Plan Overview

A Data Management Plan created using DMP Tool

DMP ID: https://doi.org/10.48321/D16P5M

Title: Brazilian Citizen Science Open Platform

Creator: Bruno Albertini - **ORCID:** 0000-0003-3738-6448

Affiliation: Universidade de São Paulo (www5.usp.br)

Principal Investigator: Natália Pirani Ghilardi-Lopes, Tiago Mauricio Francoy

Contributor: Natália Pirani Ghilardi-Lopes, Tiago Mauricio Francoy

Funder: São Paulo Research Foundation (fapesp.br)

Template: Digital Curation Centre (português)

Project abstract:

Citizen science has been standing out as a form of social engagement of the population in projects with a scientific nature. With the participation of scientists, the techniques and platforms applied to citizen science provide data collection that is otherwise impractical due to the geographic reach and frequency of observation, known to be high in citizen science approaches. Current platforms take advantage of mobile personal devices, notably smartphones, due to their connectivity and dispersion among the population. The two largest citizen science platforms today, eBird and iNaturalist, are limited to capturing images for a specific niche (bird watching and biodiversity observation, respectively). In Brazil, there are several citizen science initiatives, each one developed for a specific research project, with its own hypotheses and mechanisms. There is no standardization regarding the data collected, nor open access. This project proposes to meet this need by developing an open platform for citizen science without a specific niche. In addition to the platform itself, the project also defines a standard for storing and exchanging information between citizen science platforms, focused on interoperability and preservation of digital data, so that the various initiatives can port data from one platform to another.

Start date: 01-01-2023

End date: 12-31-2026

Last modified: 07-08-2024

Copyright information:

The above plan creator(s) have agreed that others may use as much of the text of this plan as they would like in their own plans, and customize it as necessary. You do not need to credit the creator(s) as the source of the language used, but using any of the plan's text does not imply that the creator(s) endorse, or have any relationship to, your project or proposal

---

Brazilian Citizen Science Open Platform

Coleta de Dados

Que dados serão coletados ou criados?

Este projeto visa implementar uma plataforma de ciência cidadã aberta, portanto cada projeto possui suas necessidades de coleta. Os dados podem ser midiáticos (áudio, foto, vídeo) ou grandezas (e.g. distâncias, pesos, posição geográfica, etc). As grandezas podem ser medidas automaticamente (através do celular, equipamento desenvolvido para tal, etc) ou manualmente (entrada pelo usuário).

Como os dados serão coletados ou criados?

Documentação e Metadados

Que documentação e metadados acompanharão os dados?

Uma das metas do projeto é propor uma padronização para definir os metadados que acompanharão os dados. De modo geral, todos os dados coletados são acompanhados por seus metadados originais (e.g. uma imagem com dados EXIF é compartilhada com os mesmos dados). O reuso dos dados segue as ontologias da padronização do projeto e de terceiros (padronizações encapsuladas no padrão do projeto). Há um wiki com a explicação dos campos.

Ética e Conformidade Legal

Como você administrará qualquer questão ética?

Todos os projetos da plataforma devem apresentar anuência por escrito da comissão de ética à qual estão submetidos ou da "Plataforma Brasil". Denúncias de violações podem ser feitas para o email beekeep@usp.br

Os dados disponibilizados publicamente serão anonimizados no sentido de não se identificar o usuário que o coletou, incluindo posição geoespacial. Os usuários serão tratados por um ID único gerado exclusivamente para este fim e os dados pessoais não são divulgados publicamente. Exceções previstas na LGPD para pesquisa científica serão tratadas pelos responsáveis pelo projeto através do email beekeep@usp.br.

Como você vai gerenciar os direitos autorais e os direitos de propriedade intelectual (IP / IPR)?

Todos os dados da plataforma são gerados pelos próprios usuários ou por coletas automatizadas e disponibilizados por licença de atribuição CC-BY 4.0. Em caso de violação de direitos de terceiros, aciona-se o mecanismo de disputa manualmente pela plataforma ou por email (beekeep@usp.br), com consequente retirada imediata e apagamento definitivo do dado após análise pela equipe.

Armazenamento e Backup

Como os dados serão armazenados e terão backup durante a pesquisa?

Os dados serão armazenados em servidor storage dedicado para tal, com provisão de armazenamento por 5 anos. O backup é semanal em mídia offline, feito automaticamente.

Como você vai gerenciar o acesso e a segurança?

O acesso à plataforma é feito mediante cadastro (email e senha) ou login federado (redes sociais). Os servidores são acessíveis somente para a equipe técnica do projeto e contam com proteção de nível comercial (SSH/TLS). Os dados pessoais ou que possam identificar um usuário não são visíveis para outros usuários (proteção por chave de sessão). A plataforma possui proteção contra Cross site scripting (XSS), Cross site request forgery (CSRF), SQL injection, Clickjacking e todo o serviço executa sobre HTTPS.

Seleção e Preservação

Quais dados são de valor a longo prazo e devem ser mantidos, compartilhados e / ou preservados?

Os dados coletados são disponibilizados abertamente na plataforma, de forma anonimizada. Dados não textuais são acessíveis mediante login na plataforma ou por solicitação à equipe, quando serão disponibilizados para download em link específico.

Qual é o plano de preservação a longo prazo do conjunto de dados?

Na eventualidade de encerramento do projeto, todos os dados serão enviados para repositórios de dados abertos, como o DataUSP, GBIF ou outro aplicável. Uma indicação do projeto será fornecida, com uma versão congelada da documentação e da descrição do formato.

Compartilhamento de Dados

Como você vai compartilhar os dados?

Os dados são compartilhados na própria plataforma. O usuário de um projeto pode baixar todos os dados do projeto de forma anonimizada, incluindo os dados gerados por outros usuários do mesmo projeto. Não há restrição de reuso.

Existem restrições ao compartilhamento de dados requeridos?

Não existe restrição de compartilhamento dos dados anonimizados. As exceções de acesso aos dados não anonimizados previstas na LGPD seguem políticas diferenciadas, definidas pelo solicitante e submetidas a análise.

Responsabilidades e Recursos

Quem será responsável pelo gerenciamento de dados?

O proponente do projeto é o responsável pelo gerenciamento geral de dados (Bruno Albertini, balbertini@usp.br). Para cada projeto da plataforma, o criador do projeto será o o responsável.

Quais recursos você precisará para entregar seu plano?

O servidor de storage e o servidor em nuvem são os únicos recursos necessários.