Unvailing Open Science Impact: Findings and Insights from the PathOS Project

Pathos
Open Science Impact Pathways

PathOS Objective

Identify and quantify the **Key Impact Pathways** of Open Science across **academia, society, and the economy** to enhance understanding and drive informed policy-making.

Beyond state of the art

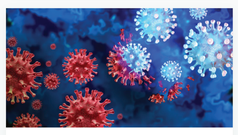
- Map the Causal Pathways for Open Science
- Design and estimate OS Impact Indicators for selected case studies
- Use data-driven, AI-assisted methodologies
- Formulate a Cost-Benefit Analysis framework for Open Science

Based on Investigative Case Studies

Accelerating collaborations within academia & industry
Portugal, Repository Network, Industry-academia

Research data and knowledge / use in non-academia
France, Societal uptake

Cross cutting effects due to Open Research data from National Repository
National Repository, SSH



COVID-19 case study: open artefacts and their role in the adoption of scientific developments
Covid-19

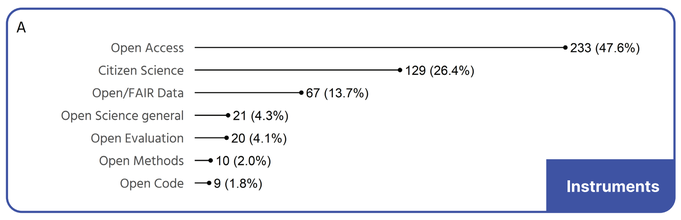
Emerging Topics:
Open Access in AI and Climate – Gender Dynamics and Industry Impact
H2020, Gender, Climate change

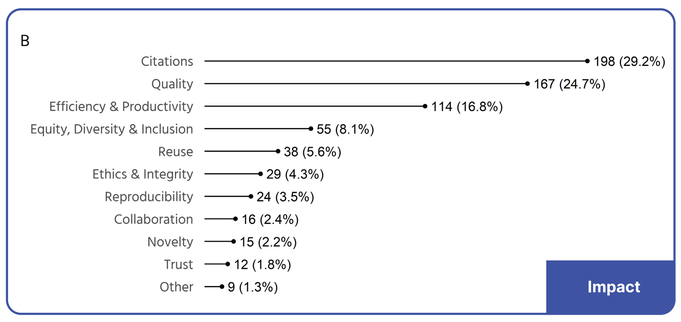
Innovation from open research resources
Life Sciences, Bioinformatics

Existing Evidence

Existing evidence for Impact
725 Scientific Papers

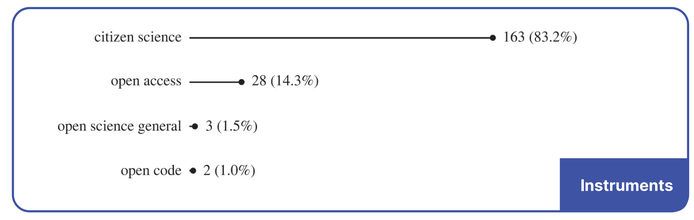
Academic Impact

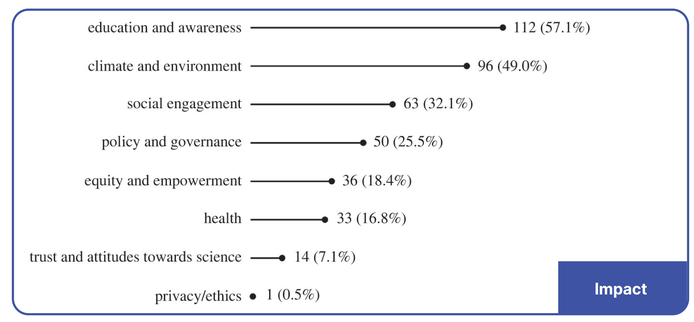




Societal Impact

Mechanisms that drive impact: public participation, collaborative creation of data, uptake of data and stakeholder engagement





Economic Impact

Scarce company data

Many theoretical papers on expected gains, but few with real evidence

Most papers on Open Science, OA and Open Data, few on Citizen Science, Open Source or Open Code

Most evidence comes from the medical and biotech sector

Challenges & Implications

Lack of Standards: No clear definitions for OS impact

Causality/Correlation: Hard to directly measure impact relationships

Knowledge Gaps: Robust evidence missing in many areas, "**streetlight effect**"

Indicators - State-of-the-Art

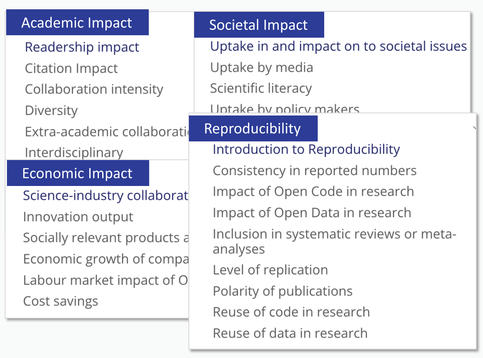
Academic: Well-developed for traditional metrics, e.g., citations

Open Science: Practices are well-covered. Training and policies need better indicators

Reproducibility: Challenging to measure. Some indicators under development - in collaboration with TIER2 project

Societal & Economic Impact: Less developed and harder to measure

Use of Proxies



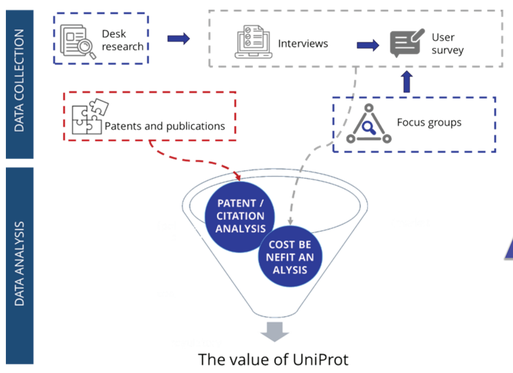
Causality - Challenges

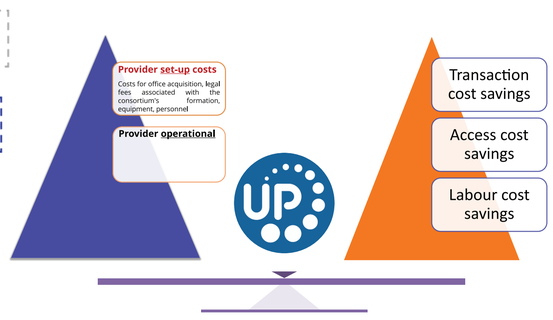
- **Complex Relationships**: Multiple factors make establishing direct causal links difficult

Example of confounding factors: Increased collaboration after Open Data policies could also be due to more funding or training, complicating attribution

- **Causal Thinking in Interpretation**: Indicators alone are insufficient—interpreting their significance requires understanding causal pathways

Cost Benefit Analysis for OS: UNIPROT





ALL USER POPULATION	LOWER BOUND	UPPER BOUND
COSTS		
Total annual average OPEX	14,664,728	
Total staff costs (EUR)	10,403,892	
Other OPEX	4,260,836	
Travel (EUR)	135,377	
Equipment (EUR)	1,455,562	
Consumables and publications (EUR)	69,465	
Overheads (EUR)	2,600,433	
Total USER costs (EUR)	7,381	
Community contribution costs	7,381	
TOTAL COSTS	14,672,109	
BENEFITS		
Transaction cost savings	1,276,382	
Access cost savings	39,870,24	
Labour cost savings	332,166,024	524,000,142
TOTAL BENEFITS	373,312,647	565,146,765
TOTAL NET BENEFIT PER USER	3,567	5,475

Note: The table reports the average annual values calculated over the period 2023-2017. All values are expressed in 2024 EUR. For total user costs, only the value of community contribution costs was included in the analysis. As explained in Section 3.2.2, training costs were considered negligible compared to alternative scenarios, while access costs were factored into the calculation of access cost savings on the benefit side to avoid duplication and thus were excluded from the total user costs.

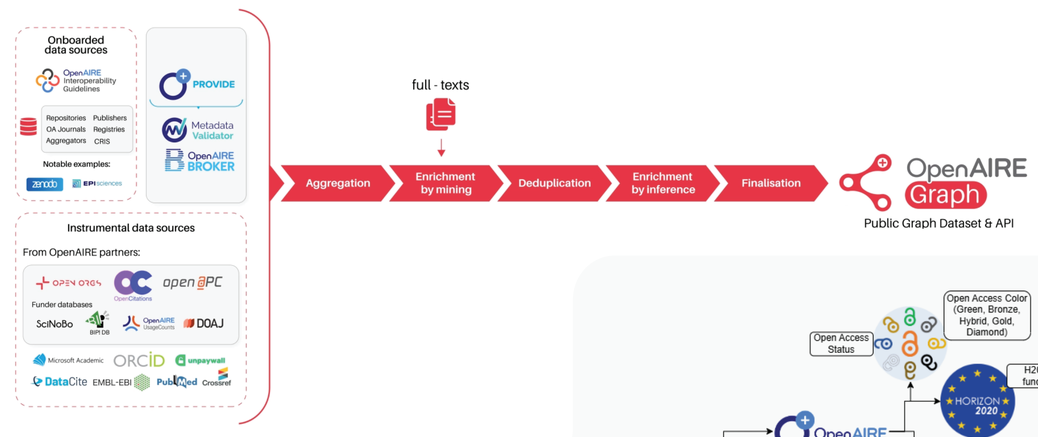
Case Study –Emerging Topics: Open Access in AI and Climate Gender Dynamics and Industry Impact

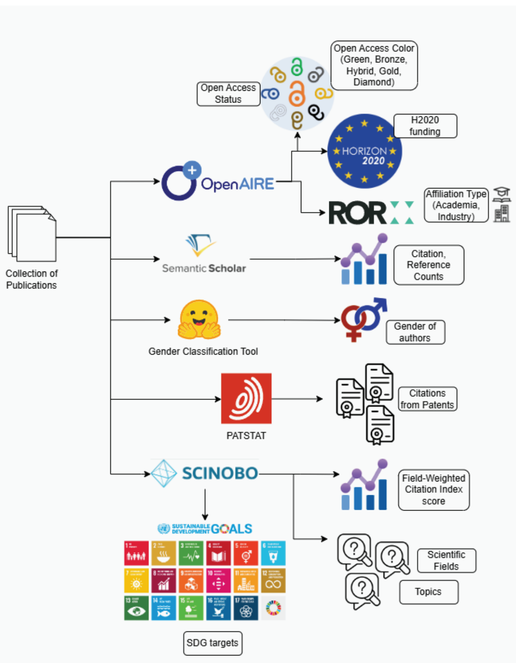
What We're Investigating

- **Innovation Impact:** Effect of Open Access (OA) on AI in climate innovations.
- **Gender Analysis:** Impact of OA on women's representation in authorship.

Methodology

- **MAIN Data Sources**
  - **OpenAIRE Graph:** 195M publications, 4M projects.
  - **PATSTAT:** 200M patent records.
- **AI-Driven Analysis:** Leveraging AI and expert guidance to link information from large datasets.t

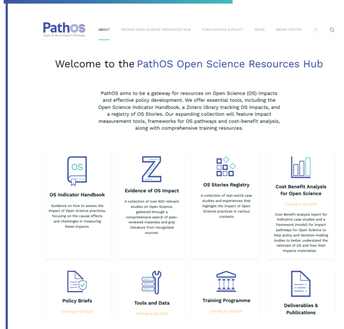




Methodological Insights

- **Causality**: Establish causal links through control groups, using big data and expert guidance.
- **Policy Analysis:** AI-driven indicators provide evidence-based, granular, historical, and predictive insights for decision-making.
- **Tracking Innovation:** Identify emerging topics in climate research, capturing new areas of innovation.
- **Transparency:** Utilize open data and transparent practices for data validation, reusability, and clear research processes.

Key Outputs



1 Frameworks
- OS Impact Pathways
- Cost-Benefit Analysis for OS

2 Handbook of OS Indicators
- Indicator "Recipes"
- Tools and Datasets

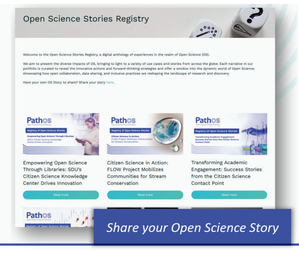
3 Literature Insights & Registry
- Lit Review on OS impacts
- Online registry of OS stories

4 Case Study Deep Dives
- OS impact assessments, Causality focus
- Cost-Benefit evaluations (Elixir UniProt & RCAAP case studies)

5 Training & Engagement
- Training for policy-makers & research administrators

6 Recommendations
- Synthesis
- Policy Briefs

Feedback & Dialogue



This project has received funding from the European Union's Horizon Europe framework programme under grant agreement No. 101058728. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Executive Agency. Neither the European Union nor the European Research Executive Agency can be held responsible for them.

Join Us

pathos-project.eu/ @PathOS_EU pathos-project