



poiesis
TRUST IN SCIENCE

D1.4: Guidance to consolidated OA data set for responsible research practices and trust in science

Author(s): *Hannah Bunt (LSE), Martin Bauer (LSE), Anne-Sophie Behm-Bahtat (WiD), Bankole Falade (LSE)*

Project title: **Probing the impact of integrity and integration on societal trust in science**

Project acronym: **POIESIS**

Grant Agreement no.: 101057253

Lead partner for this deliverable: *London School of Economics and Political Science*



This document accompanies the two Excel documents that together constitute Deliverable 1.4: a consolidated open access data set for responsible research practices and trust in science. D1.4 consists of two databases: the POIESIS National Database and the POIESIS Regional Database. The method and rationale behind the construction of all indicators in both databases are laid out in Work Package 1's previous deliverable, D1.3 (Bauer et al., 2024).

The two databases are accessible open access on the Zenodo platform under the following link: <https://doi.org/10.5281/zenodo.13373059>

The National and Regional Databases follow the same format. Both databases encompass data points available in the Eurobarometer SPSS data files (accessible via GESIS). Additionally, the National Database includes data from the Wellcome Global Monitor and the IRIS survey (Allum et al., 2023). Notably, at the country level, the data points include separate entries for East and West Germany. We ensured that the weighted results for Germany as a whole were also included in the database. On a regional level, the Eurobarometer provided data at NUTS 1, 2, or 3 levels, depending on the country. We utilised the data available in the SPSS files and retrieved corresponding socio-economic data at the same NUTS levels as each country's data. Most socio-economic data originates from Eurostat, with a handful of other indicators sourced as detailed in the indicator codebook.

Below we introduce the sheets that make up both Excel files, along with a few notes where relevant.

Cover sheet: this sheet indicates the file ID, deliverable name and authorship.

Data codebook: this sheet contains information about every single variable included in the respective datasets. Each survey question from the 'core items' related to the POIESIS model, as described in D1.2 (Bauer et al., 2023), was broken down so that every answer option became a separate variable in our datasets. These variables were given descriptive names, laid out in the column 'Variable Name'. For example, each answer option of the survey item "We have no option but to trust those who govern science and technology" in Eurobarometer 516, conducted in 2021, was given the following variable name: EB516_21_NoOption, followed by the shorthand version of the answer option, e.g., StrAgree for 'Strongly Agree'. The full variable name thus becomes EB516_21_NoOption_StrAgree. With the "No option" question having a 5-point Likert scale and 'don't know' answer option, this resulted in 6 variables in the database.

Each variable is consequently associated with the question's ID in the survey file ('Question ID'), survey name ('Survey'), year in which the survey was conducted ('Year'), phrasing of the survey question or statement ('Question/Statement'), specific answer

option that the variable refers to ('Answer Option'), as well as a binary Yes or No as to whether the datapoints were weighted for Germany, since Eurobarometer distinguishes between East and West Germany ('Germany weighted?').

Indicator codebook: this sheet contains information about all the indicators that were calculated or constructed. The columns specify the name of the indicator, the type of indicator (derived, meta, or socio-economic indicator as explained in D1.3), a description of the indicator, calculation or expression of the indicator, its source (it was either calculated by us based on the base indicators, or originates from e.g., Eurostat), and notes.

Data + indicators: this sheet contains the data points for all the variables listed in the 'Data codebook' sheet, along with the derived and meta indicators listed in the 'Indicator codebook' sheet. Any empty cells indicate missing data points for the relevant variable. The data points listed for Germany are weighted with the weights supplied in each relevant Eurobarometer data file. At the bottom of this sheet, the EU+ average (which excludes East and West Germany as separate data points), mean (which includes East and West Germany as separate data points) and standard deviation of each variable or indicator's distribution is listed.

POIESIS indicators: this sheet contains the data points for TT100 and Gdwill100, the two POIESIS indicators of trust in science of which the construction has been laid out in detail in D1.3. Alongside these scores we report average scores of Uncertainty, Ambivalence, and Polarity across each dataset. These indicators refer to, for example, levels of Uncertainty (i.e., the percentage of "don't know" responses) across all variables included in our dataset, and therefore indicate the average response pattern for 'Uncertainty' within a given geographical unit.

Socio-economic indicators: this final sheet contains the data points for the socio-economic indicators used in our analyses, as listed in D1.3. All datapoints on this sheet originated from external sources listed in the indicator codebook, such as Eurostat.

References

- Allum, N., Reid, A., Bidoglia, M., Gaskell, G., Aubert-Bonn, N., Buljan, I., Fuglsang, S., Horbach, S., Kavouras, P., Marušić, A., Mejlgaard, N., Pizzolato, D., Roje, R., Tjeldink, J., & Veltri, G. (2023). Researchers on research integrity: a survey of European and American researchers. *F1000Research*, 12, 187. <https://doi.org/10.12688/f1000research.128733.1>
- Bauer, M. W., Behm-Bahtat, A.-S. & Bunt, H. (2023). *D1.2: Dataset on core trust items, climate science, and COVID-19*. Project POIESIS, HORIZON.
- Bauer, M. W., Behm-Bahtat, A.-S., Bunt, H. & Falade, B. (2024). *D1.3: POIESIS Indicators for Trust in Science*. Project POIESIS, HORIZON.