

Co-Creating a Research Data Infrastructure with Social Policy Researchers

Research Question

What to consider when applying co-creation as a design methodology to create a data infrastructure system for social policy researchers?

Design Requirements

Over three years, we co-designed a collaborative system to support the harmonization, validation, exploration, and sharing of their research data.

Through co-creation workshops and interviews, we elicited and refined the design requirements:

- R1 Support data harmonization.
- R2 Support data validation.
- R3 Enable interactive data exploration.
- R4 Allow flexible sharing of data and resources.

The Information System

- To support data harmonization (R1), we established *Data and metadata standards*.
- We implemented a data validation pipeline (R2) to thoroughly check if the standards are fulfilled.
- We designed three interfaces for exploring the data (R3): (1) the Indicator page, shown in Fig. 1, (2) the Country profile, and (3) the Data Explorer.

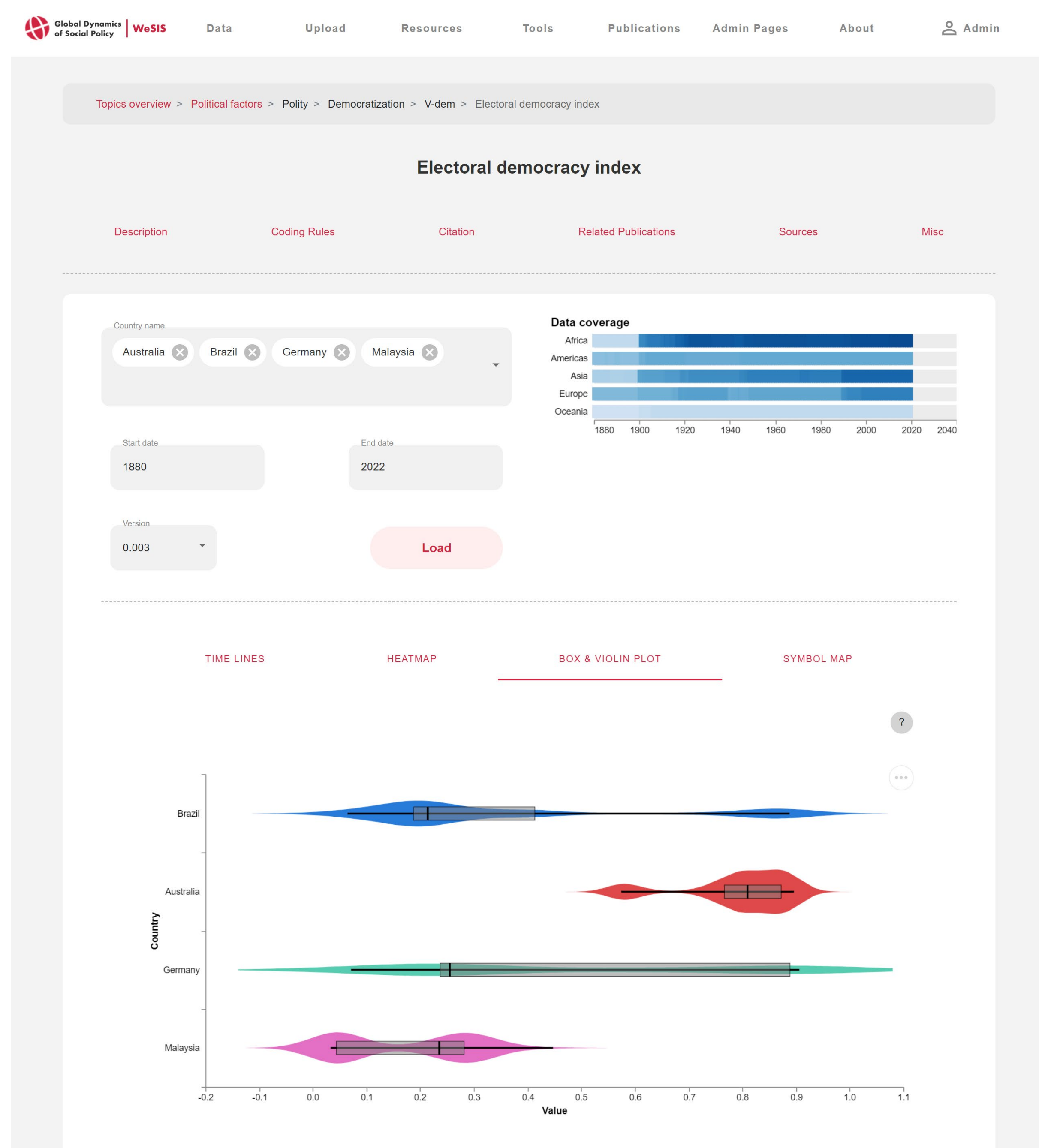


Figure 1. The Electoral democracy index indicator page (upper part of the interface).



Figure 2. Adoption in the Major Social Policy Fields for Portugal: Data visualization from the Country profile.

- For data sharing (R4), we provide data exporting options with version control, and the Community Notebooks page to share scripts.

System Evaluation

- Paper prototyping and group discussions were the most preferred activities in the co-creation workshops.
- We conducted a user study with 12 researchers to evaluate the validation and exploration features through five tasks. After each task, participants rated its difficulty. We present the difficulty answer rates in Fig. 3.

Task difficulty	Very easy	Easy	Neutral	Difficult	Very difficult
T1: Uploading a valid file	50.0%	33.3%	8.3%	8.3%	8.3%
T2: Uploading a file with errors	33.3%	41.7%	8.3%	8.3%	8.3%
T3: Browsing by indicator	41.7%	41.7%	8.3%	8.3%	8.3%
T4: Accessing a country profile	41.7%	41.7%	8.3%	8.3%	8.3%
T5: Using the Data Explorer	33.3%	33.3%	25.0%	8.3%	8.3%

Figure 3. Difficulty rank per task in the user study.

- The experts saw the benefit of a systematized workflow for future colleagues but considered that co-creating increased their workload.

Recommendations for Data Infrastructure Projects

1. Ensure a limited yet representative group of participants actively involved in the process.
2. Assess the status and amount of data available early on.
3. Connect individual and group goals, working in short iterations.
4. Define the roles and tasks of the participants early on.

Contact author:
Gabriela Molina León
molina@uni-bremen.de

