

Protocolo de estruturação de dados em pesquisas socioambientais: uma abordagem de ciência aberta aplicada aos Geoparques UNESCO do Rio Grande do Sul

Data structuring protocol in socio-environmental research: an open science approach applied to UNESCO Geoparks in Rio Grande do Sul

Luísa Crauss de Araujo

UFRGS, Porto Alegre, Rio Grande do Sul (RS), Brasil

Carlise Schneider Rudnicki

UFRGS, Porto Alegre, Rio Grande do Sul (RS), Brasil

Daniela Garcez

UFRGS, Porto Alegre, Rio Grande do Sul (RS), Brasil

GT04 - Meio ambiente, agroecologia, sustentabilidade e mudanças climáticas

Resumo: Este trabalho apresenta um protocolo metodológico para a estruturação de dados para o projeto de pesquisa socioambiental em andamento sobre os impactos de desastres climáticos e a governança nos territórios dos Geoparques UNESCO do Rio Grande do Sul, “Governança e multidimensionalidade dos riscos climáticos: abordagem multidisciplinar em Comunicação de proximidade, Interpretação geopatrimonial e Economia Ecológica aplicada aos Geoparques Unesco no Rio Grande do Sul” (UFRGS, UFSM, UNIPAMPA). Parte-se da problemática recorrente da desorganização de bancos de dados em pesquisas de longa duração, que compromete a reutilização e a transparência dos dados. O objetivo é descrever a construção de uma arquitetura de dados orientada pelos princípios da ciência aberta, visando garantir padronização, interoperabilidade e acesso ampliado à informação. Metodologicamente, o estudo baseia-se na organização dos dados em três níveis: dados brutos, dicionário de dados e dados analíticos, com aplicação de técnicas de codificação, padronização e integração com ferramentas analíticas. Destaca-se que a pesquisa está em andamento, com previsão de ampliação contínua da base de dados. Os resultados indicam que a estruturação prévia dos dados reduz o tempo de tratamento, aumenta a confiabilidade e amplia o potencial de reutilização. Conclui-se que a adoção de práticas alinhadas à ciência aberta fortalece a qualidade da pesquisa e amplia seu impacto científico e social.

Abstract: This paper presents a methodological protocol for data structuring for the ongoing socio-environmental research project on the impacts of climate disasters and governance in the UNESCO Geopark territories of Rio Grande do Sul, “Climate risk governance and multidimensionality: a multidisciplinary approach in proximity communication, geopatrimonial interpretation, and ecological economics applied to UNESCO Geoparks in Rio Grande do Sul” (UFRGS, UFSM, UNIPAMPA). It addresses the recurring issue of database disorganization in long-term research, which compromises data reuse and transparency. The objective is to describe the construction of a data architecture guided by the principles of open science, aiming to ensure standardization, interoperability, and expanded access to information. Methodologically, the study is based on the organization of data into three levels: raw data, data dictionary, and analytical data, with the application of coding techniques, standardization, and integration with analytical tools. It is highlighted that the research is ongoing, with a forecast for continuous expansion of the database. The results indicate that prior data structuring reduces processing time, increases reliability, and expands the potential for reuse. It is concluded that the adoption of practices aligned with open science strengthens the quality of the research and broadens its scientific and social impact.

Keywords: UNESCO Geoparks of Rio Grande do Sul; open science; database; climate change.

1. Introdução

A produção científica sobre mudanças climáticas e governança socioambiental exige estruturas de dados que permitam análises integradas e replicáveis. No Brasil, a maturidade dessa produção está diretamente ligada à criação de infraestruturas de dados transparentes

(NICODEMO, 2024). Este trabalho insere-se no projeto "Diagnóstico dos impactos de desastres climáticos nos territórios dos Geoparques UNESCO do Rio Grande do Sul" (ARAUJO et al., 2025), que investiga os efeitos de eventos extremos nos territórios da Quarta Colônia, Caçapava do Sul e Caminhos dos Cânions do Sul.

A problemática central deste artigo é a recorrente desorganização de bancos de dados em pesquisas de longa duração, o que compromete a transparência e a reutilização das informações, contrariando os princípios de Ciência Aberta (UNESCO, 2021). Como a pesquisa está em andamento e ainda prevê amplas coletas com gestores e indicadores socioambientais, o objetivo deste trabalho não é apresentar resultados finais, mas descrever o protocolo metodológico de estruturação de dados desenhado atualmente para este projeto. O foco é demonstrar como uma organização sistemática em três níveis (dados brutos, metadados e dados analíticos) garante que a informação seja útil e acessível para pesquisadores, gestores e para a sociedade civil interessada na resiliência climática dos Geoparques.

2. Revisão da Literatura

A Ciência Aberta (Open Science) consolidou-se como um paradigma que propõe a abertura dos processos de pesquisa e o compartilhamento amplo do conhecimento (UNESCO, 2021). Para além do acesso livre a publicações, esse modelo exige a disponibilização de dados e metodologias, sendo essencial para enfrentar desafios globais como as mudanças climáticas (UNESCO, 2021). Segundo Nicodemo (2024), a ciência aberta é uma "realidade sem volta", funcionando como condição indispensável para ampliar a maturidade, a transparência e a escala da produção científica contemporânea. Nesse sentido, a maturidade científica está diretamente associada à capacidade de criar metadados estruturados que permitam a compreensão dos dados independentemente do contexto original de produção (NICODEMO, 2024). Assim, a organização adequada dos dados atua como um instrumento de democratização do conhecimento, fortalecendo a qualidade da pesquisa e ampliando seu impacto social nas estratégias de governança territorial (NICODEMO, 2024; UNESCO, 2021).

3. Metodologia

A metodologia baseia-se na criação de um protocolo estruturado para o banco de dados do projeto "Diagnóstico dos impactos de desastres climáticos nos territórios dos Geoparques UNESCO do Rio Grande do Sul" (ARAUJO et al., 2025). Por ser uma pesquisa em andamento, o sistema foi desenhado para ser escalável, permitindo a entrada contínua de

novos dados (socioambientais e de governança) de forma padronizada e alinhada à Ciência Aberta (UNESCO, 2021).

3.1 Arquitetura do banco de dados

Para garantir a segurança e a organização, os dados são estruturados em três níveis de processamento:

- **Aba 00 – Dados Brutos:** Contém as informações exatamente como foram coletadas (via Google Forms ou KoboToolbox). Este nível é imutável, servindo como um arquivo de segurança para auditoria e preservação da fonte original.
- **Aba 01 – Dicionário de Dados (Metadados):** Funciona como um "tradutor" do banco. Cada pergunta é convertida em um código técnico (ex: RISCO_CLIM_COD) com explicações sobre o que cada número significa (ex: 1 = Sim). Isso permite que qualquer software ou pesquisador entenda a base sem erros de interpretação.
- **Aba 02 – Dados Analíticos:** É a base pronta para o trabalho. Aqui, os dados já passaram por limpeza e padronização, separando o conteúdo original do tratado para evitar falhas de análise.

3.2 Padronização e Pré-processamento

Para evitar erros manuais e preparar os dados para análises complexas, o protocolo utiliza técnicas automatizadas:

- **Conversão numérica:** Transforma respostas de texto (como gênero ou escolaridade) em números. Isso facilita o uso de ferramentas estatísticas e linguagens de programação como Python.
- **Tratamento de múltipla escolha:** Respostas múltiplas são desmembradas em colunas individuais de "sim" (1) ou "não" (0), permitindo calcular frequências com precisão.
- **Identificação única (ID):** Cada participante ou localidade recebe um código exclusivo, o que permite cruzar diferentes tabelas (ex: relacionar dados de um gestor com os impactos climáticos de sua região) de forma relacional.

3.3 Fluxo de Análise e Integração Multimodal

O protocolo organiza os dados para que sigam dois caminhos complementares:

- **Caminho Qualitativo:** Textos e depoimentos são enviados ao software NVivo para a identificação de temas e padrões de fala.
- **Caminho Quantitativo:** Dados numéricos são analisados via Python e podem ser integrados a bancos de dados como MySQL. Essa estrutura permite que, no futuro, os

resultados sejam exibidos em painéis dinâmicos e gráficos interativos no site do projeto, facilitando o acesso da sociedade.

3.4 Expansão Contínua

Como a pesquisa está em execução, este banco de dados será alimentado progressivamente com novas coletas sobre governança e indicadores socioambientais (ARAUJO et al., 2025). O uso deste protocolo rígido garante que todos os novos dados mantenham o mesmo padrão de qualidade e comparabilidade.

4. Discussão dos Resultados

Os resultados parciais evidenciam que a adoção de um protocolo estruturado de organização de dados contribui para a eficiência analítica e a qualidade da pesquisa, pois a definição prévia de padrões e metadados reduz o tempo de tratamento das informações, minimiza inconsistências e facilita o cruzamento de variáveis. Além disso, a estruturação dos dados amplia seu potencial de reutilização, permitindo sua integração com diferentes ferramentas analíticas e plataformas digitais. Esse aspecto é especialmente relevante no contexto da ciência aberta, que pressupõe a disponibilização de dados de forma acessível, interoperável e reutilizável (UNESCO, 2021).

5. Considerações finais (Seções numerada, com apenas a primeira letra em caixa alta, negrito, fonte Times 12, alinhado à esquerda)

A estruturação de dados constitui um elemento central na produção científica contemporânea, especialmente no contexto da ciência aberta. Conforme destaca a UNESCO (2021), a abertura dos dados depende diretamente de sua organização e documentação adequadas. Sendo assim, conclui-se que a adoção de práticas de ciência aberta, como as apresentadas neste trabalho, aliadas à organização sistemática dos dados, fortalece a qualidade metodológica da pesquisa, com o objetivo de ampliar seu potencial de contribuição para o entendimento e enfrentamento das mudanças climáticas.

Referências

ARAUJO, Luísa Crauss de et al. Diagnóstico dos impactos de desastres climáticos nos territórios dos Geoparques UNESCO do Rio Grande do Sul. 2025.

NICODEMO, Thiago Lima. Dados abertos de pesquisa no Brasil: diagnóstico e perspectivas futuras. 2024.

UNESCO. Recommendation on Open Science. Paris: UNESCO, 2021.