência Cidadã participativa, de acesso a todos, volta para a biodiversidade, com foco em polinização⁸⁷, onde o cidadão tem a oportunidade de criar ou participar dos projetos, para isso, basta ter uma máquina digital e tempo para fotografar, monitorar a natureza, coletar dados e postar no *site* para que outros participantes identifiquem a espécie escolhida.

Cada participante tem sua própria coleção. Para começar a coleção é necessário escolher o tipo de flor, tirar 3 fotos da planta e uma foto da paisagem.

⁸⁵ Scholarometer é uma plataforma social que estende o Google Scholar para facilitar a análise de citações acadêmicas. <https://scholarometer.indiana.edu>.

⁸⁶ <https://www.spipoll.org/>

⁸⁷ Modo de reprodução das plantas.

Elementos como temperatura, nuvens no céu, velocidade do vento, o ambiente onde se encontra (floresta, parque), fotos tiradas na sombra ou não. Para a identificação da planta é ofertado um *link* para PlantNet⁸⁸ de identificação automática usando as fotos do cientista cidadão. Os dados de identificação são validados por participantes diferentes.

Figura 34 – *Spipoll* e as últimas coleções



Fonte: *Spipoll*⁸⁹

Os dados de pesquisa estão disponibilizados, mas percebe-se que não existe uma forma de pesquisar os dados como um único conjunto (a mesma espécie pode estar em mais de uma coleção e em locais diferentes), que dificulta uma pesquisa científica mais ampla. No entanto, os dados individuais têm seu caráter fundamental nas pesquisas científicas, pois ao monitorar uma determinada espécie é possível saber de onde migraram - abelha selvagem *megachiles sculpturalis*, originária da Ásia, foi detectada em Marselha e em outras regiões da Europa -, e se estão em risco de extinção, como é o caso da abelha melífera (conforme divulgado no *site*).

Outra forma de disponibilizar os dados é abordada por Sun e outros autores (2021) que relatam em seu artigo: “*Share your data carefree: An efficient, scalable and*

⁸⁸ <https://identify.plantnet.org/pt-br>

⁸⁹ <https://www.spipoll.org/>

privacy-preserving data sharing service in cloud computing”, sobre o compartilhamento de dados em serviços em nuvem, utilizado por aplicativos, considerado como sendo uma terceirização de dados para a nuvem, como não confiáveis, pois leva a preocupação com a violação de privacidade de dados, sendo necessário o desenvolvimento ou aperfeiçoamento de estruturas confiáveis e seguras que pudesse preservar a privacidade do armazenamento de conjunto de dados criptografados na nuvem.

5 CONSIDERAÇÕES FINAIS

No apanhado do trabalho, de forma a atender o primeiro objetivo específico de identificar na literatura trabalhos que envolvem Ciência Cidadã e compartilhamento de dados, considera-se que este objetivo foi atingido, pois foi possível reconhecer que os cidadãos contribuem com suas habilidades e experiências, a Ciência Cidadã amplia o alcance da pesquisa científica, o pensamento crítico e o aumento da confiança na ciência, efetuando o seu papel social e na produção de artigos acadêmicos.

É por meio da literatura gerada dos estudos e projetos de Ciência Cidadã que ocorre a disseminação do conhecimento científico obtido, quando é efetuado o compartilhamento da produção e da recuperação de dados, produzidos por projetos, o que possibilita o acesso da informação em lugares e contextos que seria mais difícil de atingir se fossem efetuados somente por pesquisadores profissionais. Sendo o compartilhamento de dados gerados pela Ciência Cidadã uma de suas características, que possibilita a abertura e a disponibilidade desses dados para outros pesquisadores e para a sociedade em geral, por meio de sua produção científica.

Em relação à identificação de repositório de dados nos trabalhos selecionados, obteve-se um resultado limitado. O estudo buscou determinar se os dados descritos nos artigos eram compartilhados e armazenados em repositórios de dados de pesquisa. No entanto, constatou-se que nem sempre o armazenamento de dados de pesquisa ocorre em repositórios tradicionais, sendo comum encontrar outras denominações tais como: plataformas, base de dados na nuvem, serviço em nuvem, banco de dados *online*, portal de dados, interface de base na internet, entre outros.

No caso dos repositórios e similares, alguns artigos os mencionam apenas como exemplos teóricos, utilizados para ilustrar possibilidades, mas não como foco principal do estudo. Portanto, esses artigos não abordam detalhes sobre a implementação prática de armazenamento dos dados de pesquisa ou oferecem sugestões específicas para esse fim. Em vez disso, se concentram em aspectos conceituais ou em considerações gerais relacionadas à importância do compartilhamento de dados, gerados nos projetos de Ciência Cidadã.

Vale destacar a importância da disponibilidade e acessibilidade dos dados de pesquisa, pois a promoção de práticas de Ciência Aberta, incluindo o

compartilhamento de dados em repositórios é uma meta importante para fortalecer a integridade e a confiabilidade da pesquisa.

Somente em sete artigos foram identificados projetos de Ciência Cidadã com a disponibilização dos dados em repositórios de dados. Os que disponibilizam os resultados dos dados de pesquisa gerados nos projetos de Ciência Cidadã, apresentaram dinâmicas diferentes. Os dados voltados para a biodiversidade, na sua maioria são apresentados pelos resultados individuais de cada participante que disponibiliza seu enfoque de estudo *online*, como por exemplo no Zooniverse, que é considerada uma das maiores plataformas do mundo, no uso da participação do cidadão em diferentes áreas do conhecimento. Onde é possível visualizar e participar dos projetos, assim como constatar os resultados dos dados quando disponibilizados pelos seus desenvolvedores, mas o armazenamento em dados brutos, não estão disponíveis. Todavia em repositório como o *Smithsonian Research Online* e *MFN Data Repository* é possível enxergar esses dados, conforme análise no escopo desta pesquisa.

Na análise sobre como estão armazenados os dados de pesquisa oriundos dos artigos de Ciência Cidadã e dos tipos de dados disponibilizados pelos repositórios de dados que os disponibilizam, conclui-se que existe uma variedade de interpelação e flexibilização no uso de coleta, compartilhamento e armazenamento dos dados procedentes dos artigos e que sim, são disponibilizados na sua totalidade ou parcialmente, variando com o propósito do repositório e a natureza das pesquisas ou projetos.

Entretanto, devido a diversidade e abordagens distintas dos temas da Ciência Cidadã, acabam por se refletir em uma variedade de pesquisas conduzidas em diferentes disciplinas e campos de conhecimento. Isto faz com que o tamanho dos conjuntos de dados apresente tamanho e complexidades diferentes, alguns com uma grande quantidade de informações bem detalhadas e outros menores e sucintos, portanto as análises apresentadas nessa pesquisa, contemplam tamanhos diferentes de dados.

A necessidade de estabelecer a padronização dos dados é essencial, pois se faz necessário para que seja possível estabelecer maneira de garantir a consistência e interoperabilidade entre os diferentes conjuntos de dados gerados, visto que a padronização favorece a comparação e a combinação de dados dos diferentes projetos de Ciência Cidadã, o que não foi possível constatar em todos os projetos

analisados, pois em alguns foi necessário fazer buscar em *links* e *hiperlinks* distintos.

Dos projetos que disponibilizam seus dados em acesso aberto, visando promover a transparência do que foi analisado e permitindo que outros pesquisadores verifiquem e utilizem esses dados, mas, no entanto, dos projetos analisados, somente em três essas informações estavam disponíveis e completas. Nos quatro restantes, as informações estavam presentes, mas disponibilizadas parcialmente ou com acesso restrito, não permitindo a análise na íntegra.

Para essa análise levou-se em consideração o grande volume de dados, devendo o armazenamento ser capaz de crescer de forma organizada e de garantir a qualidade e preservação a longo prazo, o que pode ser perceptível nos projetos apresentados.

Na colaboração e compartilhamento, as ferramentas utilizadas devem ser acessíveis e fáceis de usar, visto que a análise dos dados coletados é efetuada posteriormente. Nos projetos analisados, nem todos deixam claro quais as ferramentas que são utilizadas, mas nos projetos de interação *online* foi possível verificar o uso dessas ferramentas e sua acessibilidade.

Os dados disponibilizados nas análises, foram diversos e nem todos os projetos apresentavam os tipos de dados esperados, uns disponibilizavam seus dados do tipo *DataSet* em formatos planilhas (Excel), arquivos CSV, imagens, áudios e vídeos, atendendo as necessidades de armazenamento e análise de cada projeto e outros não foi possível identificá-los, conforme a análise nos repositórios estudados no *corpus* dessa pesquisa.

Os dados brutos variaram de acordo com o projeto, face refletirem os diferentes tipos de pesquisas e projetos científicos, costumam ser originais coletados antes de qualquer processamento ou análise. Os dados contemplados nos projetos analisados continham dados geoespaciais, como informações geográficas, mapas, gráficos, endereços, imagens estáticas ou em movimento, informações relacionadas ao clima (temperatura, umidade do ar, medição de poluentes), entre outras, podendo explorar os dados de forma intuitiva e visual.

Do formato de dados identificou-se arquivos de *download*, nas formas mais comuns como CSV, Excel, XML, TIFF, ZIP, XLSX, entre outros. Alguns repositórios ou similares como o Zooniverse e Spipoll permitem ao público visualizarem e interagirem com os dados diretamente no navegador sem a necessidade de baixar os arquivos, conforme imagens apresentadas na análise do *Smithsonian Research*

Online/Plataforma Dryan, Zooniverse e Internet Archive WayBack Machine.

Ressalta-se que o armazenamento dos dados de pesquisas que envolvem a Ciência Cidadã, é norteado pela colaboração, acessibilidade, privacidade e segurança, assim como a necessidade de preservar e compartilhar o conhecimento, gerando desta forma um retorno para a sociedade.

O estudo realizado buscou compreender a relação entre a Ciência Cidadã e a participação do cidadão não cientista em artigos acadêmicos, a partir do armazenamento de dados de pesquisa da Ciência Cidadã.

De forma geral, no decorrer da análise dos artigos e projetos, ficou evidente a importância da participação do cientista cidadão em projetos de Ciência Cidadã. Isto se deve a diversos fatores como a inclusão do papel desse cidadão nas práticas e o grau de envolvimento nos projetos, a compreensão do perfil dos cidadãos engajados e o impacto de suas contribuições para a ciência. Ainda deve-se levar em conta a relevância da interação e colaboração entre os cidadãos e os cientistas e os pesquisadores envolvidos.

Os repositórios desempenham um papel importante na transparência, na colaboração e na reutilização de conjunto de dados, motivando os avanços em diferentes áreas do conhecimento, se bem utilizados.

Quando a sociedade tem acesso as informações geradas pelos projetos, muitos cidadãos se sentem motivados a se envolverem em observações, monitoramento e coleta de dados ou até mesmo proporem novos projetos, entretanto dos 134 artigos recuperados, somente em sete foi possível identificar os dados de pesquisa armazenados em repositórios de dados acessíveis pelo cidadão.

Pode-se concluir que os objetivos foram atingidos visto que foi possível identificar as relações nas fontes analisadas. É possível identificar a relação entre a ciência e a participação do cientista cidadão em artigos acadêmicos, no contexto da Ciência Cidadã e do depósito de dados de pesquisa colaborativa, gerados nos projetos desenvolvidos, embora ainda incipiente, visto o número de projetos que disponibilizam seus dados em repositórios, conforme demonstrado na análise da pesquisa aqui apresentada.

Os cientistas cidadãos participam ativamente dos diferentes estágios do processo científico, ligados a Ciência Cidadã, por meio de coleta de dados, análise e interpretação dos resultados e dependendo das propostas de projetos até em publicações de artigos acadêmicos, onde tem o reconhecimento e valorização e sua

contribuição para a pesquisa. A participação dos cidadãos na ciência cidadã, pode promover a divulgação científica e fortalecer a relação entre a ciência e a sociedade.

A Ciência Cidadã possui vários pontos positivos, mas ao mesmo tempo, alguns podem ser melhorados para que haja mais eficiente. Dos pontos positivos, destaca-se que é possível participar de projetos a nível global, envolvendo cidadãos de diferentes regiões do planeta, como também aumentar a quantidade e diversidade de dados disponíveis para a pesquisa, assim como já mencionado, o envolvimento do cidadão no processo científico, na ampliação da capacidade de pesquisas, na diversidade de conhecimentos, em pesquisa de novas áreas de pesquisa e na redução de custos associados a pesquisa, viabilizando projetos que de outra forma, seria inviável economicamente.

Das melhorias, por exemplo para garantir a qualidade dos dados, a necessidade de procedimentos adequados, que podem ser por meio de treinamento e supervisão adequada. O acesso aberto aos dados de pesquisa, aumentando a transparência e a reprodutibilidade, disponibilizando-os em repositórios de acesso aberto, permitindo que pesquisadores e a sociedade possam reutilizá-los. Estimular a integração as pesquisas em Ciência Cidadã com as pesquisas tradicionais, garantindo que os dados coletados por cidadãos sejam incorporados de forma correta nos estudos científicos, bem como a valorização e reconhecimento da capacidade de participação dos cidadãos nos projetos e estudos científicos.

Por fim, medir e avaliar o impacto dos projetos de Ciência Cidadã, demonstrando assim sua efetividade na resolução de problemas e na contribuição para o conhecimento científico, o que não foi identificado nos projetos analisados nesse escopo.

Para estudos futuros, almeja-se mapear os projetos de Ciência Cidadã na região sul do Brasil, em diferentes áreas científicas, com foco nas práticas de acesso aberto, disponibilização dos dados, publicação de conteúdo, existência de repositórios, participação de projetos internacionais de Ciência Cidadã, pois entende-se que é uma abordagem valiosa para o cenário da Ciência Cidadã na região.

O estudo poderia buscar interagir com pesquisadores e cientistas cidadãos envolvidos nos projetos, por meio de visitas, entrevistas, questionários, para obter informações mais detalhadas sobre os projetos e seus resultados.

Em uma abordagem combinada de pesquisa qualitativa e quantitativa permitiria compreender a dinâmica da Ciência Cidadã na região sul do Brasil. Além disso,

pressupõem-se, o contato direto com os pesquisadores e os cidadãos possibilitaria ter uma visão mais profunda sobre as percepções e impactos dos projetos.

Os resultados desse estudo poderiam contribuir significativamente para o avanço da Ciência Cidadã na região, fornecendo informações relevantes para aprimorar práticas, incentivar uma maior colaboração dos cidadãos e promover a transparência e abertura na pesquisa científica.

REFERÊNCIAS

ALBAGLI, S. Ciência aberta em questão. In: ALBAGLI, S.; MACIEL, M. L.; ABDO, A. H. **Ciência aberta, questões abertas**. Brasília: IBICT; Rio de Janeiro: UNIRIO, 2015. p. 9-26. Disponível em: <http://livroaberto.ibict.br/handle/1/1060>. Acesso em: 31 jun. 2022.

ALBAGLI, S.; CLINIO, A.; RAYCHTOK, S. Ciência aberta: correntes interpretativas e tipos de ação. **Liinc em Revista**, Rio de Janeiro, v. 10, n. 2, p. 434-450, nov. 2014. DOI: <https://doi.org/10.18617/liinc.v10i2.749>. Acesso em: 22 jun. 2022.

ANHALT-DEPIES, C *et al.* Tradeoffs and tools for data quality, privacy, transparency, and trust in citizen science. **Biological Conservation**, v. 238, 2019. DOI: <https://doi.org/10.1016/j.biocon.2019.108195>.

APOLLINÁRIO, F. **Dicionário de metodologia científica**: um guia para a produção do conhecimento científico. São Paulo: Atlas, 2004.

ARNOLD, M. *Multilingual Research Projects: Non-Latin Script Challenges for Making Use of Standards, Authority Files, and Character Recognition*. **Digital Studies/Le champ numérique**, v. 12, p. 1–36, 2022. DOI: <https://doi.org/10.16995/dscn.8110>.

AUDUBON. **History of the Christmas Bird Count**. New York (USA), [s.d]. Disponível em: <https://www.audubon.org/conservation/history-christmas-bird-count>. Acesso em: 13 jul. 2022.

BARDIN, Laurence. **Análise de conteúdo**. Edição revista e ampliada. São Paulo: Edições 70 Brasil, 2016.

BARTKY, I. R. *The Adoption of Standard Time*. **Technology and Culture**, v. 30, n. 1, p. 25–56, 1989. JSTOR. Disponível em: <https://www.jstor.org/stable/3105430?origin=crossref>. Acesso em: 12 jun. 2023.

BBC. **WW2 People's War**. *An archive of World War Two memories – written by the public, gathered by the BBC*. 2014. Disponível em: <https://www.bbc.co.uk/history/ww2peopleswar/>. Acesso em: 23 jun.2023.

BONNEY, R *et al.* Citizen science. Next steps for citizen science. **Science. New York, Science**, v. 343, n. 6178, p. 1436–1437, 2014. Disponível em: <https://www.science.org/doi/10.1126/science.1251554>. Acesso em 08 abr. 2023.

BORDA, A.; Gray, K.; Fu, Y. Research data management in health and biomedical citizen science: Practices and prospects. **JAMIA open**, v.3, n.1, Dec. 2019, p. 113-

125. Disponível em: <https://academic.oup.com/jamiaopen/article/3/1/113/5670773>. Acesso em: 12 jul. 2023.

BORGMAN, C. L. *Research data: who will share that, with whom, when, and why?* **Ratswd Working Paper**, Berlin, n. 161, p.1-21, out. 2010. Disponível em: https://www.ratswd.de/download/RatSWD_WP_2010/RatSWD_WP_161.pdf. Acesso em 12 abr. 2023.

BRUYERE, B.; RAPPE, S. *Identifying the motivations of environmental volunteers.* **Journal of Environmental Planning and Management**, v. 50, n. 4, 2007. p. 503-516. Disponível em: DOI: [10.1080/09640560701402034](https://doi.org/10.1080/09640560701402034).

CHAN, L.; OKUNE, A.; SAMBULI, N. O que é ciência aberta e colaborativa, e que papéis ela poderia desempenhar no desenvolvimento? In: ALBAGLI, S.; MACIEL, M. L.; ABDO, A. H. **Ciência aberta, questões abertas**. Brasília: IBICT; Rio de Janeiro: UNIRIO, p. 93-119, 2015. Disponível em: <http://livroaberto.ibict.br/handle/1/1060>. Acesso em: 03 jun. 2023.

CIVIS - **Plataforma de Ciência Cidadã**. Disponível em: <https://www.civis.ibict.br/>. Acesso em: 12 jun. 2023.

COHEN, C. M *et al.* *Identifying Opportunities in Citizen Science for Academic Libraries.* **ISTL – Issues in Science and Technology Librarianship**. [s.l.], n. 79, 2015. Disponível em: <http://www.istl.org/15-winter/article1.html>. Acesso em: 02 abr. 2023.

CONDIT, R *et al.* *Complete data from the Barro Colorado 50-ha plot: 423617 trees, 35 years*, **Dryad**, Dataset, 2019. Disponível em: <https://datadryad.org/stash/dataset/doi:10.15146/5xcp-0d46>. Acesso em 01: ago. 2023.

CONRAD, C. C.; HILCHEY, K. G. *A review of citizen science and community-based environmental monitoring: issues and opportunities.* **Environmental Monitoring and Assessment**, v. 176 (1-4), p. 273-291, 2010. Disponível em: <https://link.springer.com/article/10.1007/s10661-010-1582-5>. Acesso em: 09 jun. 2023.

COSTA, I. A. P. M. **Ciência Cidadã: envolvimento do público na investigação e divulgação em astronomia**. 2020. 689 f. Tese (Doutorado em Ensino e Divulgação das Ciências) - Faculdade de Ciências da Universidade do Porto, Porto, 2020. Disponível em: <https://repositorio-aberto.up.pt/bitstream/10216/127896/2/409652.pdf>. Acesso em: 09 jun. 2022.

CRALL, A. W *et al.* *Improving and integrating data on invasive species collected by citizen scientists.* **Biological Invasions**, v. 12, n.3, mar. p. 531-539, 2010. DOI: <https://doi.org/10.1007/s10530-010-9740-9> Acesso em: 18 jun. 2022.

CROWDSOURCING WEEK. **O que é Crowdsourcing?** USA. 2019. Disponível em: <https://crowdsourcingweek.com/about/>. Acesso em: 20 jun. 2022.

CULTURA - UNESCO. **Recomendação da UNESCO sobre Ciência Aberta**. Distrito Federal: Representação da UNESCO no Brasil, 2022. Disponível em: https://unesdoc.unesco.org/ark:/48223/pf0000379949_por. Acesso em: 17 jul. 2023.

EAVES, D. **The three laws of Open Government Data**. Califórnia, USA, 2009. Disponível em: <https://eaves.ca/2009/09/30/three-law-of-open-government-data/>. Acesso em: 12 ago. 2022.

ECSA (*European Citizen Science Association*). 2015. **Ten Principles of Citizen Science**. Berlin. DOI: <http://doi.org/10.17605/OSF.IO/XPR2N>.

EITZEL, M. V *et al.* **Citizen Science Terminology Matters: Exploring Key Terms. Citizen Science: Theory and Practice**. v. 2, n. 1, p.1. 2017. DOI: <http://doi.org/10.5334/cstp.96>.

EKLUND, L.; STAMM, I.; LIEBERMANN, W. K. *The crowd in crowdsourcing as a pragmatic research method*. **First Monday**. Chicago, v. 24, n. 10, Oct. 2019. DOI: <https://doi.org/10.5210/fm.v24i10.9206>.

ELLENBURG, J. A *et. al.* *Global Ozone (GO3). Project and AQTreks: Use of evolving technologies by students and citizen scientists to monitor air pollutants, Atmospheric Environment: X*, v. 4, 2019. DOI: <https://doi.org/10.1016/j.aeaoa.2019.100048>.

ESTELLÉS-AROLAS, E.; GONZÁLEZ-LADRÓN-DE-GUEVARA F. Rumo a uma definição integrada de *crowdsourcing*. **Journals of Information Science**. 2. ed., v. 38, p. 189-200, Mar. 2012. DOI: doi.org/10.1177/0165551512437.

EU-CITIZEN.SCIENCE. **Green Paper on Citizen Science**. [s.l.], 2013. Disponível em: <https://eu-citizen.science/faq/>. Acesso em: 27 jul. 2022.

FORCE 11. *The FAIR data principles*. **NBDC Research Team**. 2016. Disponível em: <https://force11.org/info/the-fair-data-principles/>. Acesso em: 10 ago. 2022.

FOSTER. **O que é o Open Science?** Introdução. Disponível em: <https://www.fosteropenscience.eu/content/what-open-science-introduction>. Acesso em: 02 jun. 2022.

FREITAS, H. M. R.; CUNHA JUNIOR, M. V. M.; MOSCAROLA, J. Aplicação de sistemas de *software* para auxílio na análise de conteúdo. **Revista de Administração da USP**, São Paulo, v. 32, n. 3, p. 97-109, 1997. Disponível em:

http://gianti.ea.ufrgs.br/files/artigos/1997/1997_052_RAUSP_Freitas_Cunha_Moscarola.pdf. Acesso em 18 ago. 2022.

FRITZ, S.; SEE, L.; GREY, F. *The grand challenges facing environmental citizen science*. **Frontiers in Environmental Science**, v.10, 2022. Disponível em: <https://www.frontiersin.org/articles/10.3389/fenvs.2022.1019628/full>. Acesso em: 23 jun. 2023.

FUNDAÇÃO OSVALDO CRUZ. Disponível em: <https://portal.fiocruz.br/>. Acesso em: 23 ago. 2022.

GABRIEL JUNIOR, R. F *et al.* Acesso aberto a dados de pesquisa no Brasil: mapeamento de repositórios, práticas e percepções dos pesquisadores e tecnologias. **Ciência da Informação**, [S. l.], v. 48, n. 3, 2019. Disponível em: <http://revista.ibict.br/ciinf/article/view/4958>. Acesso em: 27 ago. 2022.

GALVÃO, T. F.; PANSANI, T. DE S. A.; HARRAD, D. **Principais itens para relatar Revisões sistemáticas e Meta-análises: A recomendação PRISMA**. *Epidemiologia e Serviços de Saúde*, v. 24, n. 2, p. 335–342, jun. 2015. DOI: <https://doi.org/10.5123/S1679-49742015000200017>. Acesso em: 04 ago. 2022.

GANZEVOORT, W *et al.* *Sharing biodiversity data: citizen scientists' concerns and motivations*. *Biodivers Conserv*, v. 26, p. 2821–2837, 2017. Disponível em: <https://link.springer.com/article/10.1007/s10531-017-1391-z>. Acesso em: 10 jul. 2023.

GERHARDT, T. E.; SILVEIRA, D. T. (org.). **Métodos de Pesquisa**. Porto Alegre: Editora da UFRGS, 2009. Disponível em: <http://hdl.handle.net/10183/52806>. Acesso em: 16 ago. 2022.

GIDE, J. F. **Ciência Colaborativa: tudo o que você precisa saber**. São Paulo, 2019. Disponível em: <https://fia.com.br/blog/ciencia-colaborativa/>. Acesso em 12 jul. 2022.

GIL, A. C. **Como elaborar projetos de pesquisa**. 5.ed. São Paulo: Atlas, 2010.

GREY, F. *Viewpoint: The age of citizen cyberscience*. **CERN Courier**. v. 29, 2009. Disponível em: <http://cerncourier.com/cws/article/cern/38718>. Acesso em: 17 jul. 2022.

GUANAES, P. C. V. **Abertura e compartilhamento de dados de pesquisa subjacentes a artigos científicos: questões do direito autoral**. Tese (Doutorado em Ciência da Informação) - Instituto Brasileiro de Informação em Ciência e Tecnologia; Universidade Federal do Rio de Janeiro/Escola de Comunicação, Rio de Janeiro, 2020. Disponível em:

https://www.arca.fiocruz.br/bitstream/icict/46188/4/Tese_Paulo%20Cezar%20Vieira%20Guanaes.pdf. Acesso em: 07 jun. 2022.

HAKLAY, M. **Citizen Science & SDG**, Berlin 14-15 de outubro –hybrid conference. 2020. Disponível em: <https://povesham.wordpress.com/2020/10/17/citizen-science-sdg-berlin-14-15-oct-a-hybrid-conference/>. Acesso em: 03 jan. 2022.

HAKLAY, M. *Citizen science and volunteered geographic information: overview and typology of participation*. IN: SUI, D.; Elwood, S.; Goodchild, M.(eds) **Crowdsourcing geographic knowledge**. Springer: Dordrecht, p. 105-122, 2013. Disponível em: https://www.researchgate.net/publication/278640308_Citizen_Science_and_Volunteered_Geographic_Information_Overview_and_Typology_of_Participation. Acesso em: 23 jun. 2022.

HAKLAY, M et al. *Contours of citizen science: a vignette study*. **Royal Society Open Science**. 8.ed. v.8, 2021. Disponível em: <https://royalsocietypublishing.org/doi/10.1098/rsos.202108>. Acesso em: abr. 2022.

HEIGL, F et al. *Opinion: Toward an international definition of citizen science*. **Proc. National Academy of Sciences**. USA. v. 116, n. 17, p. 8089–8092. 2019, abr. 2019. DOI: [10.1073/pnas.1903393116](https://doi.org/10.1073/pnas.1903393116). Disponível em: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6486707/>. Acesso em: 02 an. 2022.

HOWE, J. (2006a) ‘Crowdsourcing: A Definition’, *Crowdsourcing: Tracking the Rise of the Amateur (weblog, 2 June)*, URL (accessed 24 November 2006): Disponível e: http://crowdsourcing.typepad.com/cs/2006/06/crowdsourcing_a.html. Acesso em 23 abr. 2023.

IGNAT, T. E AYRIS, P. *Built to last! Embedding open science principles and practice into European universities*. *Insights: the UKSG journal*, v. 33, p. 9, 2020. Disponível em: <https://insights.uksg.org/articles/10.1629/uksg.501>. Acesso em: 23 mar. 2023.

IGNAT, T et al. *Merry work: libraries and citizen science*. **Insights-the UKSG Journal**. v. 31, n. 35, 2018. Disponível em: <https://insights.uksg.org/articles/10.1629/uksg.431/>. Acesso em 06 fev. 2022.

INSTITUTO BRASILEIRO DE CIÊNCIA E TECNOLOGIA. **IBICT lança manifesto de Acesso Aberto a dados de pesquisa brasileira para ciência cidadã**. Rio de Janeiro, 2016. Disponível em: <https://ibict.br/sala-de-imprensa/noticias/item/478-ibict-lanca-manifesto-de-acesso-aberto-a-dados-da-pesquisa-brasileira-para-ciencia-cidada>. Acesso em: 26 ago. 2022.

IRWING, A. *Citizen Science: a study of people, expertise and sustainable development*. Londres: Routledge, 1995.

JORAYEV, G. *Monumento Kuyruktobe*, Otrar Oasis, fotografia aérea de 2016 e derivados digitais. **University College London**. Conjunto de dados, 2023.

Disponível em:

https://rdr.ucl.ac.uk/articles/dataset/Kuyruktobe_monument_Otrar_Oasis_aerial_photography_of_2016_and_digital_derivatives/22701175/1. Acesso em: 25 jun. 2023.

JORGE, V. A.; ALBAGLI, S. Compartilhamento de dados de pesquisa em saúde: iniciativas do *National Institutes of Health* (NIH). **Revista Eletrônica de Comunicação, Informação e Inovação em Saúde**, Rio de Janeiro, v. 12, n. 4, 2018. Disponível em: <https://brapci.inf.br/index.php/res/v/130562>. Acesso em: 15 jul. 2022.

KAUR, J *et al.* *Scholarometer: A social framework for analyzing impact across disciplines*. **PLoS ONE**, v. 7, n. 9, 2012, p.1-13. DOI: 10.1371/journal.pone.0043235
KITCHENHAM, B.; CHARTERS, S. **Guidelines for performing systematic literature reviews in software engineering**. nº EBSE 2007-001. [s.l.]: Keele University and Durham University Joint Report, 2007. Disponível em: https://www.elsevier.com/_data/promis_misc/525444systematicreviewsguide.pdf. Acesso em 02 mar. 2022.

KOCHE, J. C. **Fundamentos da metodologia científica**: teoria da ciência e iniciação à pesquisa. Petrópolis: Editora Vozes, 2011.

LAENDER, A. H. F *et al.* Abertura e gestão de dados: desafios para a ciência brasileira. Rio de Janeiro: **Academia Brasileira de Ciências**, 2020. Disponível em: <https://www.arca.fiocruz.br/handle/icict/43346>. Acesso em: 06 jul. 2022.

LEAL, L. R. C. **Influência e contribuições da Ciência Cidadã na Ciência da Informação**. Orientador: Rene Faustino Gabriel Junior, 2019. 84 f. Trabalho de conclusão de graduação - Curso de Biblioteconomia, Faculdade de Biblioteconomia e Comunicação, Universidade Federal do Rio Grande do Sul, Porto Alegre, 2019. Disponível em: <http://hdl.handle.net/10183/212445>. Acesso em: 15 ago. 2022.

LEAL, L. R. C.; SILVA, L. M.; GABRIEL JUNIOR, R. F. Mapeamento dos projetos brasileiros em Ciência Cidadã e biodiversidade. In: **Actas de la X Conferencia Internacional de Bibliotecas y Repositorios Digitales** (BIREDIAL-ISTEC), Modalidade Virtual, out. 2021. p. 150-170. Disponível em: <http://sedici.unlp.edu.ar/handle/10915/136534>. Acesso em: 23 jul. 2022.

MARCONI, M. A.; LAKATOS, E. M. **Fundamentos de metodologia científica**. 8. ed. São Paulo: Atlas, 2017.

MARTINS, B. **Afinal, o que é Ciência Cidadã?** Em rede. 2017. Disponível em: <http://www.em-rede.com/site/ciencia-aberta/afinal-o-que-%C3%A9-ci%C3%A9ncia-cidad%C3%A3>. Acesso em 10 jul. 2022.

MAUND, P. R *et al.* *What motivates the masses: Understanding why people contribute to conservation citizen science projects.* **Biological Conservation**. v. 246, Jun. 2020. Disponível em: DOI: <https://doi.org/10.1016/j.biocon.2020.108587>.

MCKINLEY, D.C *et al.* *Citizen science can improve conservation science, natural resource management, and environmental protection.* **Biological Conservation**. v. 208, p. 15-28, 2017. DOI: <https://doi.org/10.1016/j.biocon.2016.05.015>.

MINAYO, M. C. S. (Org.). **Pesquisa social: teoria, método e criatividade**. 26. ed. Petropolis: Vozes, 2007.

MOCZEK, N *et al.* *A Self-Assessment of European Citizen Science Projects on Their Contribution to the UN Sustainable Development Goals (SDGs).* *Sustainability*, v. 13, n. 4:1774, 2021. Disponível em: <https://www.mdpi.com/2071-1050/13/4/1774>. Acesso em: 02 jul. 2023.

MOCZEK, N. *Collection: A Self-Assessment of European Citizen Science Projects on Their Contribution to the UN Sustainable Development Goals (SDGs).* [Datasef] Data Publisher: **Museum für Naturkunde Berlin (MfN)** – Leibniz Institute for Evolution and Biodiversity Science, 2020. DOI: <https://doi.org/10.7479/mzy9-wp90>.

MONS, B *et al.* *The FAIR Guiding principles for scientific data management and stewardship.* **Scientific Data**. v. 3, n. 160018, mar. 2016. Disponível em: <https://www.nature.com/articles/sdata201618#citeas>. Acesso em: 01 ago. 2022.

MORAIS, A. M. M. **Extracting behavioral profiles from citizen science usage logs**. Orientador: Rafael Duarte Coelho dos Santos. 2016. 104 p. Dissertação (Mestrado em computação Aplicada. Programa de Pós-Graduação do INPE em Computação Aplicada. INPE: São José dos Campos, 2016. Disponível em: <http://urlib.net/sid.inpe.br/mtc-m21b/2016/07.06.18.43>. Acesso em 05 fev. 2021.

MORESI, E. A. D *et al.* *O emprego do aplicativo Sci-Hub em projetos de ciência cidadã.* *Revista Iberoamericana de Sistemas, Cibernética e Informática*, v. 14, n. 2, p. 45-52. 2017. Disponível em: <http://www.iiisci.org/Journal/riSCI/FullText.asp?var=&id=CA301DY17>. Acesso em: 23 abr. 2023.

NASA SCIENCE. **Citizen Science**. 2021. Disponível em: <https://science.nasa.gov/citizenscience>. Acesso em 03 ago. 2022.

NATIONAL SCIENCE BOARD. *Long-lived digital data collections: enabling research and education in the 21st century.* **National Science Foundation**, Arlington, Set. 2005. Disponível em: <https://repository.arizona.edu/handle/10150/105473>. Acesso em: 23 jul. 2020.

NEVES, E. B. Escrevendo a metodologia do estudo. In: NEVES, Eduardo Borba; DOMINGUES, Clayton Amaral (Org.). **Manual de metodologia da pesquisa científica**. Rio de Janeiro: EB/CEP, p. 45-67, 2007.

NPCRS - Núcleo de pesquisa clínica do RS. **Voluntários de pesquisa**. Porto Alegre, 2018. Disponível em: <https://www.npcrs.com.br/pesquisa-clinica/voluntarios-de-pesquisa/>. Acesso em 20 ago. 2022.

OECD – *OECD Principles and Guidelines for Access to Research Data from Public Funding*. **OECD Publishing**, Paris, 2007. DOI: <https://doi.org/10.1787/9789264034020-en-fr>

OLIVEIRA, A. J. Cientista eu? Como cidadãos comuns também podem fazer ciência. **Galileu Digital**, jul. 2015. Disponível em: <https://revistagalileu.globo.com/Revista/noticia/2015/07/cientista-eu-como-cidadaos-comuns-tambem-podem-fazer-ciencia.html>. Acesso em: 23 abr. 2022.

OPEN GOVERNMENT DATA. **The Annotated 8 Principles of Government Data**. 2007. Disponível em: <https://opengovdata.org/>. Acesso em 23 jun. 2023.

OPEN KNOWLEDGE FOUNDATION. **We are open advocates**. 2004. Disponível em: <http://okfn.org/about/>. Acesso em: 01 dez. 2021.

OPEN KNOWLEDGE FOUNDATION. **Open Definition: defining open in open data, open content and open knowledge**. Versão 2.1. [S. l.: s. n.], [2020?]. Disponível em: <http://opendefinition.org/od/2.1/en/>. Acesso em 18 jul. 2022.

OPEN KNOWLEDGE INTERNATIONAL. O que são dados abertos? In: *Open Knowledge International*. **Open Data Handbook**. [s.l.]: Open Knowledge International], [2020?]. Disponível em: https://opendatahandbook.org/guide/pt_BR/. Acesso em: 17 jul. 2022.

OpenDACHS – **OpenDACHS: A Citation Repository for the Sustainable Archiving of Cited Online Sources**. *The Heidelberg Centre for Transcultural Studies*. 2020. Disponível em: <http://www.asia-europe.uni-heidelberg.de/index.php?id=4425>. Acesso em: 01 jul. 2023.

OPENSIENTIST. **Finalizando uma definição de “Ciência cidadã” e “Cientistas Cidadãos”**. Washington DC, 2011. Disponível em: <http://www.openscientist.org/2011/09/finalizing-definition-of-citizen.html>. Acesso em: 01 mar. 2023.

OPENDOAR. **Directory of Open Access Repositories**. Disponível em: <https://v2.sherpa.ac.uk/opensoar/>. Acesso em: 02 jul. 2023.

PARRA, H. Z. M. Ciência cidadã: modos de participação e ativismo informacional. In: ALBAGLI, S.; MACIEL, M.L.; ABDO, A. (Org). **Ciência aberta, questões abertas**. Brasília: IBICT; Rio de Janeiro: UNIRIO, p. 121-143, 2015. Disponível em: <http://livroaberto.ibict.br/handle/1/1060>. Acesso em: 12 fev. 2022.

PARSONS, M. J. G *et al.* Sounding the Call for a Global Library of Underwater Biological Sounds. *Frontiers in Ecology and Evolution*. v. 10, 2022. DOI: 10.3389/fevo.2022.810156.

PERKINS, N. I. **Realising the potential of citizen science**. SciDevNet, sep. 2013. Disponível em: <https://www.scidev.net/global/editorials/realising-the-potential-of-citizen-science/>. Acesso em 07 jan. 2022.

PRICE-JONES, V *et al.* *Eyes on the aliens: citizen science contributes to research, policy and management of biological invasions in Europe*. **NeoBiota**, ed. 78, 2022. p. 1-24. <https://doi.org/10.3897/neobiota.78.81476>

RICE, R.; SOUTHALL, J. *The Data Librarian's Handbook*. [Facet Publishing Book Webpage]. **Facet Publishing**. 2016. Disponível em: DOI: <https://doi.org/10.25607/OBP-1539>. Acesso em 04 jul. 2022.

RIES, L.; OBERHAUSER, K, A. *Citizen Army for Science: Quantifying the Contributions of Citizen Scientists to our Understanding of Monarch Butterfly Biology*. **BioScience**, v. 65, n. 4, p. 419–430, 2015. Disponível em: <https://academic.oup.com/bioscience/article/65/4/419/255008>. Acesso em: 23 jun. 2022.

RNP – Rede Nacional para Educação e Pesquisa. **O que é Ciência Aberta e como ela pode facilitar a vida de cientistas**. 2021. Disponível em: <https://www.rnp.br/noticias/o-que-e-ciencia-aberta-e-como-ela-pode-facilitar-vida-de-cientistas>. Acesso em: 10 out. 2022.

ROLLO, M. F. Desafios e responsabilidades das humanidades digitais: preservar a memória, valorizar o patrimônio, promover e disseminar o conhecimento, o programa memória para todos. **Estudos históricos**, Rio de Janeiro, v.33, N. 69, 2020, p. 19-44. DOI:10.1590/S2178-149420200001000003. Disponível em: <https://www.scielo.br/j/eh/a/5qB3jG5kdsL3MS5pVBrfHzn/?lang=pt#>. Acesso em 02 jan. 2022.

ROMAR, C. T. M.; LENZA, P (org). **Direito do trabalho**. 5. ed. São Paulo: Saraiva Educação, 2018. Disponível em: <https://www.direitonet.com.br/dicionario/exibir/2044/Trabalho-voluntario>. Acesso em: 23 jun. 2022.

RYAN, S. F *et al.* *The role of citizen science in addressing grand challenges in food and agriculture research. **Proceedings of Royal Society B**, [s.l.], v. 285, n.1891, Nov. 2018. DOI: <https://doi/10.1098/rspb.2018.1977>.*

SAYÃO, L. F.; SALES, L. F. **Guia de gestão de dados de pesquisa para bibliotecários e pesquisadores**. Rio de Janeiro: CNEN, 2015. Disponível em: https://www.researchgate.net/publication/286455028_Guia_de_gestao_de_dados_d_e_pesquisa_para_bibliotecarios_e_pesquisadores/citation/download. Acesso em: 21 jun. 2022.

SAYÃO, L. F.; SALES, L. F. Algumas considerações sobre os repositórios digitais de dados de pesquisa. **Informação & Informação**, Londrina, v. 21, n. 2, p. 90–15, 2016. Disponível em: <https://brapci.inf.br/index.php/res/v/34650>. Acesso em: 04 jul. 2022.

SERRANO, F.; SANZ, F.; SCHAEFER, T.; SILVA, C.; KIESLINGER, B. **White Paper on Citizen Science**. (2014). Disponível em: https://www.researchgate.net/publication/271130935_White_Paper_on_Citizen_Science. Acesso em: 27 abr. 2023.

SILVA JÚNIOR, L. P.; SANTOS, T. H. N. Repositórios de dados científicos: um panorama teórico-prático. In: DIAS, G. A.; OLIVEIRA, B. M. J. F. (Org). **Dados científicos: perspectiva e desafios** - João Pessoa: Editora UFPB, 2019.

SILVA JÚNIOR, L. P.; BORGES, M. M. Preservação digital no Repositório Científico de Acesso Aberto de Portugal. **Revista Eletrônica de Comunicação, Informação & Inovação em Saúde**, v. 8, n. 4, p. 567-574, 2014. Disponível em: <http://www.reciis.iciict.fiocruz.br/index.php/receis/article/view/911>. Acesso em: 04 jul. 2022.

SILVEIRA, J. I. **Acesso aberto a dados de pesquisa em repositórios universitários internacionais: um estudo sobre políticas de depósito, acesso e uso**. Orientador: Samile Andrea de Souza Vanz. 2022. 145 f. Dissertação (Mestrado) - Programa de Pós-Graduação em Ciência da Informação, Faculdade de Biblioteconomia, Universidade Federal do Rio grande do Sul, Porto Alegre, 2018. Disponível em: <http://hdl.handle.net/10183/222095>. Acesso em: 22 jul. 2022.

SOARES, M. D., SANTOS, R. Ciência Cidadã: o envolvimento popular em atividades científicas. **Ciência Hoje**, v. 47. 2011. Disponível em: https://www.researchgate.net/publication/236952978_Ciencia_Cidada_o_envolvimen_to_popular_em_atividades_cientificas. Acesso em: 20 fev. 2021.

SOUZA, L. A. G. E. S.; ARAÚJO, E.; RIBEIRO, J. R. Ciência aberta: os desafios para concepção da ciência cidadã. **II Workshop de Informação, Dados e Tecnologia**. João Pessoa: Universidade Federal da Paraíba. 2018, p. 606-613. Disponível em: <http://hdl.handle.net/1843/39749> Acesso em: 25 fev. 2022.

SPIPOLL. Disponível em: <https://www.spipoll.org/>. Acesso em: 04 Jul. 2023.

SUN, J *et al.* *Share Your Data Carefree: An Efficient, Scalable and Privacy-Preserving Data Sharing Service in Cloud Computing.* **IEEE Transactions on Cloud Computing**, v. 11, n. 1, p. 822-838, 2021. DOI: 10.1109/TCC.2021.3117998.

TAKASHINA, N.; ECONOMO, E. P. *Developing generalized sampling schemes with known error properties: the case of a moving observer.* **Ecography**, v.44, ed. 2, p. 293-306, 2021. Disponível em: <https://doi.org/10.1111/ecog.05198>.

TEIXEIRA, L. A.; COSTA, L.T.; ZAHER, E. H. Vital Brazil: um pioneiro na prática da Ciência Cidadã. **Cadernos de História da Ciência**, v.10, n. 1, 2014. p. 33-55. DOI: <https://doi.org/10.47692/cadhisciecienc.2014.v10.33904>. Acesso em: 13 jun. 2022.

TRIVIÑOS, A. N. S. **Introdução à pesquisa em ciências sociais**: a pesquisa qualitativa em educação. São Paulo: Atlas, 1987.

VERHULST, S. *Crowdcrafting.* **The Living Library**. 2016. Disponível em: <https://thelivinglib.org/crowdcrafting/>. Acesso em: 25 jan. 2022.

WOLF, M *et al.* *Best Practices in Data Management at Ocean Networks Canada: a Citizen Scientist case study.* **OCEANS 2019 MTS/IEEE SEATTLE**, Seattle, WA, USA, 2019, p. 1-6. DOI: 10.23919/OCEANS40490.2019.8962800.

WOLKORTE, R.; HEESINK, L.; MICHELLE M.A. *As open as possible, as closed as necessary: how to find the right balance in sharing citizen science data for health? In: Proceedings of Science*, v. 418, n. 28, 2022. Disponível em: <https://pos.sissa.it/418/028/pdf>. Acesso em: 23 jun. 2023.

WU, Y.; WASHBOURNE, V.; HAKLAY, M. *Citizen science in China's water resources monitoring: current status and future prospects,* **International Journal of Sustainable Development & World Ecology**, v. 29, n. 3, 2022. p.277-290. DOI: <https://doi.org/10.1080/13504509.2021.2013973>. Disponível em: <https://www.tandfonline.com/doi/10.1080/13504509.2021.2013973>. Acesso em: 03 abr. 2022.

YAMAMOTO, E. Programa Ciência Cidadã premia oito projetos para aproximar a pesquisa da comunidade. **Jornal da USP**, São Paulo, 22 jun. 2020, p. 14. Disponível em: <https://jornal.usp.br/institucional/oito-projetos-que-aproximarao-a-pesquisa-cientifica-da-comunidade/>. Acesso em: 18 jan. 2022.

APÊNDICE A

Tabela 3 - Identificação dos Títulos por meio do ID disponibilizado pelo *software StArt*

| ID Start | Artigo |
|----------|--|
| 2 | Community-driven citizen science approach to explore cardiovascular disease risk perception, and develop prevention advocacy strategies in sub-Saharan Africa: a programme protocol. |
| 3 | Opportunities for Understanding MS Mechanisms and Progression With MRI Using Large-Scale Data Sharing and Artificial Intelligence. |
| 13 | Built to last! Embedding open science principles and practice into European universities. |
| 16 | A self-assessment of European citizen science projects on their contribution to the UN sustainable development goals (SDGs). |
| 17 | Developing generalized sampling schemes with known error properties: the case of a moving observer. |
| 18 | Sharing Data Not with Us! Distrust as Decisive Obstacle for Public Authorities to Benefit From Sharing Economy. |
| 19 | Design and use of crowdsourcing geographic data collection system during the monitoring phase of environmental impact assessment reports. |
| 22 | Research data management challenges in citizen science projects and recommendations for library support services. A shopping review and case study. |
| 27 | Data Sharing, Citizen science; Content analysis; Data repositories; Driving forces; Literature reviews; Mutual development; Research programs; Scientific data sharing, Open Data. |
| 29 | Enhancing citizens trust in technologies for data donation in clinical research: Validation of a design prototype. |
| 31 | Privacy Protection Scheme Based on CP-ABE in Crowdsourcing-IoT for Smart Ocean. |
| 33 | Crowdsourcing; Data Sharing; Information systems; Information use; Quality management, Citizen science; Data quality; Humanities; Information quality; Science, Information analysis. |
| 37 | Crowdsourcing; Data Sharing; Information systems; Information use; Quality management, Citizen science; Data quality; Humanities; Information quality; Science, Information analysis. |
| 38 | From eDNA to citizen science: emerging tools for the early detection of invasive species. |
| 39 | Does design matter when visualizing Big Data? An empirical study to investigate the effect of visualization type and interaction use |
| 42 | Data Sharing in the Context of Health-Related Citizen Science. |
| 43 | Toward Secure Data Sharing for the IoV: A Quality-Driven Incentive Mechanism with On-Chain and Off-Chain Guarantees. |
| 44 | Citizen science and invasive alien species: An analysis of citizen science initiatives using information and communications technology (ICT) to collect invasive alien species observations. |
| 48 | Citizen science in community-based watershed management: An institutional analysis in Ethiopia. |

| | |
|-----|--|
| 49 | A New paradigm in health research: FAIR Data (Findable, Accessible, Interoperable, Reusable) [Um Novo Paradigma em Investigação em Saúde: Dados FAIR (Localizáveis, Acessíveis, Interoperáveis, Reutilizáveis)]. |
| 52 | Modeling observational crowdsourcing. |
| 57 | Responsible processing of crowdsourced tourism data. |
| 59 | Research data management in health and biomedical citizen science: Practices and prospects. |
| 65 | Traffic Congestion Estimation Based on Crowd-Sourced Data. |
| 67 | Best practices in data management at ocean networks Canada: A citizen scientist case study. |
| 69 | Global Ozone (GO3) Project and AQTreks: Use of evolving technologies by students and citizen scientists to monitor air pollutants. |
| 70 | Tradeoffs and tools for data quality, privacy, transparency, and trust in citizen Science. |
| 71 | Opportunities and challenges in interpreting and sharing personal genomes. |
| 74 | Blockchain-based secure and fair crowdsourcing scheme. |
| 75 | Open Humans: A platform for participant-centered research and personal data exploration. |
| 76 | Blockchain-enabled data collection and sharing for industrial iot with deep reinforcement learning. |
| 81 | Public participation, trust and data sharing: gardens as hubs for citizen science and environmental health literacy efforts. |
| 89 | How citizens can encourage scientific research: The case study of bottlenose dolphins monitoring. |
| 93 | Toward a participatory VGI methodology: crowdsourcing information on regional food assets. |
| 100 | A social collaborative urban distribution integration platform. |
| 107 | Privacy-based recommendation mechanism in mobile participatory sensing systems using crowdsourced users preferences. |
| 108 | Defining the role of libraries in the Open Science landscape: A reflection on current European practice. |
| 112 | Sharing the archive: Using web technologies for accessing, storing and re-using historical data. |
| 113 | Big data for ecological models. |
| 115 | Social-Aware Device-to-Device Offloading Based on Experimental Mobility and Content Similarity Models. |
| 118 | A conceptual building-block and practical openstreetmap-interface for sharing references to hydrologic features. |
| 120 | Barriers to sharing water quality data: experiences from the Shale Network. |
| 122 | Sharing biodiversity data: citizen scientists concerns and motivations. |
| 130 | New frontiers on open standards for geo-spatial Science. |
| 131 | Is citizen science an open science in the case of biodiversity observations? |
| 132 | Is open science the future of drug development? |
| 133 | Crowdsourcing as a tool for research: Implications of uncertainty. |
| 136 | Expert profiling for collaborative innovation: big data perspective. |
| 140 | Big data to smart data in Alzheimer's disease: Real-world examples of advanced modeling and simulation. |

| | |
|-------|---|
| 141 | The open community data exchange: Advancing data sharing and discovery in open online community Science. |
| 143 | Citizen science in risk communication in the era of ICT. |
| 144 | To share or not to share: Drivers and barriers for sharing data via online amateur weather networks. |
| 147 | A journey of citizen science data in an online environment. |
| 148 | Ethical Dimensions of Dynamic Consent in Data-Intense Biomedical Research Paradigm Shift, or Red Herring? |
| 149 | Collective knowledge and creativity: The future of citizen science in the humanities. |
| 150 | Bringing citizen monitoring into land management: A case study of the bosque ecosystem monitoring program. |
| 152 | Citizen Science on Your Smartphone: An ELSI Research Agenda. |
| 155 | Crowdsourcing the General Public for Large Scale Molecular Pathology Studies in Cancer. |
| 165 | Digital repository of associations between environmental variables: A new resource to facilitate knowledge synthesis. |
| 169 | Adoption of geospatial systems towards evolving sustainable himalayan mountain development. |
| 183 | Scholarometer: A Social Framework for Analyzing Impact across Disciplines. |
| 21364 | As open as possible, as closed as necessary: how to find the right balance in sharing citizen science data for health? |
| 21365 | Science learning in biodiversity citizen science: Inputs from the analysis of online social interactions within a contributory project for pollinators' monitoring. |
| 21366 | Citizen science in environmental and ecological sciences. |
| 21370 | Recruitment and Retention in Remote Research: Learnings From a Large, Decentralized Real-world Study. |
| 21372 | The grand challenges facing environmental citizen Science. |
| 21376 | Open Access to the Digital Biodiversity Database: A Comprehensive Functional Model of the Natural History Collections. |
| 21377 | Geomagnetism, Paleomagnetism and Electromagnetism Perspectives on Integrated, Coordinated, Open, Networked (ICON) Science. |
| 21380 | Sounding the Call for a Global Library of Underwater Biological Sounds. |
| 21382 | Eyes on the aliens: citizen science contributes to research, policy and management of biological invasions in Europe. |
| 21385 | Multilingual Research Projects: Non-Latin Script Challenges for Making Use of Standards, Authority Files, and Character Recognition. |
| 21392 | Online data sharing with virtual social interactions favor scientific and educational successes in a biodiversity citizen science project. |

Fonte: Da Autora (2023)

APÊNDICE B

Tabela 4 - Dados informativos dos artigos selecionados para estudo sobre projetos em Ciência Cidadã

| ID Start | Crowdsourcing | Projetos | Armazena Dados | Observações |
|----------|---------------|----------|----------------|--|
| 2 | | | | Descreve a necessidade de armazenamento seguro de dados de pesquisa, mas não menciona em quais repositórios. |
| 3 | | | | Menciona a importância do compartilhamento de dados, mas não menciona em qual plataforma. Dados abertos |
| 13 | | | | Enfatiza a importância da qualidade de dados em CC. Desenvolvimento do repositório RDR. |
| 16 | | | | Projeto de monitoramento hídrico, tendo como desafio a tomada de decisões nos resultados dos projetos. Armazenamento repositório do <i>Museum für Naturkunde Berlin</i> . |
| 17 | | | | Projeto de ecologia, observações de campo pelos cidadãos, resultados apresentados no <i>Smithsonian Research Online</i> . https://repository.si.edu/handle/10088/784 |
| 18 | | | | Apresenta a necessidade de compartilhamento de big data em plataformas de <i>crowdsourcing</i> , mas não mencionam quais. |
| 19 | | | | Preocupação com extração de dados geográficos em projetos de <i>crowdsourcing</i> , para solução sugerem o uso do <i>software</i> GeoServer de código aberto, que permite visualização e edição dos dados PostgreSQL. |
| 22 | | | | Apointa os desafios de RDM em CC em como sustentar a interoperabilidade de <i>software</i> ou tecnologia usada em projetos, sugerem a criação de plataformas ou <i>software</i> de código aberto, de acordo com a necessidade dos usuários. Dados abertos. |
| 27 | | | | Estudo sobre a confiabilidade dos repositórios onde os dados de pesquisa são armazenados. Citam alguns exemplos: GS Cloud. GSA e GCdataPR. Dados abertos. |
| 29 | | | | Produção de dados de pesquisa sobre saúde, por meio da doação de dados dos cidadãos em dispositivos móveis. Sugerem o uso de plataformas como Mobistudy. |
| 31 | | | | Coleta e compartilhamento de dados marinhos na Internet das Coisas. Uso de plataforma de nuvem combinadas com tecnologia de IA. Não cita exemplos. |
| 33 | | | | Análise sobre a qualidade dos dados extraídos em projetos de CC, no monitoramento hídrico. Sugerem armazenamento em banco de dados, mas não especificam quais. |
| 37 | | | | Compartilhamento de dados na área de saúde e o uso de dispositivos versáteis, como o serviço na nuvem o DataSHIELD. Dados abertos. |
| 38 | | | | Estudo sobre compartilhamento e integração de dados na detecção precoce de espécies invasoras. Exemplificam o uso de plataformas como a iNaturalist. |
| 39 | | | | Coleta de dados para experimento quantitativo em larga escala, usando a plataforma Amazon Mechanical Turk, para compartilhamento dos dados. |
| 42 | | | | Compartilhamento de dados de CC relacionados à saúde, ênfase na importância |

| | | | | |
|----|--|--|--|--|
| | | | | das plataformas como as de Citsci.org e Open Humans.org. |
| 43 | | | | Compartilhamento de dados em aplicativos para automóvel - uso da inteligência do veículo (Lov). Sugere o armazenamento em banco de dados, não menciona quais. |
| 44 | | | | Análise sobre iniciativas de CC na coleta de dados de IAS e que a falta de compartilhamento de dados em IAS gera a fragmentação desses dados. Sugerem o compartilhamento em plataformas como o CitSci. |
| 48 | | | | Estudos em coleta de dados hidrológicos e meteorológicos. <i>Feedback</i> em plataformas de Inovação (IPs), não menciona quais. |
| 49 | | | | Armazenamento de dados de pesquisa clínica, biométrica e científica. Uso de dados FAIR. Não menciona onde os dados serão armazenados. |
| 52 | | | | Estudos sobre <i>crowdsourcing</i> observacional, com foco em interfaces de usuário, coleta de dados, compartilhamento de dados e interoperabilidade. Cita como exemplo a plataforma iSpotNature. |
| 57 | | | | <i>Crowdsourcing</i> como fonte de informação para os stakeholders ⁹⁰ do turismo e o uso de Plataformas como AirBnB e TripAdvisor. |
| 59 | | | | Análise em práticas de gerenciamento de dados de pesquisa em estudos biomédicos de CC. Menciona armazenamento nas bases de dados: PubMed, Scopus, ACM Digital Library, IEEE Xplore e Web of Science e as plataformas: Hubzero, Zooniverse e PatientsLikeme. Dados abertos. |
| 65 | | | | Proposta sobre os dados de tráfego serem coletados de fontes de multidão: sistemas de IoT fixos e de detecção humana. Menciona o banco de dados MongoDB para armazenamento. |
| 67 | | | | Práticas de gerenciamento de dados ONC no <i>Salish Sea Marine Survival Project</i> . Armazenamento na base de dados PostgreSQL.org. |
| 69 | | | | Uso de tecnologias por cientistas cidadãos para monitorar poluentes atmosféricos. Uso de plataforma <i>online</i> e repositório. Armazenamento no <i>WayBack Machine</i> . |
| 70 | | | | Trata da qualidade dos dados gerados em projetos de CC. Não menciona armazenamento de dados, mas apresenta a plataforma snapshotwisconsin.org como uma opção de compartilhamento. Dados abertos. |
| 71 | | | | Promover abertura e o compartilhamento ético de dados de genoma pessoal. Apresenta o repositório openSNP.org e a plataforma Naveed Aziz, como exemplo. Dados abertos. |
| 74 | | | | Compartilhamento de dados com base em <i>blockchain</i> ⁹¹ . Não menciona onde armazenar os dados. |

⁹⁰ Grupos e indivíduos que, de uma forma ou de outra, apresentam algum nível de interesse nos projetos, atividades e resultados de uma determinada organização. <https://www.siteware.com.br/gestao-estrategica/o-que-sao-stakeholders/>.

⁹¹ Blockchain é um livro de registros, compartilhado e imutável, que facilita o processo de gravação de transações e rastreamento de ativos em uma rede de negócios. <https://www.ibm.com/br-pt/topics/blockchain>

| | | | | |
|-----|--|--|--|---|
| 75 | | | | Estudo sobre plataformas que permitem a coleta de dados pessoas, com controle de autorizações para compartilhamento. Ex. Open Humans. |
| 76 | | | | Coleta e compartilhamento de dados eficientes e seguros para a IoT. EX. o serviço na nuvem MySQL. |
| 81 | | | | Projeto de CC em locais de lixo perigoso e o envolvimento de cidadãos. Cita compartilhamento de dados por meio de <i>design</i> de informação, não informa qual. |
| 89 | | | | Monitoramento e compartilhamento de dados de redes meteorológicas. Compartilha dados em plataformas <i>online</i> , não menciona quais. |
| 93 | | | | Abordagem sistemática usada para criar VGI ⁹² . Ênfase na necessidade de repositórios de dados abertos não governamentais. Não menciona onde dos dados são hospedados, mas cita a Plataforma Carto. |
| 100 | | | | Proposta de um protótipo de plataforma como um portal de informações colaborativas e compartilhadas. Desenvolvimento de aplicativo por meio do programa R e Shiny ⁹³ . |
| 107 | | | | Preocupação com a privacidade dos usuários no compartilhamento no uso de dispositivos móveis. Desenvolvimento do app PriRe ⁹⁴ , com armazenamento no banco de dados MySQL. |
| 108 | | | | Gerenciamento de dados de algumas bibliotecas universitárias. Menciona a importância de armazenamento e cita as plataformas como Discovery, OAPEN e JSTOR. |
| 112 | | | | Planejamento de construção de banco de dados para pesquisa colaborativa, também descreve as oportunidades e desafios enfrentados pelos projetos. Embora ainda não armazenem mencionam interface na nuvem e as plataformas Wordpress e Omeka. Compartilhamento de código aberto. |
| 113 | | | | Estudo sobre arquivamento e compartilhamento de dados em repositórios de <i>big data</i> . Não menciona nenhum repositório. |
| 115 | | | | Compartilhamento de dados D2D e coleta no uso de aplicativos móveis, com armazenamento na Interface na nuvem. |
| 118 | | | | Pesquisa sobre a importância dos repositórios para armazenamento e compartilhamento. Apresenta como exemplo a plataforma Datahub ⁹⁵ . |
| 120 | | | | Resistência no compartilhamento de dados do projeto desenvolvido pela Shale Network, sobre a qualidade da água. Base de dados indicada é o <i>site</i> CUAHSI HIS ⁹⁶ . |
| 122 | | | | Pesquisa sobre voluntários e as motivações nos projetos de CC. Uso da plataforma <i>online</i> LimeSurvey ⁹⁷ para responder ao questionário |

⁹² Informações Geográficas Voluntárias

⁹³ Shiny é um *framework* em linguagem R para a criação de aplicativos da *web*. <https://programando-em-shiny.curso-r.com/>.

⁹⁴ PriRe é um mecanismo de recomendação baseado em privacidade, que pode medir a privacidade dos usuários com precisão e fornece recomendações eficazes aos usuários para compartilhamento de dados em sistemas móveis de detecção participativa.

⁹⁵ <https://datahub.io/>

⁹⁶ https://hiscentral.cuahsi.org/pub_network.aspx?n=228

⁹⁷ O LimeSurvey é uma ferramenta de pesquisa *online* para criar rapidamente questionários,

| | | | | |
|-----|--|--|--|---|
| 130 | | | | Estudo sobre a facilidade de compartilhamento de dados, face aos avanços da interface na nuvem. Citam como exemplo: Ushahidi e GitHub. |
| 131 | | | | Pesquisa sobre identificação na abertura relativa de dados sobre biodiversidade em CC. Sugestão do uso de repositórios digitais públicos, não informa quais. Dados abertos. |
| 132 | | | | Disserta sobre a colaboração e compartilhamento de dados e código aberto em estratégias no desenvolvimento de medicamentos. Cita plataformas <i>online</i> e repositórios públicos, não exemplificam. |
| 133 | | | | Análise sobre o <i>crowdsourcing</i> como uma ferramenta útil, viável ou desejável para pesquisa no uso da tecnologia baseada na <i>web</i> . Cita o Zooniverse, Amazon Mechanical Turk, como plataforma científica. |
| 136 | | | | Aborda sobre inovação colaborativa em plataformas de redes sociais de pesquisa baseada em big data. Ex. LinkedIn. |
| 140 | | | | Projeto de parceria público-privadas focado no compartilhamento de dados, inferência causal e análise baseada em caminhos, na área da saúde. Não menciona armazenamento de dados. Dados abertos. |
| 141 | | | | Desenvolvimento de infraestrutura para cientistas computacionais no avanço da compreensão das interações humanas <i>online</i> , o GenBank. Possibilidade de armazenamento no Github ⁹⁸ . Não apresenta dados. |
| 143 | | | | Comunicação descentralizada, colaborativa e baseada em rede que permite aos cidadãos gerar dados e compartilhar informações em mídias sociais, por meio do software ArcGIS ⁹⁹ . |
| 144 | | | | Incentivo aos cidadãos na coleta e compartilhamento de dados sobre o meio ambiente, por meio de plataformas. Não cita quais. |
| 147 | | | | Estudo de caso onde foram examinados os dados do projeto de CC Biocubes ¹⁰⁰ , que está armazenado no INaturalist ¹⁰¹ . |
| 148 | | | | Uso de dados de pesquisa biomédica e normas sociais de compartilhamento de dados no desenvolvimento de um <i>link</i> para projeto de CC. Mencionam biobancos, plataformas e repositórios virtuais, mas não cita quais. |
| 149 | | | | Terminologia da CC e as contribuições dos voluntários em projetos. Foco no desenvolvimento de plataformas como a <i>Citizen Cyberscience Center</i> (desativada) ¹⁰² e a OKF ¹⁰³ . |
| 150 | | | | Estudo de caso, baseado nos dados coletados por voluntários e sua qualidade. Menção a publicações em banco de dados ou em plataformas não informando quais |

enquetes de votação e pesquisas. <https://www.limesurvey.org/pt-br>

⁹⁸ <https://github.com/>

⁹⁹ <https://www.arcgis.com/index.html>

¹⁰⁰ <https://www.inaturalist.org/projects/biocubes>

¹⁰¹ <https://www.inaturalist.org/>

¹⁰² <https://github.com/citizen-cyberscience-centre>

¹⁰³ <https://okfn.org/>

| | | | | |
|-------|--|--|--|--|
| 152 | | | | Participação e contribuição dos voluntários em projetos sobre saúde. O uso de plataformas como Ape Researchkit. e o Software Sage Bionetworks. |
| 155 | | | | Projeto de CC, voltado para compartilhamento de imagens, para digitação e análise de amostras, na área da saúde. Cita o uso de plataformas, mas não cita quais. |
| 165 | | | | Estudo sobre biodiversidade e compartilhamento de dados em repositórios. Cita somente a plataforma IPBES. Dados abertos. |
| 169 | | | | Adoção de <i>drones</i> em projetos de CC para o desenvolvimento de sistemas automatizados de monitoramento ambiental. Uso de banco de dados regional (RDS), mas não informa qual. |
| 183 | | | | Uso de métricas para avaliar o impacto de publicações acadêmicas. Para a coleta e o compartilhamento de dados é sugerido o modelo de plataforma da ferramenta da <i>web</i> social, o <i>Scholarometer</i> . |
| 21364 | | | | Importância do compartilhamento de resultados de pesquisa ou dados agregados. Sugerem abertura de dados por meio de repositórios como o Zenodo. |
| 21365 | | | | Estudo sobre o impacto na participação em projetos de CC no programa Spipoll de biodiversidade. Plataforma Spipoll. |
| 21366 | | | | Iniciativa de CC envolvendo o cidadão em pesquisas para gerar resultados científicos. Ênfase na publicação dos dados em repositórios, menciona GBIF, Dryad, OSF e Zenodo. |
| 21370 | | | | Engajamento de voluntários pagos em estudo, por meio de app, acerca de saúde (resfriado). Menciona que os dados estão disponíveis no repositório de código GitHub, de código aberto, no repositório Prolific e no Amazon Mechanical Turk. |
| 21372 | | | | Combinação de CC com IA, no uso da qualidade dos dados de pesquisa ambientais, por governos e autoridades locais. Não cita onde os dados seriam compartilhados. Dados abertos. |
| 21376 | | | | Projeto de digitalização de acervo de coleções de História Natural, sobre biodiversidade e taxonomia. banco de dados AMUNATCOLL e BioCAsE Provider Software |
| 21377 | | | | Abordam a CC, o envolvimento da comunidade e compartilhamento de dados de geomagnetismo da Terra. A maioria dos dados não estão disponíveis em bancos de dados e sem descrições e estatísticas de interpretações. Repositório mencionado foi o IRIS. |
| 21380 | | | | Uso da CC em projetos de monitoramento de sons de ambientes aquáticos, e desenvolvimento de uma plataforma de acesso aberto baseada na <i>web</i> . Citam repositórios acústicos e portais de dados para armazenamento, mas não informam quais. |
| 21382 | | | | Relatos de projetos em CC sobre espécies exóticas, onde foram produzidos artigos revisado por pares e $\frac{1}{3}$ destes projetos compartilharam seus dados em repositórios como os Global Biodiversity Information Facility ou a <i>European Alien Species Information Network</i> . Dados abertos. |

| | | | | |
|-------|--|--|--|---|
| 21385 | | | | Aborda pesquisa acadêmica sobre dados de pesquisa de escrita digital não latina NLS. Estudos de caso sobre a expansão de um padrão de metadados e dos desafios na produção de textos de jornais chineses. Uso do repositório de citações OpenDACHS e Github Open Dachs. |
| 21392 | | | | Ênfase na coleta de grandes conjuntos de dados e monitoramento por longo prazo, na área de biodiversidade. Uso da plataforma Spipoll para identificação de insetos ou plantas por qualquer voluntário. |
| 21415 | | | | Aborda a CC na gestão de serviços hídricos e ecossistêmicos, assim como a importância da coleta e análise de grande quantidade de dados. Cita o uso de banco de dados, mas não especifica quais. |

Fonte: Da autora (2023)