Assessing the Impact of Open Research Data Sharing – Contributions and Tools from the PathOS Project

Authors (Organisation):

Lennart Stoy, Technopolis Group Belgium https://orcid.org/0000-0002-3827-9815
Izabella Martins Grapengiesser, Technopolis Group Belgium https://orcid.org/0009-0006-9821-8236
Elisa Seminaroti, Technopolis Group Belgium https://orcid.org/0009-0003-1476-4052
Ioanna Grypari, Athena Research Center and OpenAIRE https://orcid.org/0000-0002-7485-1591

Description

FAIR and Open research data sharing is increasingly seen as a cornerstone of modern research practices. The rationale for this shift includes improving efficiency and reproducibility, advancing interdisciplinary research, and fostering innovation. However, the impacts of Open Research Data, particularly in academic and economic contexts, are less well understood than those of open government data. The PathOS project, a 3-year Horizon Europe initiative, addresses this gap by expanding knowledge on the impacts of Open Science, with a focus on Open Research Data. It does so through three main pillars:

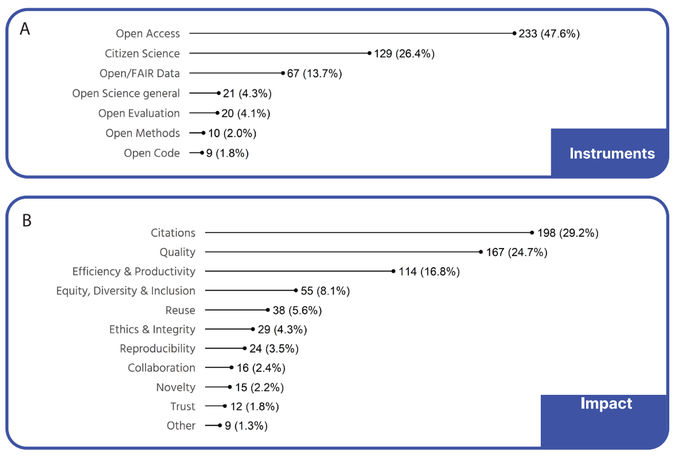
1. **Modelling Impact Pathways:** Mapping the key outputs, outcomes, and impacts of Open Science, including Open Research Data, using evidence-based frameworks.
2. **Operationalising Indicators:** Developing and testing tools, data flows, and indicators to measure Open Science impacts, validated through six diverse case studies.
3. **Cost-Benefit Analysis Framework:** Evaluating the net economic value of Open Science practices, offering insights into the efficiency, innovation, and economic growth facilitated by Open Data.

UNDERSTANDING THE IMPACT OF OPEN DATA

Systematic screening of over 30,000 records identified 479 relevant studies

The PathOS project conducted three scoping reviews to explore the broader impacts of Open Science, focusing on academic, societal, and economic dimensions. While these reviews examined Open Science as a whole, the evidence specifically indicating the impact of Open Data was primarily found in two areas:

- **Academic Impacts**: Evidence suggests that Open Data improves research transparency, reproducibility, and collaboration. By making high-quality datasets more accessible, it facilitates faster scientific discovery and provides new opportunities for research, supporting productivity and the uptake of results in academia and industry.

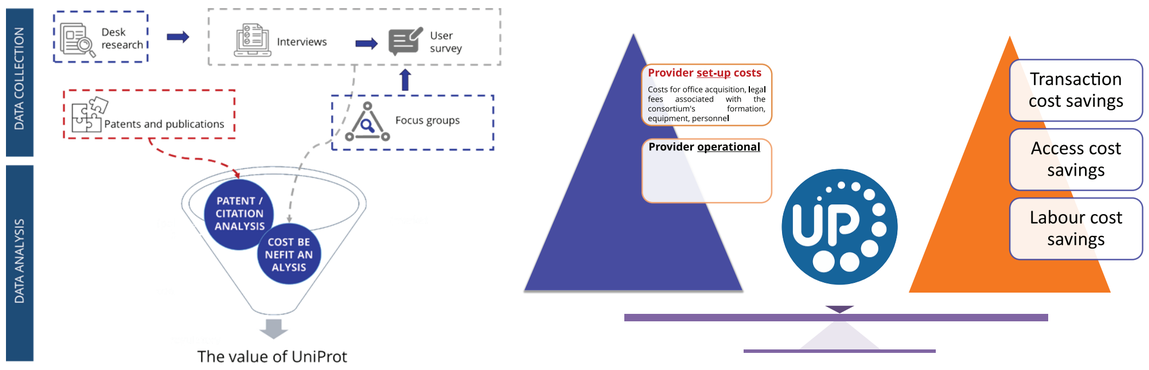


Academic Impact

the economic paper is not published yet

- **Economic Impacts:**
  - **Efficiency Gains:** Open Data has been linked to cost reductions in reproducing, collecting, and processing data, enabling researchers to save time and resources. These efficiencies may support faster research activities and uptake of results by industry.
  - **Enhancing Innovation:** Evidence indicates that Open Data lowers barriers to accessing critical information, promoting collaboration and discovery. This may enable diverse applications and commercialisation across sectors, fostering innovation.
  - **Economic Growth:** Open Data appears to contribute to broader economic impacts through improved R&D efficiency and innovation-driven growth. Preliminary findings suggest that the returns on investment in Open Data often outweigh the associated costs of data creation and curation.

EVALUATING THE COSTS AND BENEFITS OF OPEN DATA



Cost-Benefit Analysis (CBA) Framework

- Provides a structured approach to quantify the direct costs and benefits of Open Data.
- Focuses on measurable benefits such as time savings, transaction cost reductions, and enhanced research productivity.
- Essential for understanding the net impact of Open Data by comparing its outcomes to a counterfactual scenario where the data is not openly available.

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.

UniProt, a freely accessible protein knowledge base, was assessed using the CBA framework. Key benefits identified include:

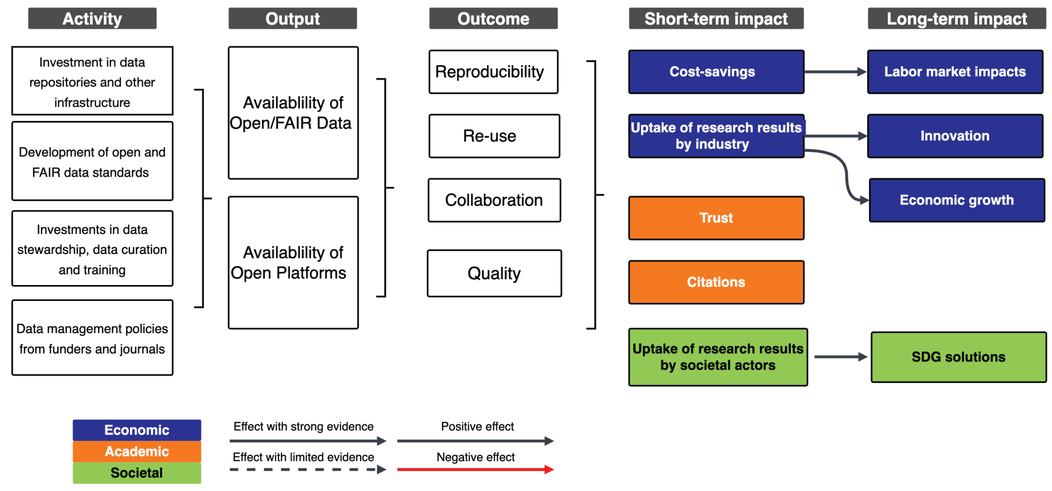
- **Efficiency Gains:** Significant time savings for users in accessing, integrating, and analyzing protein data.
- **Cost Savings:** Avoided expenses related to alternative data access methods and enhanced labor productivity.

The study highlighted the broader impacts of UniProt, such as facilitating innovation in biotech and pharmaceutical industries and enabling downstream research applications.

CONCEPTUALISING PATHWAYS FOR OPEN DATA IMPACT

Key Impact Pathways: Methodological Approach

- **Evidence-Based Framework:** Built using findings from a scoping review of articlesexploring the impacts of Open Science.
- **Stepwise Approach:**
  - Identify pathways by analysing activity-level evidence.
  - Aggregate pathways to highlight broader academic, societal, and economic trends.
- **Systematic Coding:** Ensures consistent mapping of relationships between Open Science activities, outputs, outcomes, and impacts.
- **Validation Through Stakeholder Input:** Pathways are iteratively refined to align with practical insights and real-world contexts.
- **Clear Connections to Impact:** Visualised pathways demonstrate links from activities, like Open Data sharing, to short-term benefits such as cost savings and long-term impacts like innovation and economic growth.

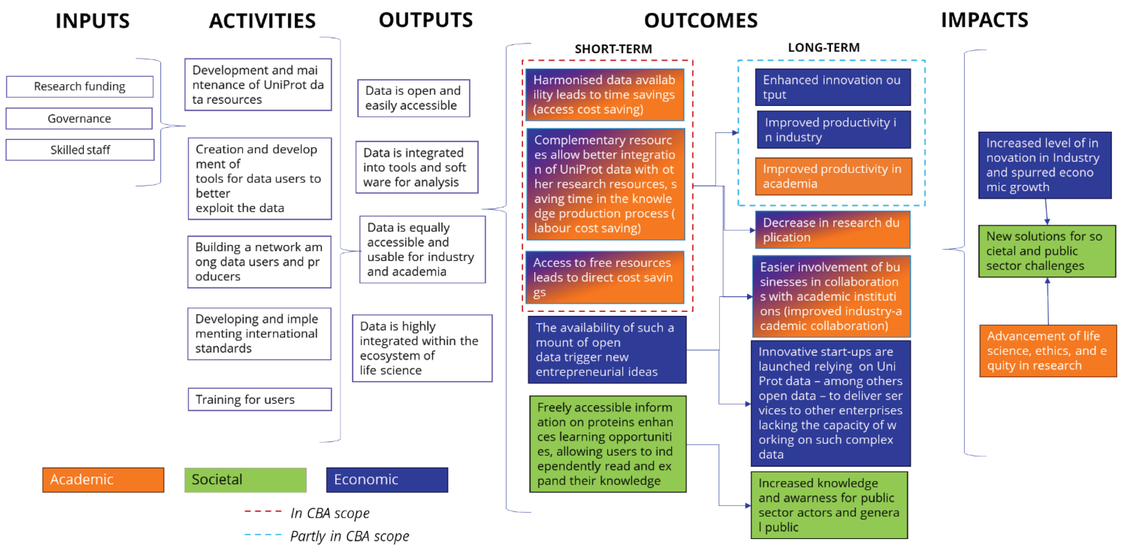


MEASURING OPEN DATA IMPACT THROUGH INDICATORS AND CASE STUDIES

The PathOS project has developed a comprehensive handbook of Open Science indicators aimed at understanding and measuring the impacts of Open Science. This handbook outlines the methodologies, metrics, and frameworks used to systematically evaluate Open Science academic, societal, and economic impacts. It serves as a tool for researchers and policymakers to assess and promote the benefits of Open Science practices.

The project focuses on three key case studies related to Open Data:

- **RCAAP (Portugal):** Examines how Open Data facilitates academic-industry collaborations, leveraging national open science platforms to drive knowledge exchange and innovation within and beyond academia.
- **EASY (The Netherlands):** Investigates the cross-cutting effects of Open Data shared through a national repository, highlighting its influence on research accessibility and interdisciplinary collaboration.
- **UniProt (ELIXIR):** Explores the economic and societal impacts of Open Data use in the biomedical sector, particularly through the ELIXIR Open Science platform, focusing on innovation and data reuse by industry stakeholders



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.