

PolPaL (Political Party Links)

Tutorial “Linking”

Paul Bederke & Kyra Meyer, October 2026

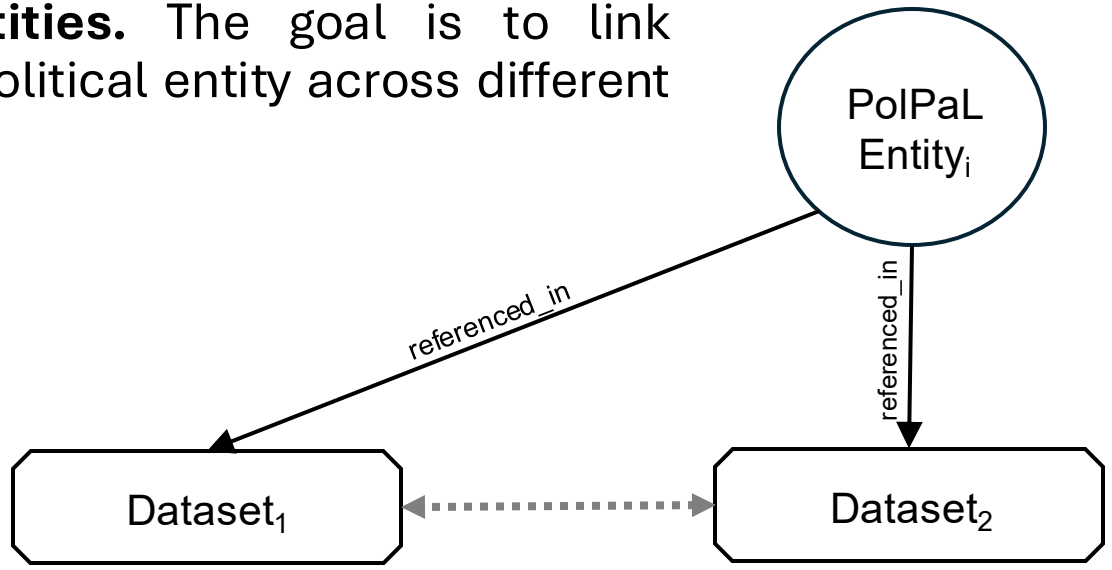
Content

- **Introduction**
- **Data Organization**
- **Linking Tool Interface**
  - Navigation
  - External Entries
  - Selecting PolPaL entities
  - Linked PolPaL entities
- **Linking External Observations**
  - Aggregated Format
  - Detailed Format
  - Observations without a Time Dimension
  - Special Cases

Introduction

PolPaL provides a linkage platform for identifying political entities, such as political parties or electoral alliances, across social science datasets. With PolPaL, researchers can easily combine information on political entities across various datasets, save time during data preparation, increase transparency and reproducibility, reduce linking errors, and simplify complex analyses.

This tutorial explains the linking tool interface and the steps for linking external dataset observations to PolPaL entities. The goal is to link observations from external datasets to the same political entity across different datasets by linking them to a common PolPal entity.



Data Organization

Data Project, Dataset, and Data Version

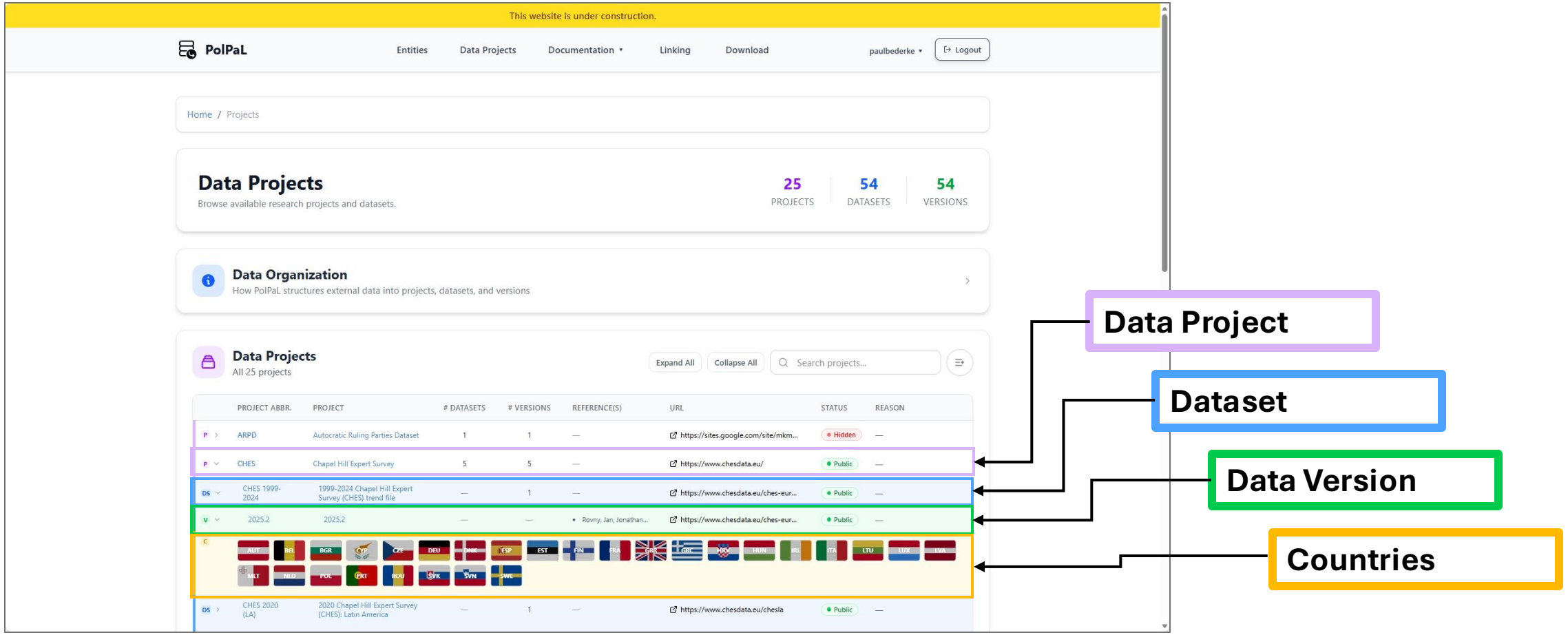
Data Organization

PolPaL links observations of political entities across datasets. Those datasets are provided by a variety of scholars, institutions, and research infrastructures — some updated regularly, others published and archived once. To accommodate these differences, PolPaL structures external data into three levels: **data projects**, **datasets**, and **data versions**.

- **Data Project:** The overarching research project, institution, or infrastructure responsible for one or more datasets. For example, the *Varieties of Democracy Institute* publishes several datasets and is represented as a single data project. When a dataset is published independently — outside the context of a broader project — the dataset itself also serves as the data project. For example, the *One-Party Membership Dataset (OPAMED)* is both the data project and the dataset.
- **Dataset:** A specific collection of data nested within a data project, typically corresponding to one data file (plus any supplementary files). For example, while the European Social Survey is the data project, *ESS Round 1* and *ESS Round 2* are separate datasets within it. If the same data is split across multiple country files, each file is treated as a separate dataset.
- **Data Version:** A release of the same dataset reflecting minor corrections or major updates. For example, *ESS11 edition 4.0* and *ESS11 edition 4.1* are two versions of the same dataset.

Within each data version, external observations are nested in **countries**.

Data Organization



Linking Tool Interface

Navigation, External Entries, Selecting PolPaL entities, and Linked PolPaL entities

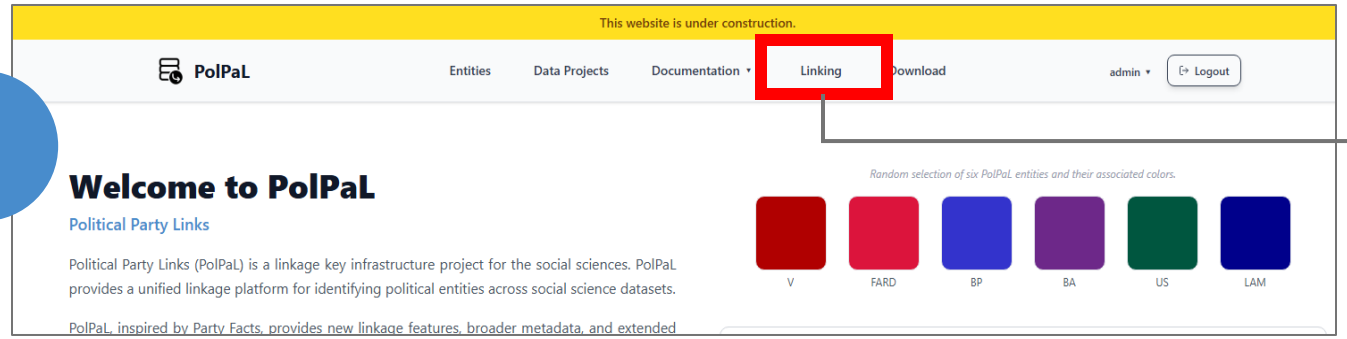
Linking Tool Interface

Linking is the process of assigning PolPaL entities to observations in external datasets. For the linking process, PolPaL provides a **linking tool interface** that displays the external dataset observations and allows users to assign PolPaL entities to them.

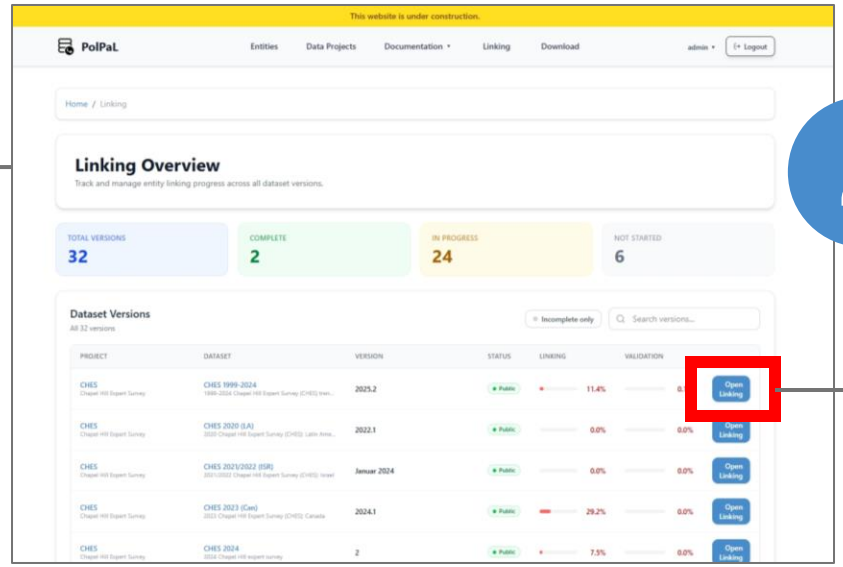
External dataset observations can be shown in either an **aggregated** or **detailed format**. In the aggregated format, external dataset observations with the same ID and ID variable are grouped together and shown only once. In the detailed format, each observation is displayed separately in a list, sometimes accompanied by additional information provided by the external dataset that may facilitate the linking process.

Linking Tool Interface: Navigation

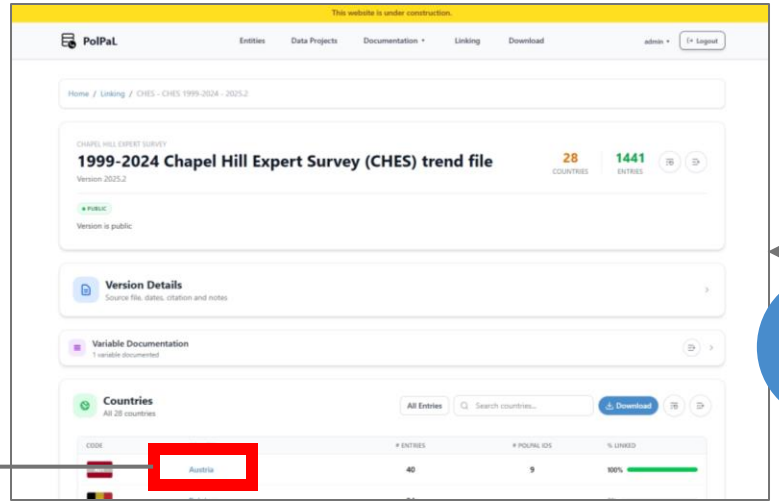
1



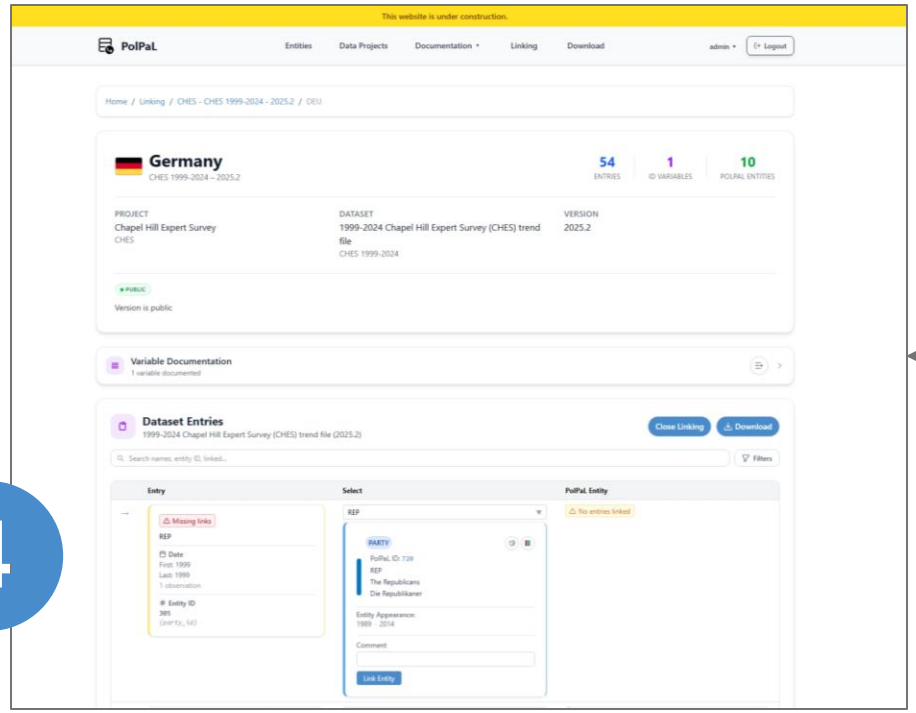
2



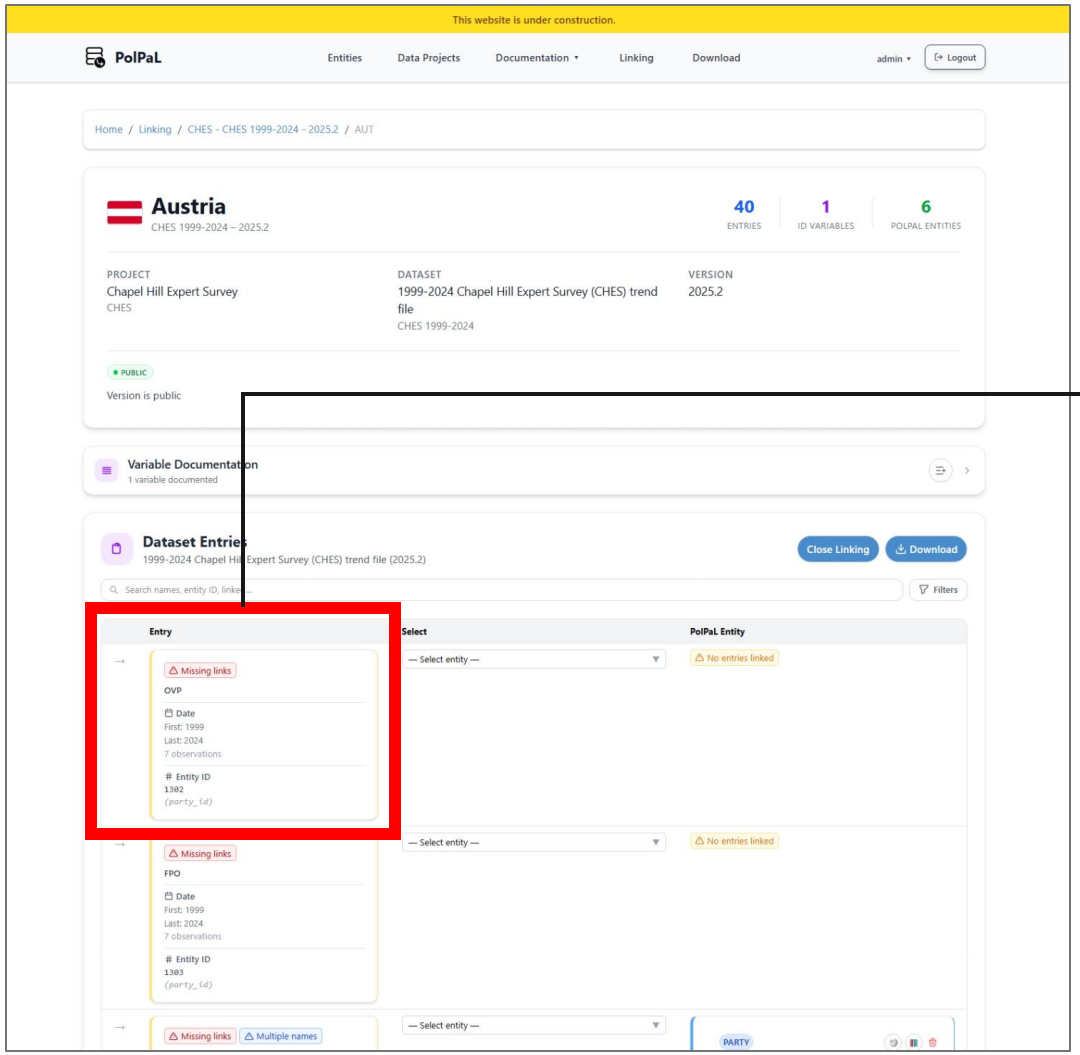
3



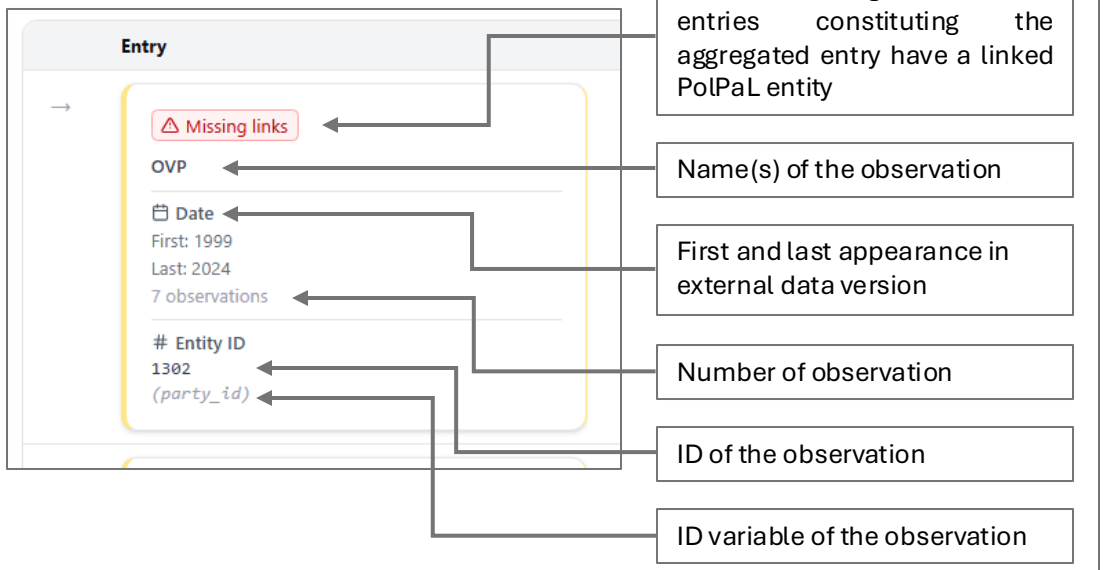
4



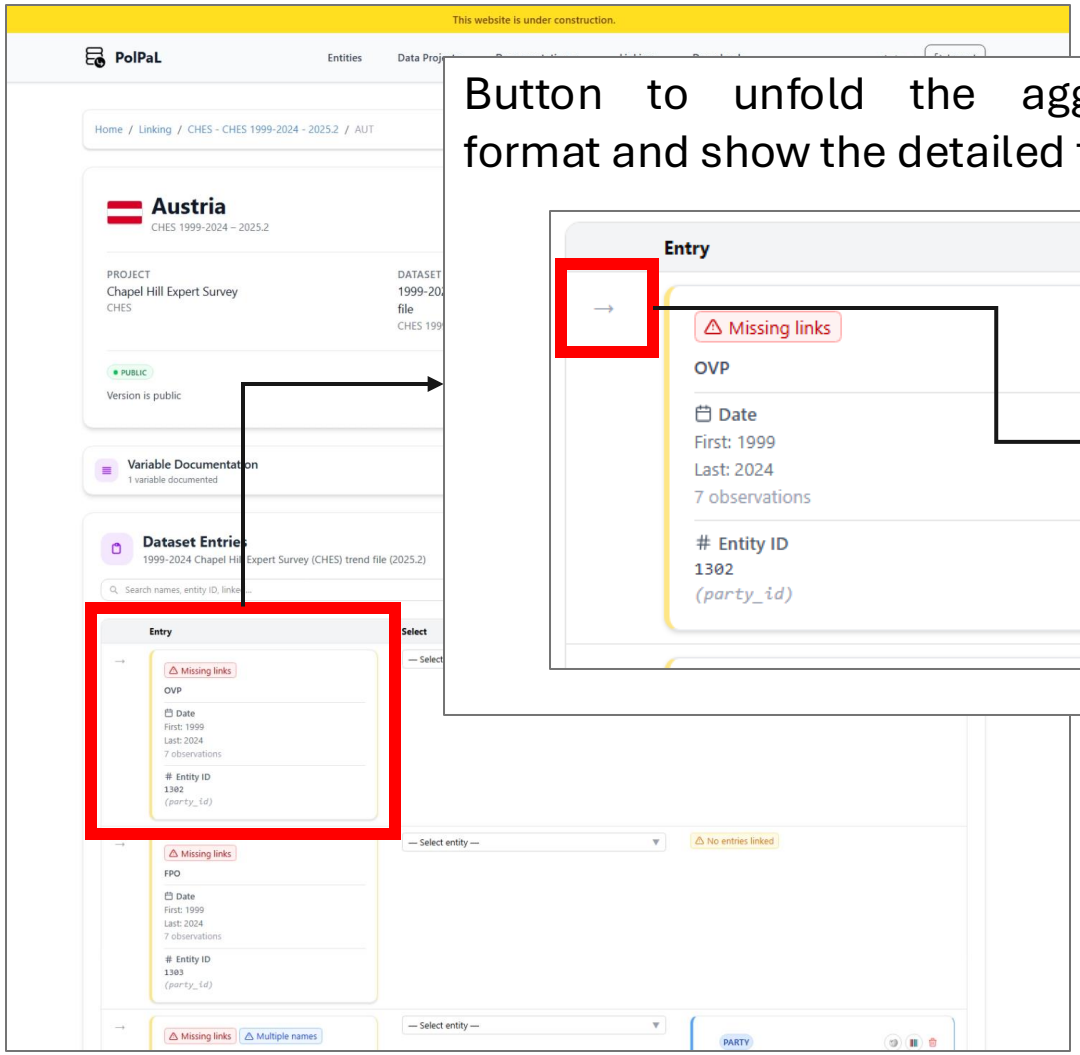
Linking Tool Interface: External Entries



Aggregated format: Entries from the external dataset are aggregated based on the ID variable and the ID of the observation.

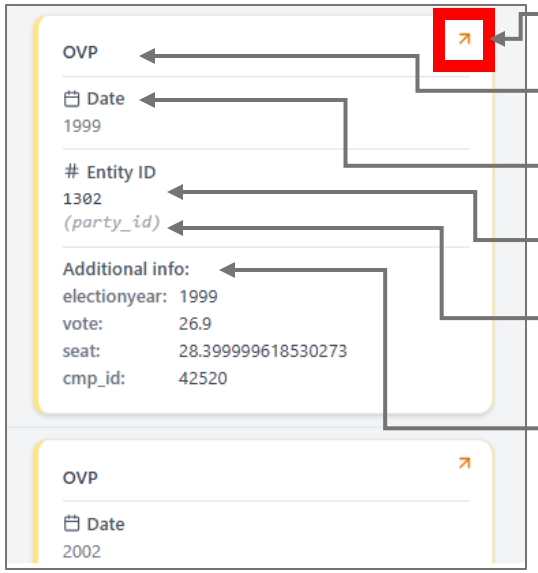


Linking Tool Interface: External Entries



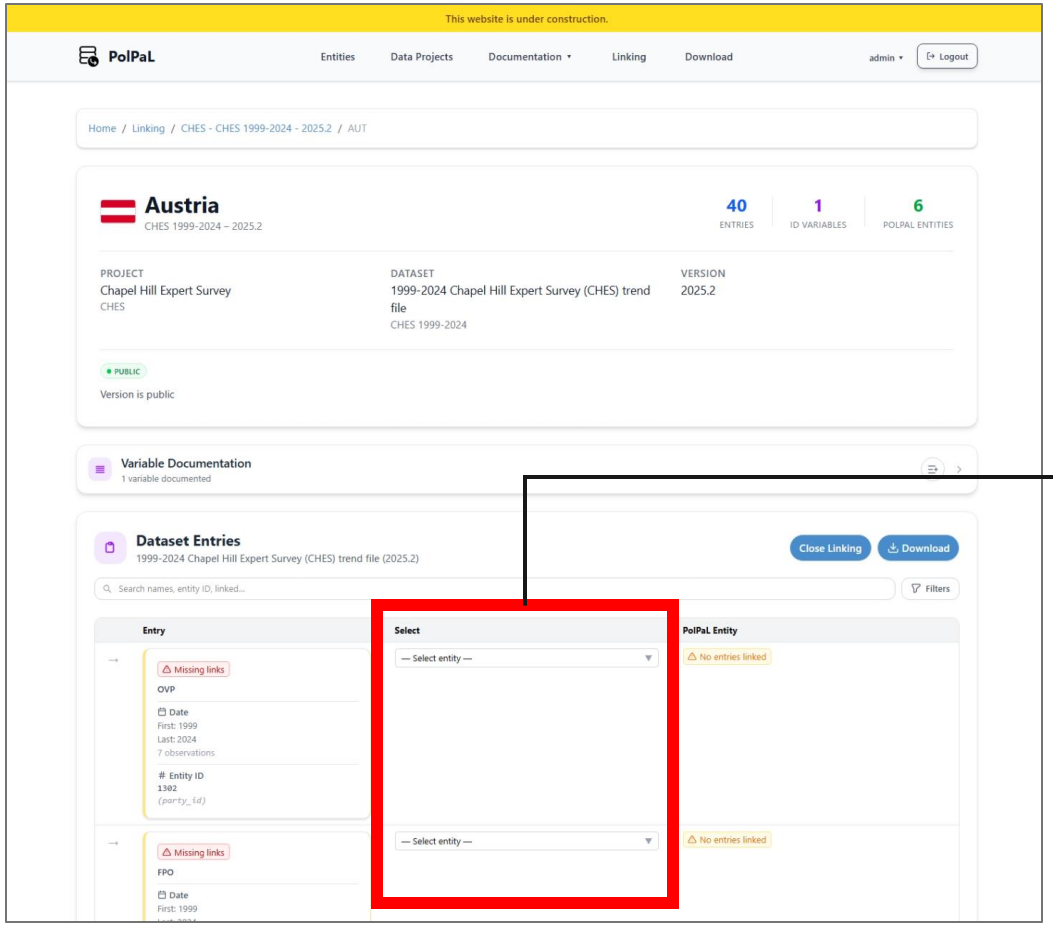
Button to unfold the aggregated format and show the detailed format.

Unfold for **detailed format**, one row per external dataset observation.

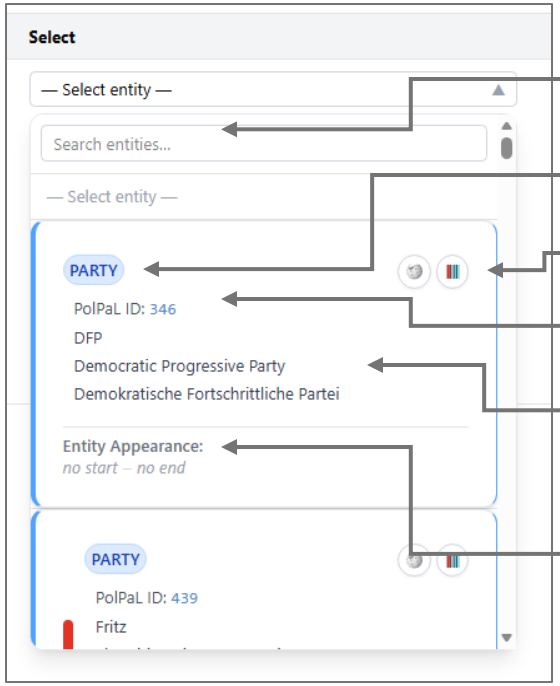


- Redirect to detailed entry page
- Name(s) of the observation
- Specific date of the observation
- ID of the observation
- ID variable of the observation
- Additional entry information extracted from the external dataset

Linking Tool Interface: Selecting PolPaL entities

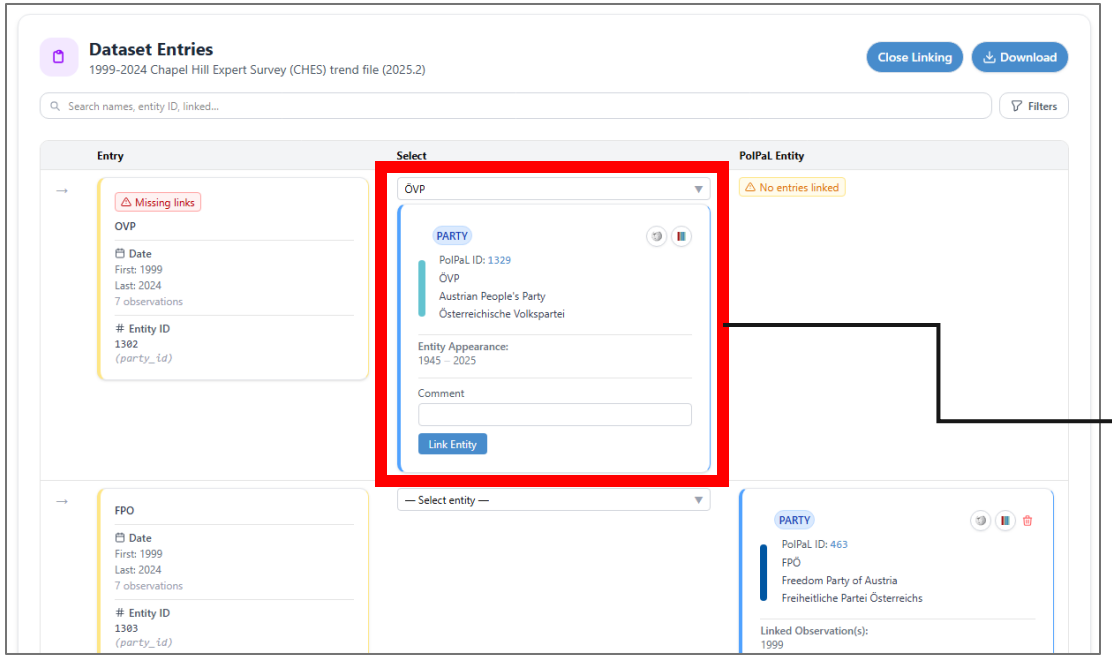


Searching and selecting a PolPaL entity

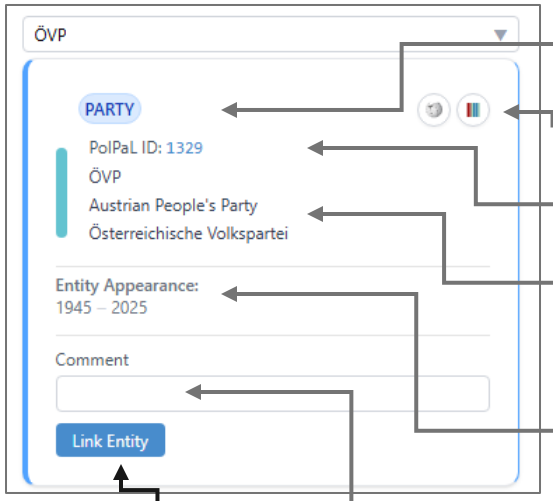


- Search field to find a PolPaL entity (from the respective country)
- Entity type
- Wikipedia and Wikidata links
- PolPaL ID
- Name(s) of PolPaL entity
- If available, first and last year for the appearance of the PolPaL entity across all external datasets (calculated automatically)

Linking Tool Interface: Selecting PolPaL entities

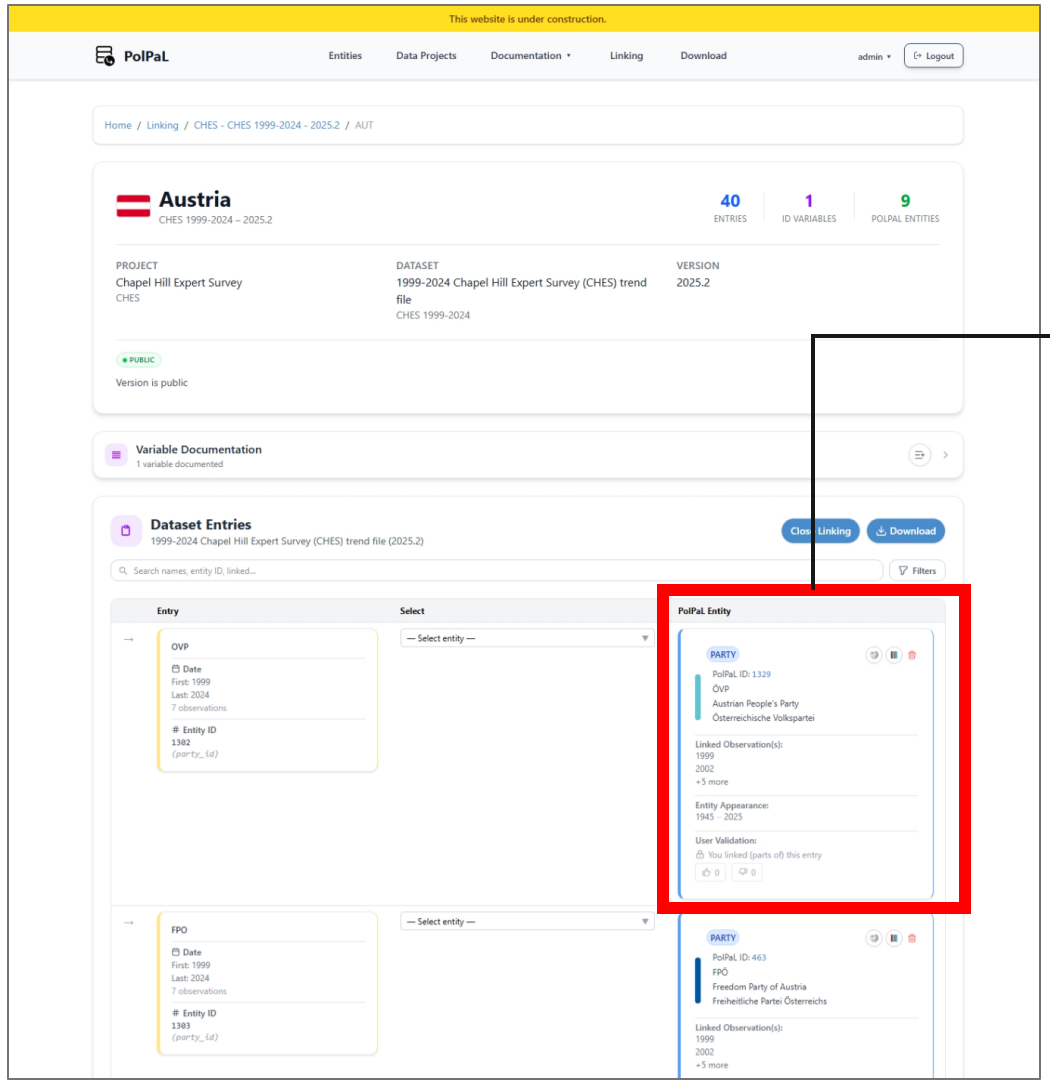


Selected PolPaL Entity

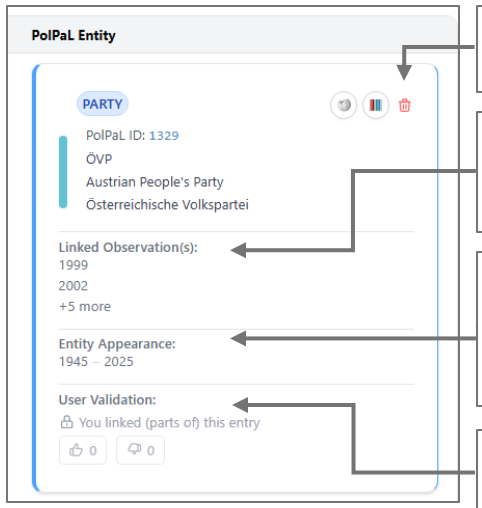


- Entity type
- Wikipedia and Wikidata links
- PolPaL ID
- Name(s) of PolPaL entity
- If available, first and last year for the appearance of the PolPaL entity across all external datasets (calculated automatically)
- Button to link the PolPaL entity and the external dataset observation
- Comment field, for any additional information that helps understand complicated cases.

Linking Tool Interface: Linked PolPaL entities



The linked PolPaL entity



Delete link of entries to PolPaL Entity

Date(s) of the external dataset observation(s) linked to the PolPaL entity

If available, start and end year for the appearance of the PolPaL entity across all external datasets

User validation scores and comments

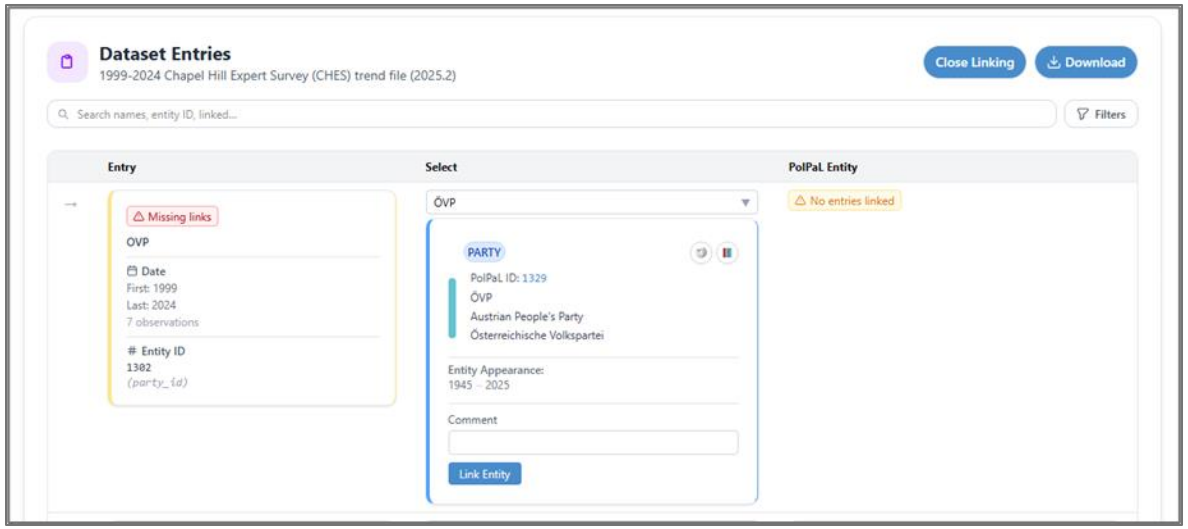
Linking External Observations

Aggregated Format, Detailed Format, and Observations without a Time Dimension

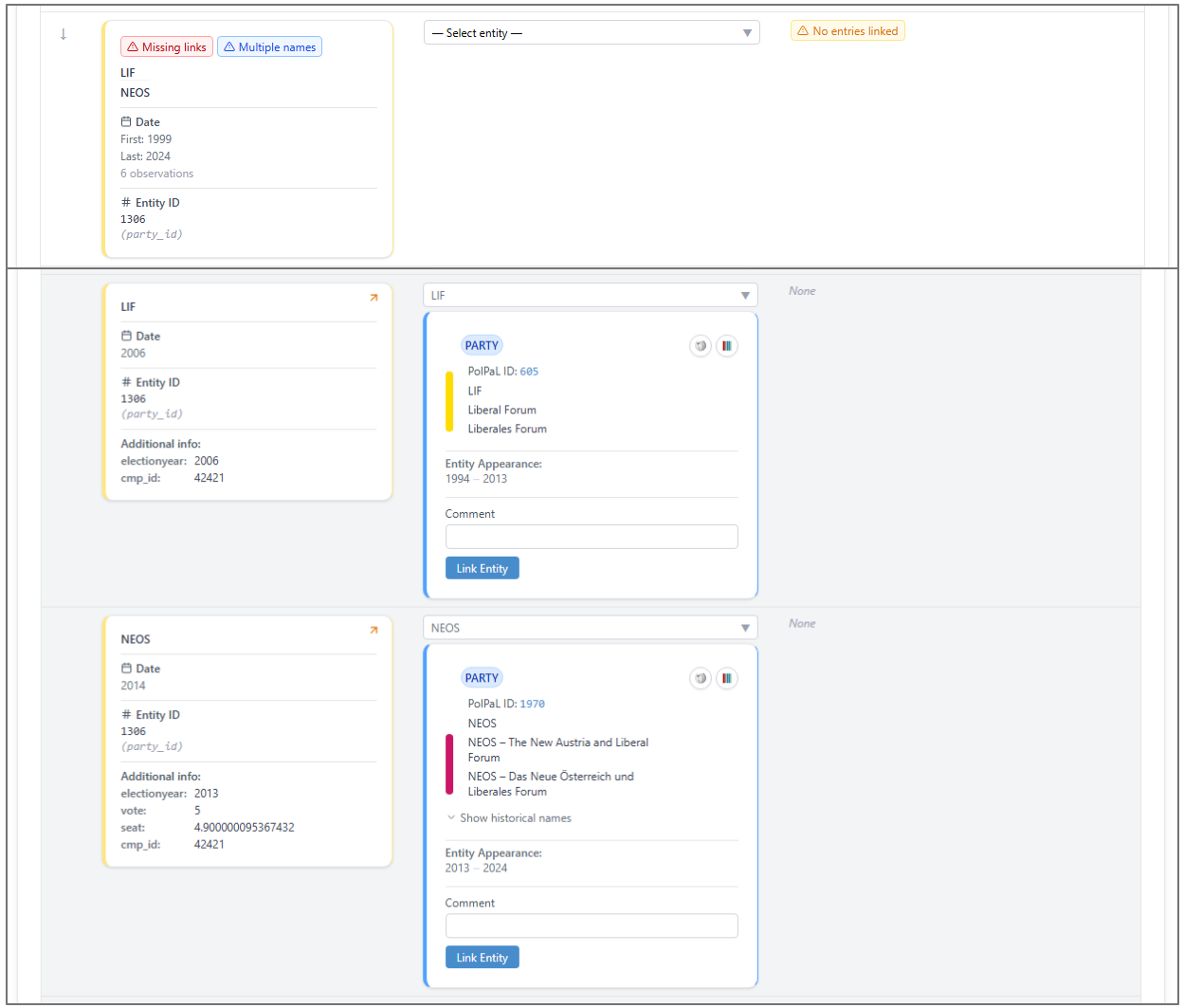
Linking External Observations: Aggregated Format

It is possible to link all external dataset observations which are aggregated into one group (by external ID and ID variable) to a PolPaL entity at once.

Be sure that all the observations that are aggregated should indeed be linked to the respective PolPaL entity.



Linking External Observations: Detailed Format



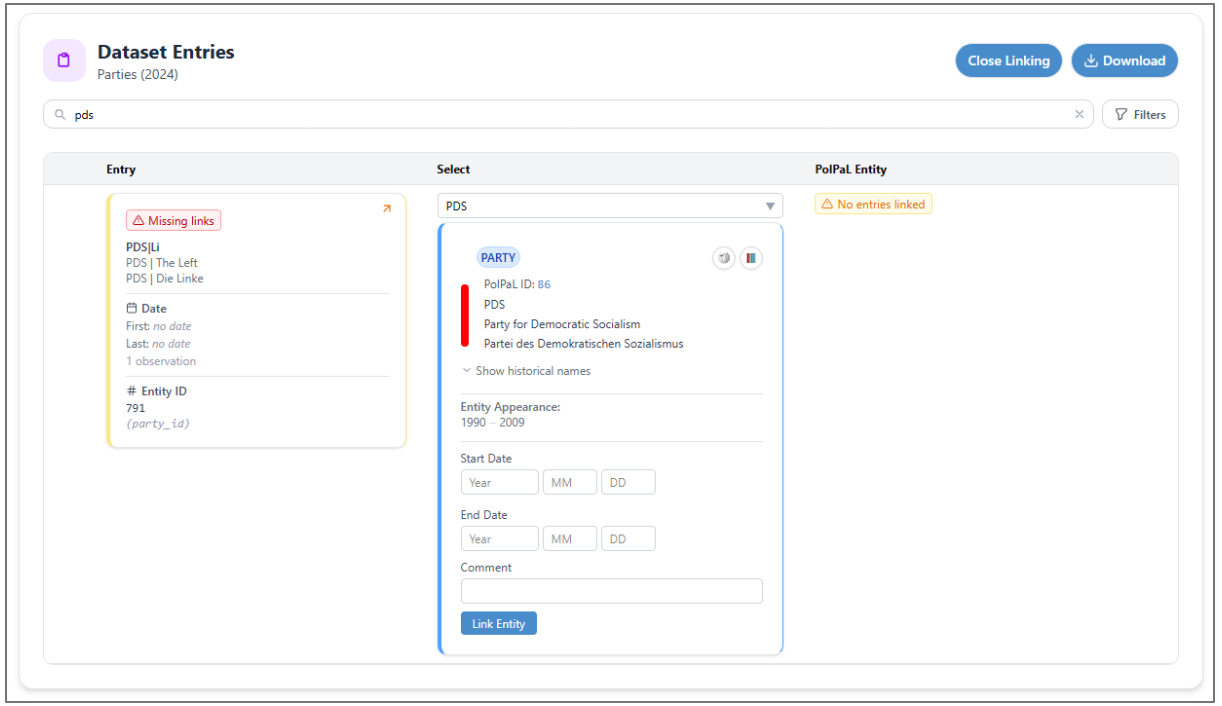
In some cases, it may be necessary to link multiple PolPaL entities to a single external dataset observation, which is shown in an aggregated format.

Unfolding allows the display of the detailed entries. You can subsequently link PolPaL entries by observation. This may be necessary, for example, when an external dataset treats two PolPal entities, such as a current party and its predecessor, as a single entity and thus assigns them the same ID.

Linking External Observations: Missing Time

In some cases, the external dataset observation does not contain a temporal dimension.

In most cases, the corresponding PolPaL entity can just be linked. However, sometimes the external dataset observation codes two parties that are related to each other as one. In those cases, it might be required to specify the temporal dimension for each PolPaL entity linked.



Linking External Observations: Special Cases

PolPal includes several special entities to accommodate external dataset observations that do not correspond to political parties or electoral alliances, or that cannot be linked to a specific political entity.

Polpal independent covers two types of observations in external datasets: observations of individual independent politicians (e.g., Olaf Scholz) and observations of aggregated groups of independents (e.g., Independents).

PolPal others covers external dataset observations that group multiple political parties into a single category without specifying the individual parties (e.g., others (etc), other right, other left).

PolPal residual covers various types of external dataset observations that cannot be linked to any other PolPal entity. These are often specific to the external dataset, such as survey responses (e.g., Abstained, Did not vote, Don't know), electoral outcomes (e.g., Invalid votes, No-seat, Appointees), and leadership affiliations (e.g., King, Queen, Military).

Importantly, external dataset observations labeled Unknown are also linked to the PolPal residual entity.

PolPal unknown covers external dataset observations whose referents cannot be identified, making it unclear which entity they refer to.

If something is unclear, please discuss with another PolPaL Team Member!

lea.kaftan@gesis.org or paulbederke@ps.au.dk