Open Science Impact Pathways

Deliverable 1.1

Open Science Intervention Logic

Deliverable Number and Name	D1.1 Open Science Intervention Logic
Due Date	28/02/23
Delivery Date	27/03/23 (Updated on 20/10/23)
Work Package	WP1
Type	R — Document, report
Author	Ron Dekker (TGB), Istvan Karasz (TGB), Lennart Stoy (TGB),
Reviewers	Tony Ross-Hellauer (KNOW), Lena Tsipouri (Opix), Nicki Lisa Cole (KNOW)
Approved by	Ioanna Grypari
Dissemination Level	PU
Version	2.0
Number of Pages	53

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.

Revision History

VERSION	DATE	REASON	REVISED BY
0.1	26/01/23	Agreement on structure & References	Ron Dekker
0.2	10/20/23	Intermediate version	Lennart Stoy, Istvan Karasz
0.3	16/03/23	Peer review	Tony Ross - Hellauer Lena Tsipouri
0.4	23/03/23	Peer review comments addressed	Lennart Stoy, Istvan Karasz, Ron Dekker
0.9	26/03/23	Final editing	Ron Dekker
1.0	27/03/23	Final Version after proofreading	Ioanna Grypari
1.1	07/08/23	Version 1.1 with updates throughout the text, new section on intervention pathways in part 3.2.1, additional information on case studies, consideration for results of the Scoping review in D1.2, updated figures 8 and 9	Lennart Stoy, Istvan Karasz,
1.2	14/08/23	Review	Nicki Lisa Cole
1.3	21/08/23	Peer review comments addressed	Istvan Karasz, Lennart Stoy
1.4	29/08/23	Final Review	Ioanna Grypari
1.5	14/10/23	Final review comments addressed	Istvan Karasz Lennart Stoy
1.6	16/10/23	Final editing	Istvan Karasz
2.0	25/10/23	Final Version	Istvan Karasz Ron Dekker

Table of contents

List of Tables

List of Figures

Disclaimer

This document contains description of the PathOS project findings, work and products. Certain parts of it might be under partner Intellectual Property Right (IPR) rules so, prior to using its content, please contact the consortium head for approval.

In case you believe that this document harms in any way IPR held by you as a person or as a representative of an entity, please do notify us immediately.

The authors of this document have taken any available measure to ensure that its content is accurate, consistent and lawful. However, neither the project consortium as a whole nor the individual partners that implicitly or explicitly participated in the creation and publication of this document hold any sort of responsibility that might occur as a result of using its content.

This publication has been produced with the assistance of the European Union. The content of this publication is the sole responsibility of the PathOS consortium and can in no way be taken as a reflection of the views of the European Union.

PathOS is a project funded by the European Union (Grant Agreement No 101058728).

Abbreviations

APC	**Article Processing Charge**
EC	European Commission
EOSC	European Open Science Cloud
EU	European Union
FAIR	Findable, Accessible, Interoperable, Re-usable
KIP	Key Impact Pathway
MS	Member State (of the EU)
AC	Associated Country (to the EU)
OA	Open Access
OECD/DAC	Organisation for Economic Cooperation and Development – Development Assistance Committee
OS	Open Science
PathOS	Open Science Impact Pathways (Horizon Europe project)
RFO	Research Funding Organisation
RI	Research Infrastructure
RPO	Research Performing Organisation
RI-PATHS	Charting Impact Pathways of Investment in Research Infrastructure (Horizon 2020 project)
ToC	Theory of Change
WP	Work Package

Executive Summary

This deliverable aims to provide the methodological framework for the impact pathways developed by the PathOS project. In this paper we explain the overarching theoretical framework that provides the basis on how to develop key impact pathways.

To be useful, the pathways need to have sufficient detail and include context, expected causal relations and provide empirical evidence. However, there is a big discrepancy between theory and setting up the pathways that are useful for policy evaluations.

Our approach is systematic in the sense that we take the EC Open Science policy actions as a starting point and that we distinguish between academic, societal and economic impacts. We start from the Theory of Change and Impact Analysis and take the RI-PATHS approach (Griniece et al., 2020) as a baseline model. To address criticisms on the impact analyses, e.g., oversimplification of relations and causalities, lack of feedback-loops due to interactions and not including responses by (other) stakeholders, we use the methods that were applied in the environmental impact of research and innovation (European Commission, 2014) and we introduce a stepwise approach:

1. As a starting point, we identify individual intervention pathways.
   These are based on the EC Open Science strategy that distinguishes eight Open Science priorities.
2. Next, we will aggregate and zoom out to a higher level
   This is done by clustering the outcomes and impacts from step 1.
   In this step the focus is on the impacts – in our case covering academic, economic, and societal impact.
3. Then, we zoom in for narrative overviews and causal relations.
   This will give room for providing details on how the interventions have worked (or can work), zooming in on causal relations, engagement by multiple stakeholders, etc.

This approach allows maximum use of multiple data collection methods: literature scoping, implementing indicators, but also case studies (both thematic and national) and cost-benefit analysis. It allows for a systematic approach for the uptake of empirical information.

To collect and absorb these different information types, we develop ontologies for the different stages, e.g., standardize on types of inputs, outputs, categorisation of stakeholders, defining the different impact categories.

This deliverable will be dynamic as it will be updated by upcoming tasks on indicators, case studies and cost benefit analysis. Hence, there will be feedback loops from these activities to have the final Key Impact Pathways.

1. Introduction

1.1. Scope of the project

While we are beginning to understand some of the **dynamics of OS** in the research system, evidence about how this may affect our economies and societies is limited.

PathOS wants to contribute to a better understanding and measurement of Open Science impacts and their causal mechanisms. T shape effective Open Science (hereafter: OS) policies, we need an in-depth **understanding of the mechanisms** and underpinnings of OS practices, as well as their positive and negative effects in terms of outcomes.

This deliverable aims to provide a common **theoretical framework** on how to help policy and decision-making bodies to better understand the rationale of OS, the effects of interventions, and how its impacts materialise.

Our target audiences are policy advisors and decision-makers and to contribute to better understand the rationale of OS and how its impacts materialise.

1.2. Structure of this deliverable

In Chapter 1 we will set the scene, including extensive examination regarding the European and worldwide policy categorization of Open Science (OS), and the delineation of plausible primary stakeholders and stakeholder clusters.

In chapter 2 we will explain the Theory of Change and elaborate our approach for doing Impact Analysis, followed by an outline of the diverse contributions of various work packages towards this framework.

In chapter 3, we will showcase how to operationalise impact pathways, taking the approach of a previous project called RI-PATHS as a baseline, implementing pathways. Again, we will be connecting to the other WPs by having feedback loops between this framework and the other actions.

Chapter 4 the paper provides a summary. And in the annexes are templates on how to systematically collect information from literature review and case studies.

1.3. Setting the scene

For any impact analysis, we need to start from the needs and problems that lead to policy objectives. Considering the European context of the project, as well as the potentially guiding effect of EU policy on national or institutional policies, we will use the policy categories that are being used by the EC and the EOSC Steering Board. This policy framework follows a systematic approach by taking the policy and research cycles as starting points and distinguishes eight strategic pillars[1] - see first subsection.

Our framework must allow for multiple inputs – literature, case studies, cost benefit analysis and of course connect with indicators. For this we will work together with the other WPs. Indicators will be connected to other outputs like the *Handbook of Open Science Impact Indicators* (WP2).

In order to scale up on literature review and case studies, it is important to standardise the items that are being described. This goes especially for describing the stakeholders in a consistent way – see second subsection. It is also necessary to have consistent categorisation when we want to aggregate results.

For impact, we will distinguish three types of impact: academic, economic, and societal. This is an ideal-type taxonomy, knowing that these three types of impact are often interrelated. For example, patents can be both academic and economic solutions to societal problems. Hence, it is worth mentioning this taxonomy as "ideal but interconnected".[2]

1.3.1. EC Open Science policy

In its Open Science strategy, the Commission distinguishes eight strategic 'pillars' that concern all aspects of the research cycle. Recently, new pillars have been added, like 'open software code'. Table 1 below shows the original and updated Open Science pillars, the latter being used in the EOSC Observatory.[3]

[1] https://research-and-innovation.ec.europa.eu/strategy/strategy-2020-2024/our-digital-future/open-science_en. Note: to align with EC development of these pillars there has been an update on these categories that are also used in the EOSC Observatory – see Annex 1.
[2] UNESCO (2020). *Open Science for the 21st Century:* https://en.unesco.org/sites/default/files/isc_paper_for_unesco_open_science_consultation_2020.pdf
[3] https://eoscobservatory.eosc-portal.eu/home

Table 1 Original and updated EC Open Science pillars

Original (2016) OS pillars[4]	Updated OS pillars[5]	Explanation, examples
Future of Scholarly Communication	Publications	Research publications that are available in Open Access
European Open Science Cloud & FAIR Data	Data	Research data management and research data that is FAIR/open
	Services	Services that enable research data discovery and exploitation
	Infrastructure	Data stewardship, data repositories, and data preservation
	Software	Software that enables research and is available in open source
Rewards and Incentives	Assessment	Incentives and rewards for researchers to practise Open Science
Skills and Education	Skills/training	Skills and training for researchers to practise Open Science
Citizen Science	Engagement	Research that engages and involves citizens via citizen science
Research Integrity	RRI	Refers to all pillars ('horizontal'); focus will be on reproducibility of research
Research Indicators & Next-Generation Metrics		Applies to all pillars

Source: EC 2016 and EOSC Steering Board 2022

We will take these updated OS pillars as the objectives of the EC's Open Science strategy. It is important to have these objectives because objectives are at the start of the intervention logic and act as a reference for the impact analysis. Moreover, the EOSC Steering Board annually collects data from the Member States on Open Science policies via the EOSC Observatory. This information can be used in our impact analysis and results from our analysis can be used by the EOSC Steering Board as we have aligned with their objectives.

1.3.2. Stakeholders

In general, but especially in Open Science, there will be multiple stakeholders involved. Moreover, stakeholders will respond to interventions made by the government or other stakeholders. Hence, the inclusion of key stakeholders is crucial.

We define key stakeholders as those entities that have the mandate or power to effectively change or formulate a potential outcome or to block other initiators from doing so.

Stakeholders can be involved in activities during the whole process: from initial needs to inputs, results and impacts. For example, researchers can be both important producers and users of

[4] https://research-and-innovation.ec.europa.eu/strategy/strategy-2020-2024/our-digital-future/open-science_en#ref-8-ambitions-of-the-eus-open-science-policy

[5] https://zenodo.org/record/7574165

scientific output. Key is that they play an active role in one of the stages of the impact analysis. Contributions and actions can be both positive and obstructive as some stakeholders may oppose the policy interventions.

Our initial list of stakeholders is presented in Table 2 below.

Table 2 Stakeholder categories

Main Category	Subcategory
Governments	International policy making bodies, European Commission
	National Governments
	Ministries[6]
	Regional (within country) Governments, Local governments
Funders (RFOs)	Public RFOs – for both basic and applied research
	Private RFOs
Industry	Manufactories & Service Providers
	Experts (incl. researchers) at companies
	Think Tanks
Publishers	Lobby Organisations
	Large Commercial Publishers
	Small Commercial Publishers
	Not-for-profit Publishers
Universities & Research Institutes (RPOs)	Basic Research Organisations
	Applied Research Organisations
	University Libraries
	Individual Researchers (at RPOs)
Research Communities	Formal associations
	Learned Societies
	Informal Communities
Civil Society	Community-Based Organisations
	(International) Non-Governmental Organisations
	International Groups
General Public	Citizens
	Interest Groups (incl. patient organisations)
	Citizen Scientists

Source: PathOS consortium

In developing the framework other entities will also be discussed and categorised. Based on the findings of the fieldwork, their categorisation might be updated.

[6] Ministries are a separate subcategory as there can be different positions towards Open Science

2. Theory of Change

2.1. Setting the framework

The Theory of Change (ToC) is a strategic framework that provides a systematic approach to understanding and catalysing social, widespread transformation. Rooted in the field of program evaluation and development, ToC offers a dynamic perspective on how interventions lead to desired outcomes within complex systems. This conceptual framework emphasizes the need to identify causal pathways, assumptions, and intermediary steps that connect inputs to impacts. As articulated by Anderson and Johnson (1997), the Theory of Change serves as "a comprehensive description and illustration of how and why a desired change is expected to happen in a particular context."[7]

European Research funders use the ToC framework to design their high-level interventions. This includes the Dutch NWO[8] and the British Arts and Humanities Research Council[9]; The EC uses a related intervention logic approach to describe the Horizon programme's key impact pathways.[10] Theory of Change can be used at different levels: in general (e.g., funding programmes) or specific contexts (e.g., project-level), which provides the required flexibility to analyse both system-level change as well as case-specific aspects in the wide range of activities under the umbrella of Open Science.[11]

At its core, the Theory of Change encourages practitioners, policymakers, and researchers to thoughtfully map out the **logical sequence of events** that drive change. It transcends linear cause-and-effect models by accommodating the intricate interplay of various factors, acknowledging the context in which change occurs, and recognizing the importance of both intended and unintended consequences. In essence, the ToC framework provides a roadmap for understanding how activities, outputs, and outcomes collectively contribute to the broader goals of a (policy) intervention.

ToC is at the basis of Impact Analysis. We consider the framework of ToC appropriate to assess and exploit available and future data as well as to provide an easily understandable framework

[7] There are many theories for describing and analysing change, In that sense the name for this theory is confusing.
[8] https://www.nwo.nl/en/impact-plan-approach
[9] https://www.ukri.org/who-we-are/ahrc/who-we-are/theory-of-change/
[10] https://research-and-innovation.ec.europa.eu/system/files/2023-05/swd-2023-132-monitoring-evaluation-he.pdf
[11] https://www.tipconsortium.net/wp-content/uploads/2022/02/MOTION-Handbook-180222.pdf

for a consortium, consisting of professionals from multiple disciplines. ToC also provides the required flexibility to the framework to adapt changes in an organised manner according to found evidence.

2.1.1. Definitions

Before we go into more detail, we describe the common definitions throughout the project on the elements of the ToC in Table 3.

Table 3 Elements of Theory of Change

Terminology		**Explanation and OS example**
	Problem	What we want to solve, incl. why it is a problem and who it affects, defined by the intervention initiator with wide consensus. *OS: Paywalled articles obstruct free and easy access to results of publicly financed research.*
	Context	The setting in which the intervention takes place, incl. stakeholders involved. *OS: The publications market faces a lock-in because the Journal Impact Factor raises barriers for new entrants.*
	Assumptions	Underlying conditions and expected behaviour that are assumed for the planned change. *OS: researchers are willing to cooperate in changing the current publication system.*
Intervention (in control)	Inputs	The actual collective resources that are allocated (by design or unintentionally) to the activities related to a defined intervention or initiative. Resources can be financial, human, or material. *OS: Additional funding for hybrid Open Access – reimbursing Article Processing Costs.*
	Activities	Actions that lead to the (desired) outputs. *OS: Agreements with publishers on making journals Open Access.*
	Outputs	Directly related measurable effects of activities that can be controlled by the initiators of the intervention. *OS: Increasing number of research articles are freely available.*

Results	Outcomes	The likely or achieved short-term and medium-term effects of the outputs (mostly measurable social, economic and academic changes). There is only limited (or no) control by the intervention initiator. *OS: Trend that more academic journals switch to (Gold) Open Access.*
	Impacts	Long-lasting, elementary and wide-spread change that affects different segments of society, economy, and academia, and therefore the environment in general. Impacts can be caused by multiple stakeholder interventions, and multiple stakeholders can be affected. Impacts can be direct or indirect, intended or unintended, relate to behavioural and/or systemic changes. *OS: Academic impact of Open Access is that OA articles get more visibility and more citations and re-use.* *Societal impact of Open Access is more engagement by society in research activities and increased trust in science.* *Economic impact of Open Access is that new audiences from enterprises (incl. SMEs) have better access to new knowledge and can use this to innovate and improve their products and services.*

Definitions based on OECD (2023)[12], examples by Technopolis Group

2.1.2. Understanding change

In this subsection we briefly discuss the theoretical basis for analysing change.

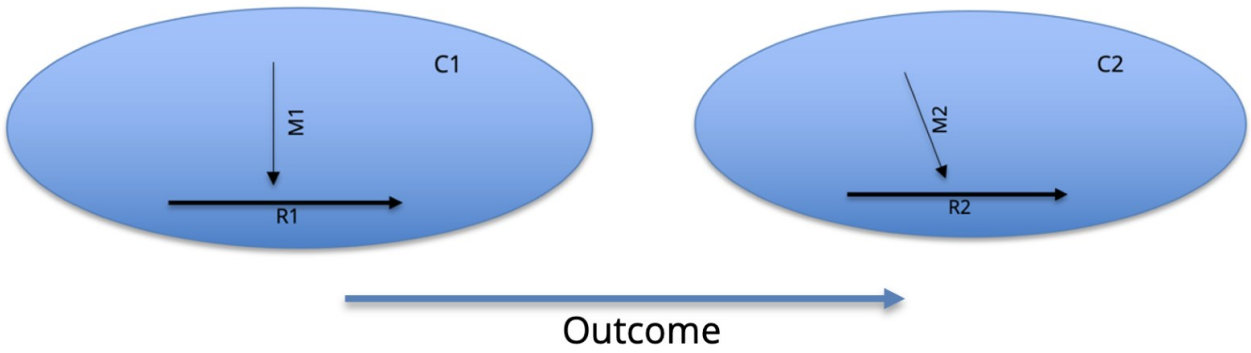
Whenever there is a problem of an ongoing regularity,[13] an agent (e.g., the government) wants to correct this by changing the mechanism that is steering the regularity. All this works in a specific context (Tilley, 2000).

[12] https://wikis.ec.europa.eu/display/ExactExternalWiki/Introduction+to+Results+and+Monitoring

[13] Regularity is something that happens in the usual way, or the fact of something doing this. In sociology it stresses the patterns and the regularities of social life which is, most of the time, orderly and largely predictable - https://dictionary.cambridge.org/dictionary/english/regularity.

Ideally, the effect of the intervention[14] can be studied in a separate context and compared with the old situation. In this case, the mechanism of the change can be researched by comparing the two situations, see Figure 1.

Figure 1 Measuring the effect of an intervention



Source: Tilley (2000).

For example, if we want to analyse the effect of introducing Hybrid Open Access for research articles, we could measure the citations of articles before and after the introduction of Article Processing Costs. In practice we could do this by comparing two regions, or over time – but in both situations there are contextual situations that are influenced by wider settings and with factors that are out of scope or were neglected during the analysis, cf. Tilley (2000).

2.2. Evaluating policies: impact analysis

After developing policies and doing interventions, it is crucial to understand whether these activities reach their objectives and create the intended impact. This is where evaluation and monitoring come into play.

In the context of Horizon Europe, for example, monitoring and evaluation of the programme are meant to provide evidence into the design of the programme, the performance of the programme, and for improving the programme.[15] Monitoring and, subsequently, evaluation, follows the steps of the Theory of Change by assigning, e.g., indicators and specific evaluation criteria to the various steps of the ToC, for example through performance and impact indicators.

[14] The act of interfering (of key decision making stakeholders) with the outcome or course especially of a condition or process (as to prevent harm or improve functioning) according to the available contextual background assessment.

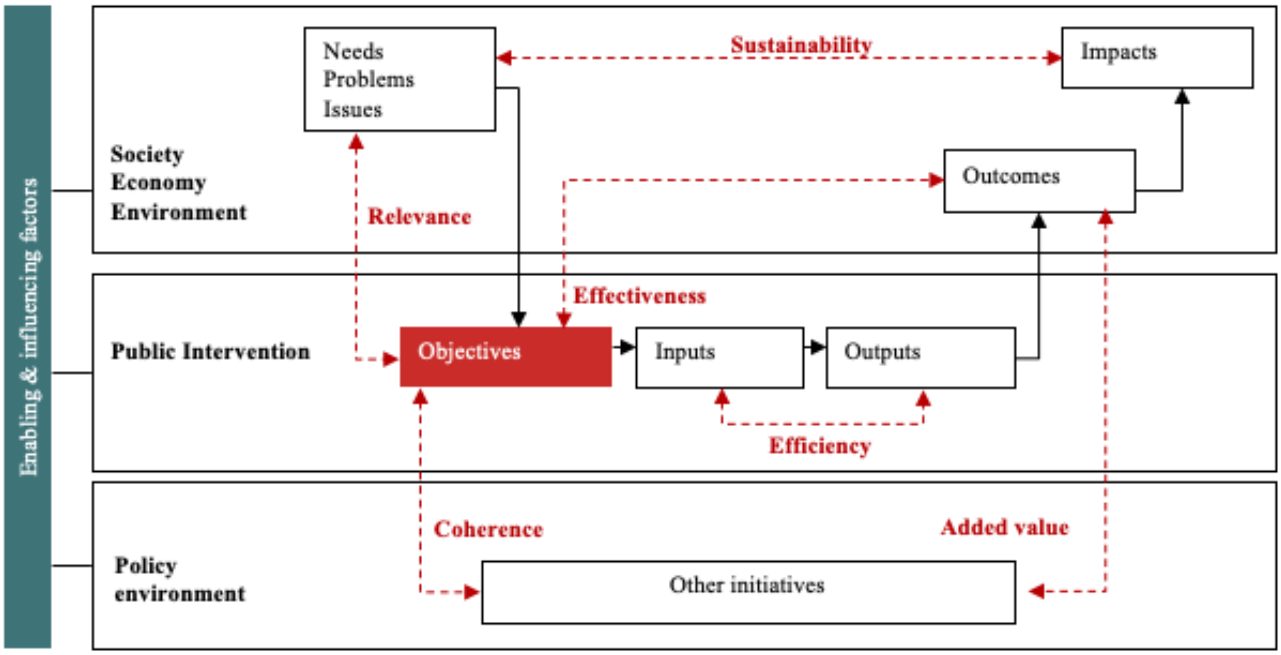
[15] https://research-and-innovation.ec.europa.eu/system/files/2023-05/swd-2023-132-monitoring-evaluation-he.pdf

According to OECD/DAC,[16] **impact analysis** is an important element of an evidence-based approach to policy making and defined as a systemic approach to critically assessing the positive and negative effects of proposed and existing regulations and non-regulatory alternatives.

OECD/DAC distinguishes six criteria in evaluating policies: **relevance, effectiveness, efficiency, sustainability, coherence and impact**.[17] Relevance is whether the intervention is doing the right things? Effectiveness is whether the intervention is achieving its objectives. Efficiency is about how well the resources are being used: how are the resources/inputs converted to results? Sustainability is about whether the benefits will last, for how long and how? Coherence is about how well the intervention does fit: is the project compatible with other objectives, plans and measures? Impact is about what difference the intervention makes. Positive as well as negative impacts, while primary and secondary (spillover) long-term effects may be produced; these can be direct or indirect, intended or unintended.

Figure 2 gives an overview of the criteria in evaluating policies, based on the ToC and the OECD/DAC Impact Analysis[18]. These can be considered stages in evaluation and impact analysis and as we will show in the next section, they can provide guidance for a framework of the Key Impact Pathways (KIPs).

Figure 2 Stages in evaluating policies



Source: Technopolis Group

[16] https://www.oecd.org/gov/regulatory-policy/ria.htm
[17] SECO/WE, Evaluation Guidelines, https://www.seco-cooperation.admin.ch/secocoop/en/home/results/evaluation.html
[18] https://www.oecd.org/dac/evaluation/daccriteriaforevaluatingdevelopmentassistance.htm

Note that impact is a final stage in the assessment cycle, while all other criteria describe relations between stages from needs until impacts. The figure above also includes a 'value added' relation in this figure, describing the relationship between outcomes and (the effects of) other initiatives. Hence, added value is defined here as the contribution of the other initiatives and relates to the context and environment. It is not a direct effect of our intervention and will not be considered as one of the evaluation criteria.

2.3. Impact pathways

The goal is to build impact pathways (D1.3) that will be used to evaluate the effects of Open Science interventions, making use of different inputs, like literature review, case studies, cost benefit analysis, allowing dynamics of different stakeholders' actions and causality claims. In order to do that we use the following building blocks: 1. Theory of Change, 2. Impact Analysis, 3. Key Impact Pathways.

Here Impact Pathways describe how the activities and outputs from these activities relate to outcomes and longer-term developmental changes (impact)[19]. Pathways are to indicate and provide assumptions of events, not under control of the executive body that implemented the intervention. In our Figure 2 this relates to the path from Inputs to Impact. Key Impact Pathways are defined as the most important or relevant pathways, where this relevance is supported by evidence. For example, one key impact pathway would describe the overall, proven impacts of Open Access scholarly publishing on academic, economic, or societal aspects. How to identify and validate these key impact pathways will be described in Section 3.

To be able to draw (key) impact pathways, we must formulate hypotheses of impacts (to set up the Theory of Change) that we can later test by finding evidence, including the causality claims (paths) between collected activities, outputs and outcomes. Hypotheses can be found in literature, especially policy documents, and via workshops with policy advisors, Open Science experts, and other stakeholders. Based on their objectives (and assumptions about policy effects) we can express hypothetical impacts of Open Science. Internally, we organised Mutual Learning Exercises to align on terminology and how to use the different types of information from literature, case studies, cost benefit analysis.

Mostly, these assumptions have a positive outlook, envisioning that Open Science can enhance the effectiveness and quality of scientific endeavours. For example, this would be achieved through increased public participation, which fosters trust through transparent practices, and

[19] NWO, Theory of Change and Impact Pathways, https://knowledge4food.net/wp-content/uploads/2020/05/2019arf-benin-countryworkshop_annex06.pdf

by involving the global community, thereby amplifying the economic advantages of scientific advancements for everyone involved.

However, it is crucial not to solely concentrate on the positive aspects. It is equally important to conduct comprehensive and critical research that identifies and addresses potential negative consequences, necessitating appropriate measures to mitigate them. For example, the introduction of Hybrid Open Access was used by (commercial) publishers to expand their market while keeping the Subscription business model, thus enlarging turnover and profits[20].

To construct and understand these pathways we need to:

1. Understand the social, economic and academic territorial context that might predetermine causal mechanisms and their degree.
2. Identify key elements of pathways (input-activity-output-outcome-impact).
3. Describe how they are linked and work together.
4. Allow for feedback loops and interactions of stakeholders.
5. Validate the causalities and links through evidence and expert groups.

2.3.1. Types of impact

Especially for impacts, we can add many characteristics and alternative definitions. We will use the OECD-UNDP (2000) version that defines impact as[21]: "Results of a programme or project that are assessed with reference to the development objectives or long-term goals of that programme or project; changes in a situation, whether planned or unplanned, positive or negative, that a programme or project helps to bring about." In addition, impacts can be direct or indirect or have primary and secondary effects. Examples include higher standard of living, increased food security, increased earnings from exports, increased savings owing to a decrease in imports. For Open Science, impacts can be increased trust in science, catalysing and improving innovation for (small) enterprises, addressing and solving problems from the Sustainable Development Goals.

When designing a framework of impact assessment, it is advisable to conceptualise what impact is to be assessed in order to make it fit for purpose (ESF, 2012[22]). We will distinguish three different types of impact, see Table 4):

[20] https://www.coalition-s.org/why-hybrid-journals-do-not-lead-to-full-and-immediate-open-access/
[21] OECD (2000), GLOSSARY OF EVALUATION AND RESULTS BASED MANAGEMENT TERMS, WORKING PARTY ON AID EVALUATION, 33rd MEETING, Paris, 22-23 November 2000
[22] ESF (2012), The Challenge of Impact Assessment, http://archives.esf.org/coordinating-research/mo-fora/evaluation-of-publicly-funded-research.html

Table 4 Types of impact

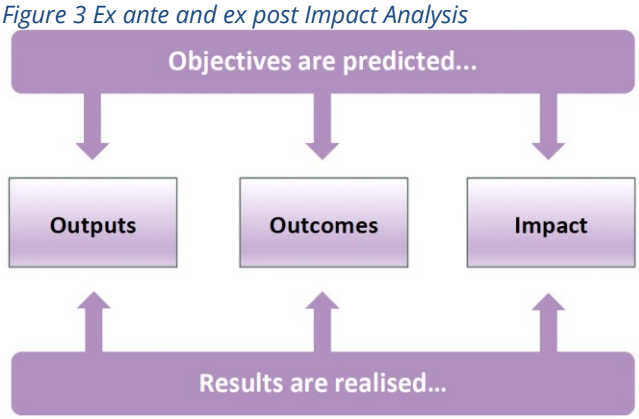
Academic impact
Scientific impact: contribution to the subsequent progress of knowledge, the formation of disciplines, academic training and capacity building.
Training impacts: contribution to curricula, pedagogical tools, qualifications.
Societal impact
Social impact: contribution to community welfare, quality of life, behaviour, practices and activities of people and groups.
Political impact: contribution to how policy makers act and how policies are constructed and to political stability.
Environmental impact: contribution to the management of the environment, for example, natural resources, environmental pollution, climate and meteorology.
Health impact: contribution to public health, life expectancy, prevention of illnesses and quality of life.
Cultural impact: contribution to understanding of ideas and reality, values and beliefs.
Economic impact
Economic impact: contribution to the sale price of products, a firm's costs and revenues (micro level), and economic returns either through economic growth, productivity growth or innovative capacity growth (macro level).
New knowledge: investment that may not be monetised immediately.
Technological impact: contribution to the creation of product, process and service innovations.

Source: based on ESF 2012, p.7, Technopolis Group

2.3.2. Timewise directionality of evaluation

The impact analysis can be ex ante or ex post, see Figure 3:

Figure 3 Ex ante and ex post Impact Analysis



Source: Technopolis Group

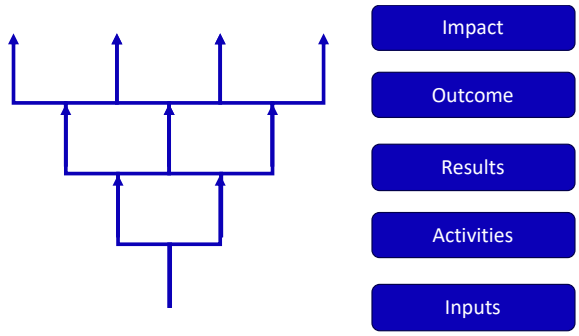
- Ex-ante[23] or predictive evaluation predicts the most likely outcomes and impacts of an intervention.
- En ex-post[24] or descriptive evaluation focuses on already realised impacts of intervention on. In a way we go back the chain towards the (most) effective interventions.

In both situations we need to collect information for the whole impact pathway, bearing in mind the multitude of inputs that influence one impact and the diluting effects of one activity (cf. section 2.3). In case of ex-ante evaluation, we need to find information by benchmarking similar interventions and what impacts it had in different contexts. In case of ex-post evaluation, however, we need to find evidence for causality between the intervention and an impact (and to measure the causality). Figure 5 is a representation of the directionalities of the two different types of approaches.

2.3.1. Diluting and converging effects

In practice the connection between actions (inputs, activities) and impact is not necessarily straightforward. On the way to impact, effects might spill over into other areas.

Figure 4 Spill-over effects

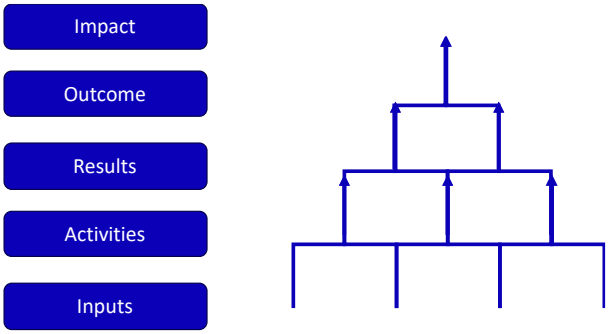


Source: Technopolis Group

Figure 4 displays how an individual input or activity can lead to different types of impact. Over time additional effects may also occur and add to the 'direct effect' of an action on impact[25].

Moreover, different inputs and activities can converge to a single impact, meaning that the individual contribution of a specific activity may be difficult to identify and measure.

Figure 5 Converging effects



Source: Technopolis Group

[23] predict outcome/impacts before the intervention
[24] Evaluation and assessment of the impact after the intervention
[25] Source: Dekker (2019, p. 112)

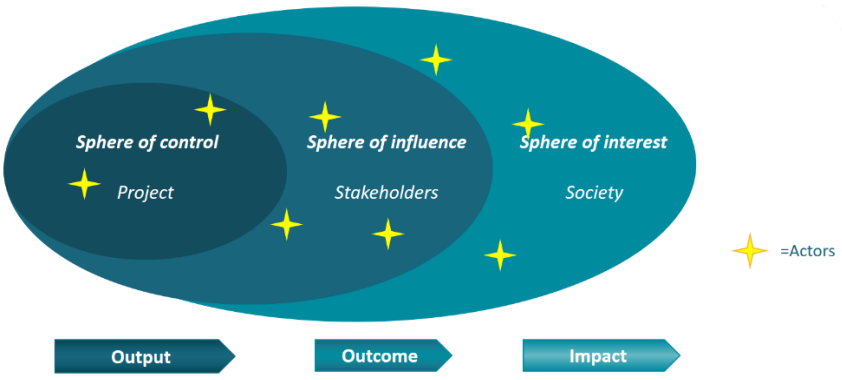
This complexity of causal chains is often neglected in impact analysis, showing linear effects and simple diagrams between action and impact and not allowing for reactions or feedback loops. This oversimplification is one of the main criticisms of the Theory of Change and Impact Pathway models. For example, Rogers & Weiss (2007, p65/66) criticise that just connecting the boxes with inputs, activities, outputs, outcomes and impacts without explaining how the relations work doesn't explain the relations.

2.3.2. Other (complexing) factors

Another complexing factor is that output from one stakeholder serve as inputs for other stakeholders. For example, publishing via Gold Open Access is an output of their intervention for research funders, but for publishers and for researchers this gold OA functions as an input to their activity of publishing.

Interventions might be obstructed by legal, technical or economic barriers. This becomes troublesome if these barriers are outside the sphere of influence, see Figure 6. For example, the Journal Impact Factor (JIF) created a lock-in situation: careers of researchers, rankings of universities are based on this JIF. Hence, researchers are hesitant of publishing in new Open Access journals, because these lack a JIF, or get assigned a low JIF in the beginning. This JIF-dependency obstructed the introduction of new journals based on new business models.

Figure 6 Spheres of control



Source: NWO

A path from inputs to impacts suggests simple and linear effects. In reality, this is hardly the case. To capture this in designing an intervention logic for a policy, we can rely on narratives as basis for describing the path and to take risks and challenges into account, spell out the underlying assumptions, etc. Narratives also allow for feedback loops and complex interactions (or reactions) between stakeholders.

2.3.3. First results

Klebel et al. (2023) analysed a large sample of academic publications on the impact of Open Science (cf. D1.3). First results of this PathOS Scoping Review of Open Science Impact indicate that many articles discuss academic impacts, with a lot of attention to citizen science as an academic impact. There were ample articles identifying societal impacts, across a variety of types, including educational, engagement and empowerment benefits for participants and their communities, and the creation of data for use in governmental monitoring and administering of environments and natural resources, although for political impact there was less evidence. For economic impacts we found less evidence. These results may be due to the fact that our research focussed mainly on academic literature. In the next phase of literature review we will focus more on grey literature and policy documents.

To fully understand the effects of interventions, we need to include causal mechanisms and provide explanations why and how impacts occur. Such approaches are particularly helpful for the production of policy-oriented recommendations and highlighting the necessary conditions for a given public initiative or intervention to be successful.

In the next chapter we will operationalise our framework for the impact pathways.

3. Operationalisation

In this chapter we will apply the theoretical approaches of Theory of Change – Impact Analysis – Pathways to set up our framework for Open Science Impact Pathways. We don't start from scratch but will make use of experiences from existing projects, like the RI-PATHS project and the EC Environmental Impact Assessment (European Commission, 2014). The strength of this latter framework is that it allows the tracking and visualisation of impacts including academic impact, economic or societal impacts in an intuitive fashion and clustering pathways across different themes and levels.

First, we will briefly explain the RI-PATHS approach and then propose a stepwise approach to apply for our analysis of pathways.

3.1. Baseline: RI-PATHS

The RI-PATHS Horizon 2020 project developed a set of impact pathways for research infrastructures (see https://ri-paths-tool.eu/en). Its approach serves as a baseline model for the approach of PathOS. It was chosen because it was designed for the impact of research activities and research organisations and therefore close to goal. Moreover, the visual frameworks and online tool provided by RI-PATHS can be used as guiding examples for PathOS outputs.Figure 1Figure 7 displays the general framework of RI-PATHS.

Figure 7 RI-PATHS example of impact pathways

(Source: Griniece et al., RI-PATHS 2020. Licensed under CC-BY 4.0 SA)

RI-PATHS complemented the framework by sets of indicators for the different pathways developed by for each step[26] such as resource/activity/outputs/outcome and impact. It further developed an online tool to compose new specific pathways based on user needs and it includes 10-15 high level key impact pathways relevant to potential users.

PathOS aims to develop a similar approach to identify high-level impact pathways useful for different user groups (e.g., national policy makers, research organisations etc.) but also to capture the complexities, and diversity of these pathways. During our project we will explore how to make our project outputs[27] searchable, interactive and dynamic, working closely with WP5 (Engagement and capacity building, especially T5.3 Exploitation of project results).

3.2. Framework: stepwise approach

To identify pathways, and to account for the complexity of causalities, interactions between stakeholders and feedback loops in these pathways, we propose a stepwise approach.

How we operationalize the identification of (key) impact pathways for Open Science in three steps is detailed in the following subsections.

3.2.1. Step 1: Identification of pathways

As a starting point, we focus on the identification of impacts of the interventions relevant in the broad range of the eight European Union's Open Science policy priorities (cf. chapter 1).[28]

For each of the eight policy priorities – which are considered fields of activities – we will identify impact sequences from activity to output, outcome and impact. This will be based on the full-text screening and themes identified through the literature review and case studies. Based on this information, impact pathways with information on inputs, activities, outputs, outcomes, and impacts can be constructed.

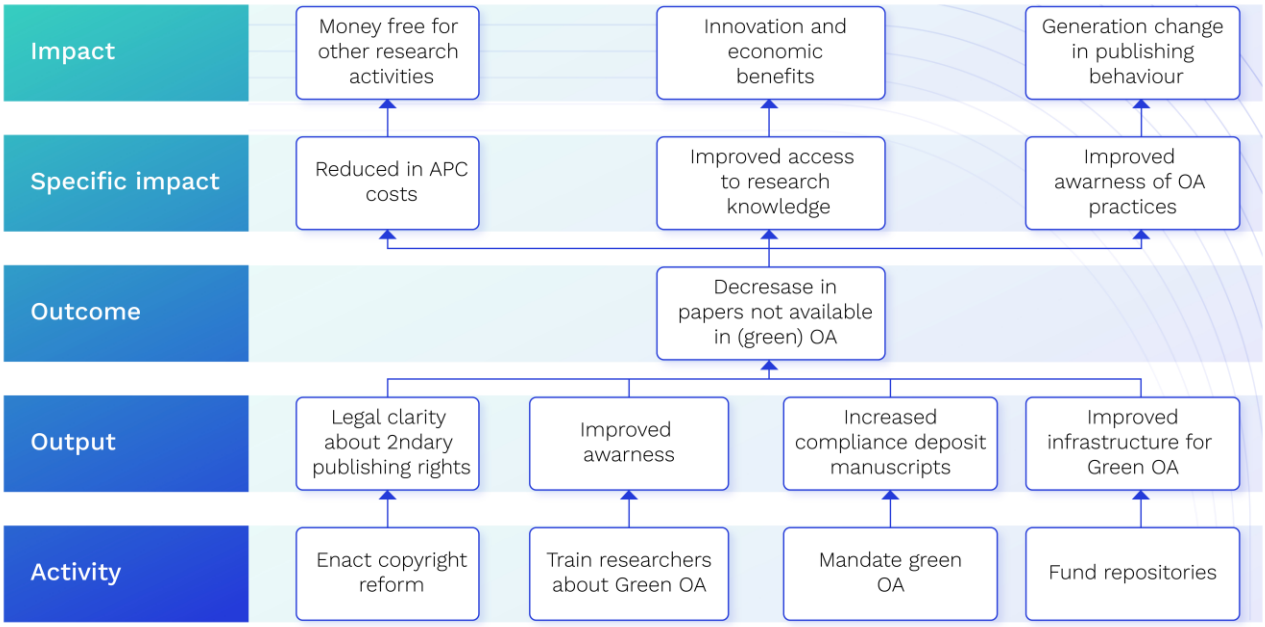
For example, stemming from a single case study or research article, but also several combined sources. It should be noted that the approach to arrive at key impact pathways focuses on **cumulating identified outcomes and impacts.**

[26] Resources + activity are slightly different in our terminology, where resources might be a part of the context (narrative, wider setting that will serve as an input, need that creates the basis for the objective) but also as an intervention if it's under direct control. Activity is an input in our terminology while we don't divide short and long-term outcomes, we rather distinguish outputs (short-term within control) from outcomes (Directly related, mostly measurable social, economic and academic changes that can be somewhat controlled by the intervention initiator).
[27] Handbook of Open Science Impact Indicators, an online registry with a portfolio of use cases and the results of a Cost Benefit Analysis framework.
[28] This does not preclude the addition of further areas under investigation later in the project.

In constructing pathways, one can start with setting up or objective trees[29]: They are the result of brainstorming, literature review or other data collection methods. Activities and outcomes give the (short term) outcome (cf. the converging effects in figure 5), and from this outcome there can be several (specific) impacts (cf. the spill-over effects in figure 4).

Figure 8 presents an example objective tree for activities fostering Green Open Access[30]. The tree lists a range of activities, their outputs, the overall outcome, and the associated impacts in this area. Activities can include legal reform, training activities, policy mandates, or funding activities. These create specific outputs, which lead to a specific outcome – the increase of papers available in Green Open Access. In this fictional example, this outcome creates a range of different impacts, for example cost savings and further economic benefits through the improved access to knowledge. It also illustrates that activities and outputs can be relatively easy to measure. Even the outcome can be measured, although the effect size of each individual activity to the outcome may be more difficult to establish.

Figure 8 Potential Open Access objective tree



Source: Technopolis Group

[29] For more information on objective trees as a tool for strategy and project design, see https://wikis.ec.europa.eu/display/ExactExternalWiki/Problem+and+objective+tree

[30] Green open access refers to self-archiving as an OA model. Earlier (and final) versions of a paper are deposited in a searchable repository.

As this example shows, literature review and desk research can likely yield extensive numbers of individual pathways – originating from these objective trees – which describe activities and outputs at a very detailed level.

But note that these objective trees are not the only way to determine pathways. Another approach is to start at a higher level, e.g., assuming FAIR data is a given and modelling the outcomes and impacts of this, without looking at specific activities and inputs.

Both approaches can be useful for this analysis. The former, because it contains information about activities which may in turn be relevant to identify risks, feedback loops and other underlying assumptions. The latter because it provides a focus on the specific impacts of Open Science activities.

Especially the objective tree approach can be used to find feedback loops by introducing multiple stakeholders. Activity of one can influence the activities of another stakeholder, and sometimes outputs of a stakeholder serve as input for other stakeholders. Feedback loops may also occur from the interaction with the wider environment. This is not only discovered by literature review, but also by having workshops high-level focus group discussion[31] where we provide space for stakeholders to comment and to make suggestions on our framework.

3.2.1.1. In practice

This first step requires a systematic collection of evidence. through literature review and in addition case studies and other methods. To work systematically and be able to scale up, this requires templates and protocols on how to process the literature and how to describe the case studies. For this we set up templates for both the literature scoping and the case studies. Both are coded with the terminology of input/activity/output/outcome/impact.

It is also important to make the assumptions and the hypotheses explicit and to indicate which part of the pathway is being covered. It can be expected that not all literature will offer information for the entire sequence. Instead, it might zoom in on and only describe the effects of one specific activity or zoom out and provide evidence of contextual determinants. This will make results comparable and suitable for aggregation at a later stage.

3.2.2. Step 2: Defining cumulative pathways

Whereas Step 1 is about collecting as much evidence as possible, Step 2 is about discovering trends and high-level or cumulative impact pathways. In other words, the aim is to find evidence

[31]PATHOS Focus Group Key Takeaways (2023): https://pathos-project.eu/a-year-in-review-the-pathos-focus-groups-key-takeaways

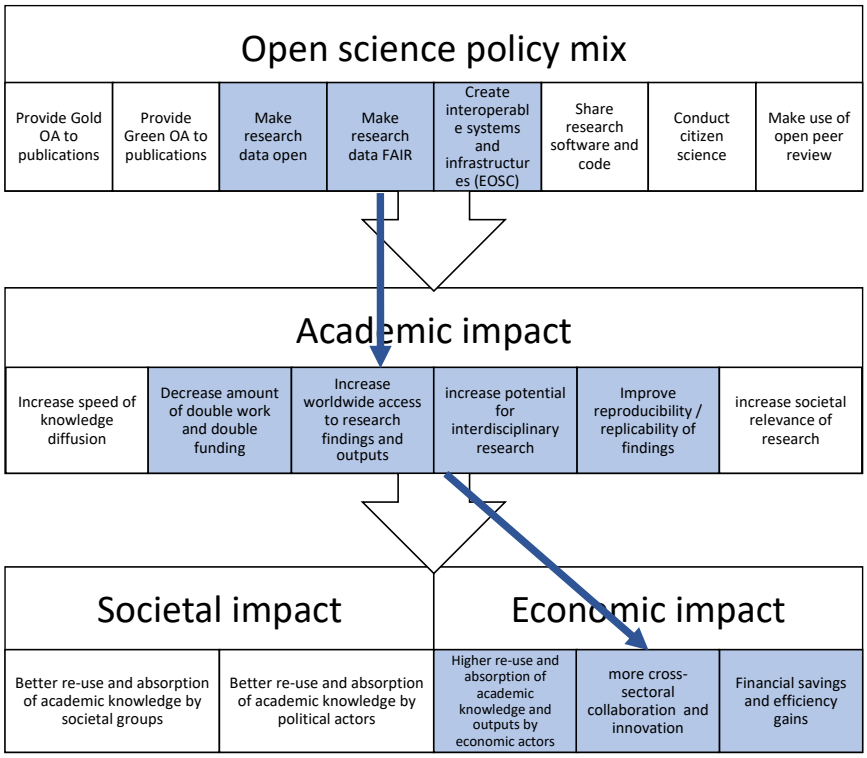
for cumulative impacts of different interventions that merit from the findings of the literature scoping papers and case studies.

We do so by adapting a framework for the evaluation of environmental impact of research and innovation (European Commission et al. 2014) to the area of Open Science. This approach focuses on identifying the outcomes and impacts of individual interventions and clustering them across different themes.

To achieve this and to define key impact pathways from the individual pathways from step 1, we will need to aggregate and zoom out to a higher level. We will do this by taking the outcomes and impacts identified in step 1 and cluster them under appropriate themes covering academic/economic/and societal impact. *Cumulative impact pathways* will exist where various activities lead to similar types of impacts. Examples could be efficiency gains or financial savings across different Open Science areas or wider reception of research results highlighted by citation advantages of OA/Open Data publications.

Examples for the framework are given in Figure 9 and Figure 10. The first hypothetical example displays impact pathways for a set of data-related interventions.

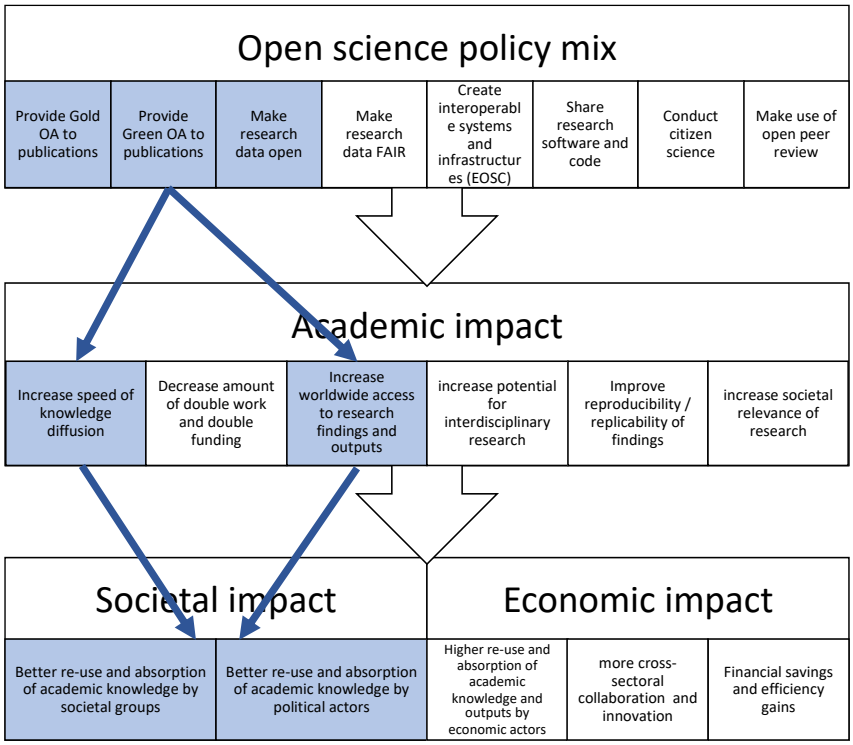
Figure 9 Example 1 of cumulative impact pathways



Source: Technopolis Group

The second example shows impacts of open access to research outputs.

Figure 10 Example 2 of cumulative impact pathways



Source: Technopolis Group

The framework allows visual tracing of the logic of impacts from intervention to impacts.

3.2.2.1. In practice

The practical work will include a detailed coding of activities and outputs to the different open science areas, and of outcomes and impacts to the areas of academic, economic, or societal impact. Next, cumulation will take place through collection of a range of such individual impact pathways and identifying the range of impacts associated with activities in a given Open Science area.

Based on empirical work, the initial classifications can be updated and validated throughout the research process in PathOS. For example, the initial hypothesis that academic impact is a mediator between intervention and societal and economic impact can be rejected. These issues will be addressed based on the information generated by the findings of the scoping review[32] and case studies. We also expect interactions between steps 1 and 2, as new information from step 1 may change the aggregated results in step 2. We will explore if the framework may also

[32] Klebel T. et al (2023). PathOS - D1.2 Scoping Review of Open Science Impact. Zenodo. https://doi.org/10.5281/zenodo.7883699

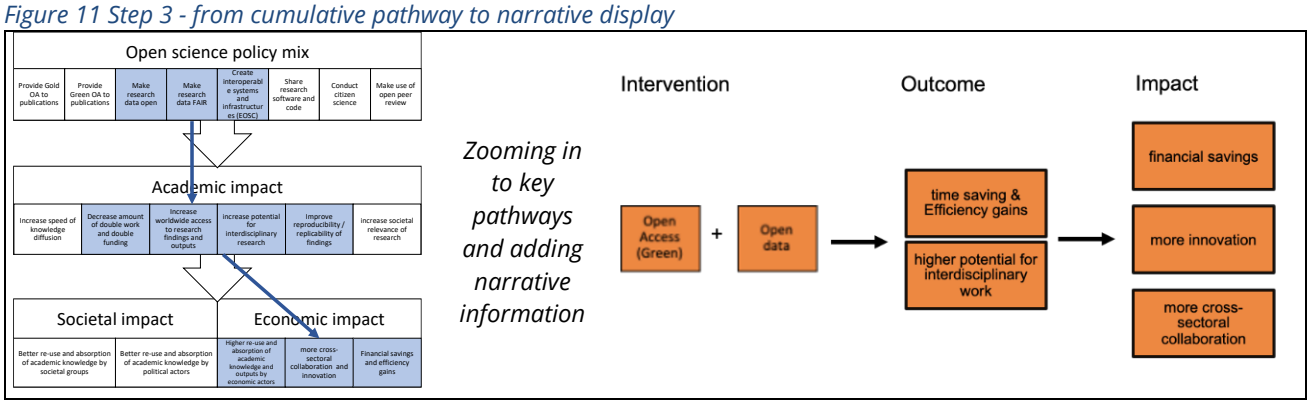
include information on the strength and directions – including feedback loops – of impacts or the stakeholders involved.

3.2.3. Step 3: Zooming in for narrative overviews

In the third step we provide detailed information about selected (Key) Impact Pathways, similar to the approach of defining high-level impact pathways in RI-PATHS. This way, we make the identified pathways a useful, actionable tool for stakeholders. For this, PathOS will zoom into specific pathways of relevance for stakeholders (to be identified) and show the relevant interventions, outcomes, and impacts. A specific impact pathway might for instance be focussing on sets of activities around openness of research results, on scholarly publishing, on research data, or on infrastructure. In this third step we can provide more detailed information, e.g., about the causalities, reactions from other stakeholders on an intervention.

Figure 11illustrates this process. We will select individual or cumulative pathways and the impacts only associated to those. Narrative details will provide information about contextual factors, indicators and other relevant information. Assumptions about the causal mechanisms, risks and relevant stakeholders can be added as well[33].

Figure 11 Step 3 - from cumulative pathway to narrative display



Source: Technopolis Group

This step is a crucial element for the use of the framework as it enables users to obtain detailed information about specific open science areas and the types of impacts these activities create. In the narrative part, we can include important contextual information, assumptions, indicators, and other details.

[33] This approach is also used in the 'Payback Framework' that was originally developed by Buxton and Hanney. In Donovan and Hanney (2011): Research Evaluation, September 2011 DOI: 10.3152/095820211X13118583635756.

3.2.3.1. In practice

Information for this 3rd step can come from literature review but also (and especially) from workshops, mutual learning exercises and other interactive meetings.

We will test this step in upcoming workshop events. In addition, we will explore whether the results can be supported by an online tool that provides an ontology of relationships between interventions, outputs, outcomes, impact, the relevant indicators and stakeholders. This would provide a learning system and speed-up upcoming impact analyses.

3.2.4. Example application of the stepwise approach

To illustrate the work explained in this stepwise approach, we present two impact pathways.

The first, depicted in Figure 12, is based on the ELIXIR case study. The specific case concerns the creation, curation and re-use of ELIXIR supported databases for research and industry.

The second intervention logic, displayed in Figure 13 is on FAIR data and derived from a report on the economic benefits of FAIR data (European Commission, 2019). The study itself focused on outputs, outcomes, and impacts and we added examples of inputs and activities.

For both cases, we associated the outcome and impact stage with the high-level areas of academic, societal, and economic, which is displayed using the colour coding indicated in the respective figures. This is a first step in identifying key impact pathways of open science in academia, society and economy.

In Figure 13, we also tried to set up a coding of activities and outputs according to the respective open science areas. By coding categories, we can scale up empirical work and it may contribute to ontologies to describe relationships (cf. Step 3). Specific areas for this case cover FAIR data [FAIR] and open data [OPEN]. It also concerns the establishment of interoperable systems and infrastructure [INTEROP] as well as re-use of data [RE-USE]. These codes are only in the early stages and will be refined in the further work.

In the next steps of PathOS, we will test this stepwise approach. In brief: Step 1: identify pathways based on intervention logic (the full cycle from intervention to impact); Step 2 zoom in and identify the paths that relate to a specific impact; Step 3: zoom out to develop causalities, feedback loops, interactions, based on narratives and mixed information sources. In the next steps we will also connect with the indicators that are being developed.

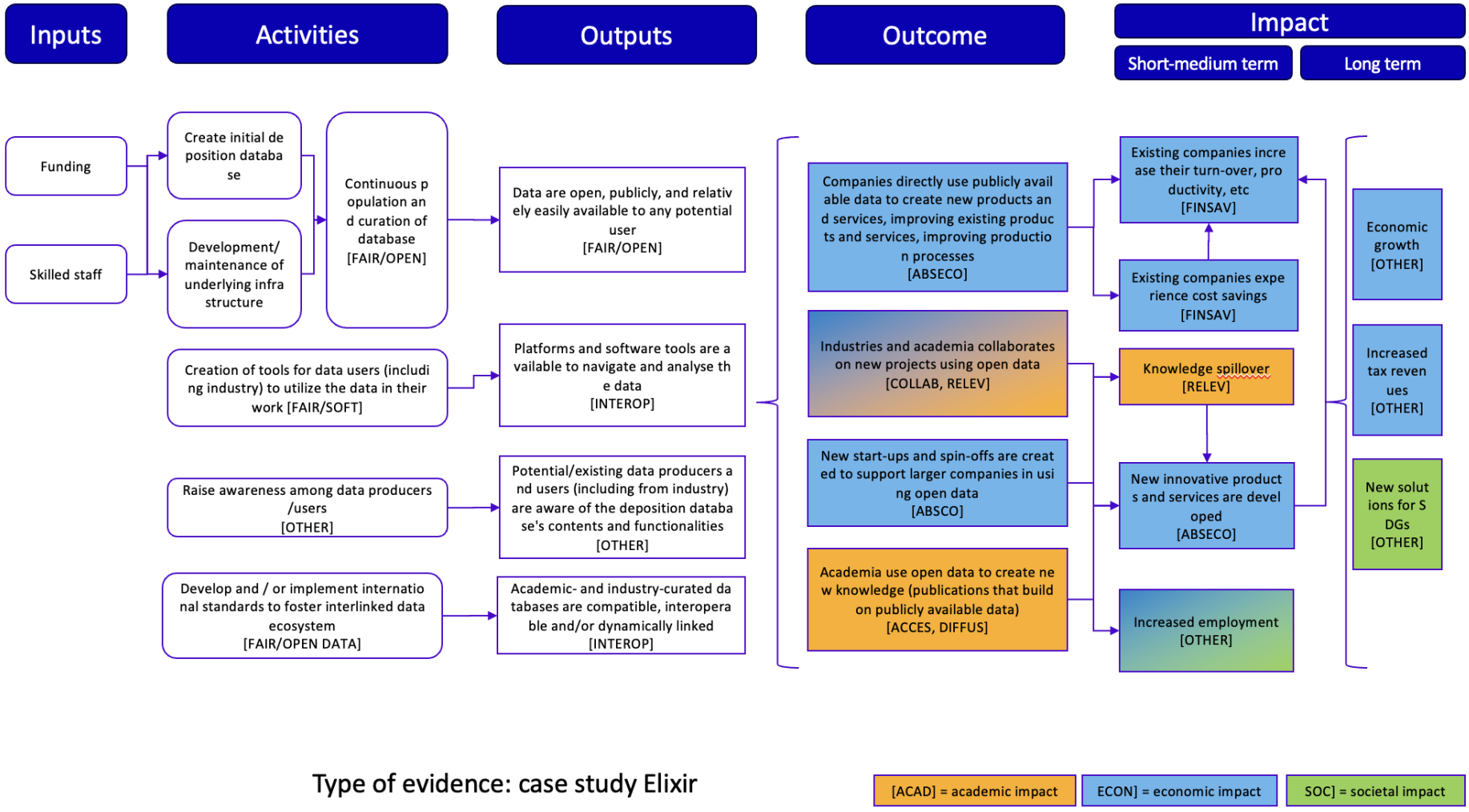


Figure 12 Intervention logic of the ELIXIR case study

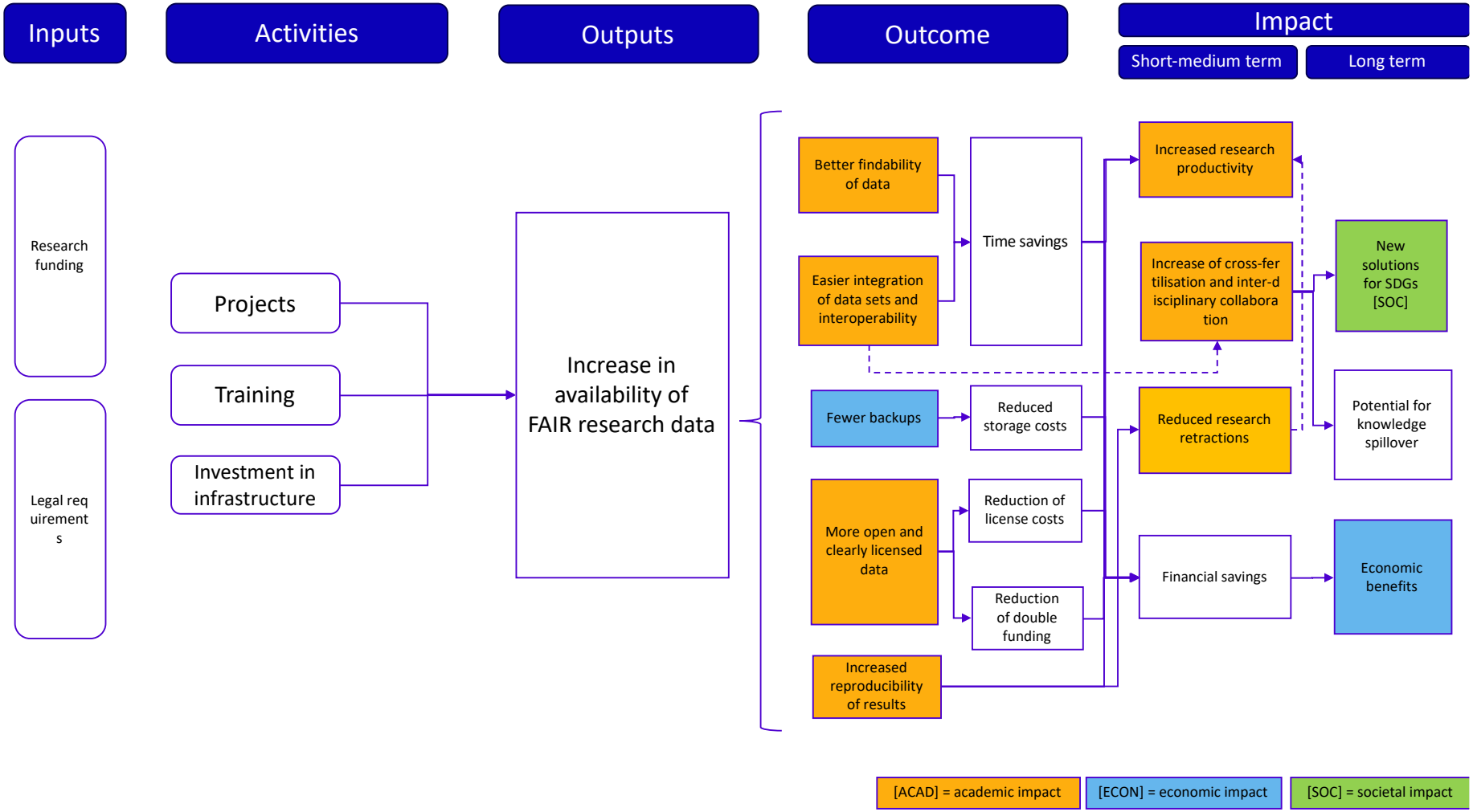


Figure 13 Intervention logic of FAIR research data

3.3. Empirical evidence and operationalisation

3.3.1. Literature review and case studies

For the first findings of the literature review, the related deliverable is available.[34] In addition, there is also a protocol how to find adjacent and related literature. For the literature review we set up a template in WP1[35] to systematically collect the information. In this way we can distribute the work[36] and still work in a consistent manner. This is explained further in D1.2, the deliverable on the literature review and in the Protocol for Scoping review document. Similarly, we set up a case studies template in WP3. See the annex for more details.

This systematic and transparent approach also gives opportunities to scale up and organise large teams for the literature review and case studies.

3.3.2. Interaction between Work Packages

Collection of evidence and operationalisation of the impact pathways are ensured throughout the project and the Work Packages have their respective methodologies and approaches to meet their described objectives and to contribute to the final objective of the project.

The objective of PathOS Work Package 1 is to develop the underlying conceptual framework for impact pathways. This is done through close collaboration with the other project WPs. WP2 will develop indicators for the impact of open science practices on different areas, WP3 will measure and test on selected cases and WP4 will carry out a Cost-Benefit Analysis. The results will provide feedback into the framework and provide important evidence, contextual factors, stakeholder feedback and other information.

Interaction with other WPs is a key element in the project. The results need to be validated both from theoretical and empirical angles, and it methods from one WP need to be suitable for other WPs to use as a guiding framework for their purposes.

[34] Klebel et Al (2023). PathOS - D1.2 Scoping Review of Open Science Impact. Zenodo. https://doi.org/10.5281/zenodo.7883699

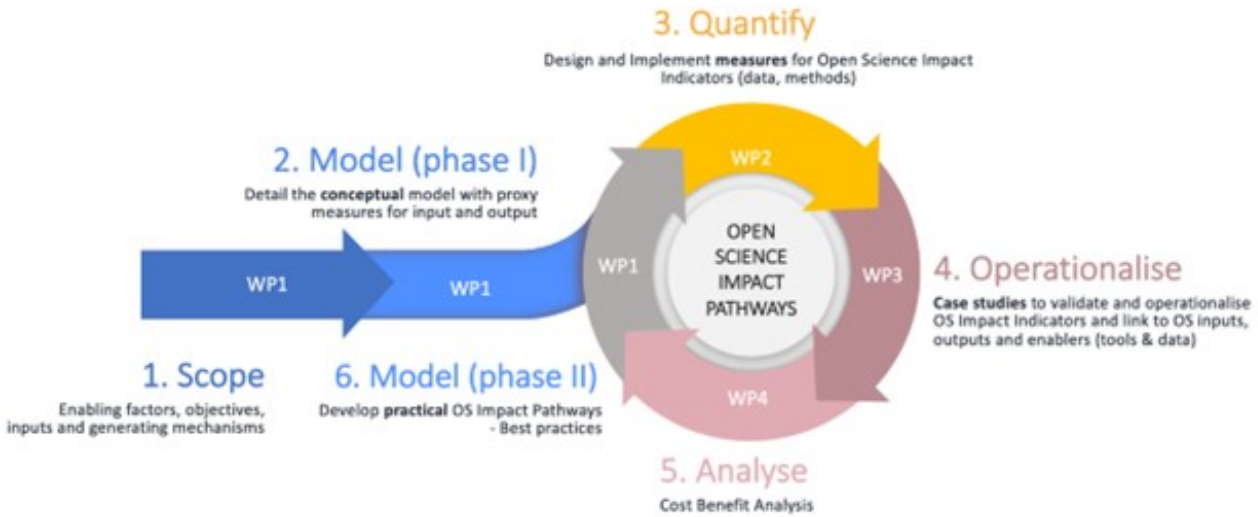
[35] Template is described under Chapter 2.4 in Klebel et Al (2023). PathOS - D1.2 Scoping Review of Open Science Impact. Zenodo. https://doi.org/10.5281/zenodo.7883699

[36] The literature scoping started with over 6,000 articles. Even after filtering the number of articles to review was too big for a single person to process – instead a team of 6 project staff members carried out the work.

The interaction is also useful for mutual learning that creates understanding for all project partners on key principles, the expected scope, our common objectives and the sustainability of the outcomes.

Feedback loops are built-in into the PathOS workflow as depicted in Figure 14. Activities 2-5 will take place in a phased approach, and centred around selected case studies, each bringing on the table a range of OS specific elements what will allow us to better evaluate and measure them and be able to construct the pathways from start-to-end.

Figure 14 PathOS workflow



Source: Technopolis Group

In phase 1, WP1 provides, through this deliverable, the initial conceptual model for Open Science impact pathways. This model informs methods and initial working definitions of the other project WPs. During phase 2, after work has been done on indicators, case studies, cost benefit analyses etc, WP1 takes stock of the collected evidence and prepares a set of evidence-based key impact pathways in collaboration with other related WPs through already used practices, as Mutual Learning Exercises, Workshops and by utilising Expert Panels and international expert fora.

4. Summary

In this report we describe the stepwise approach to find, analyse and present impact pathways and to make full use of the information that is available from different sources – literature, case studies, cost benefit analysis. First, we do 'traditional' impact analysis taking the EU policy actions as starting point, but then we zoom out and cluster impact pathways to give the overview. In the third step we zoom in again to describe causality and relationships. In a way we disentangle the impact analysis into separate steps.

We benefitted from the parallel work in WP1 on the literature scoping and the case studies in WP3 in order to test this stepwise approach. Among the lessons learned is that we should categorise and have taxonomies as much as possible: stakeholders, actions, instruments and set up templates to systematically analyse and to be able to scale up.

Ideally, our approach should also provide a framework to do future evaluations.

References

European Commission, Directorate-General for Research and Innovation, Peter, V., Arnold, E., Doranova, A., et al. (2014) A short guide to assessing environmental impacts of research and innovation policy. Publications Office. https://data.europa.eu/doi/10.2777/5779

European Commission, Directorate-General for Research and Innovation, (2019) Cost-benefit analysis for FAIR research data: cost of not having FAIR research data. Publications Office. https://data.europa.eu/doi/10.2777/02999

European Science Foundation (2012) Evaluation in Research and Research Funding Organisations: European Practices. http://archives.esf.org/index.php?eID=tx_nawsecuredl&u=0&g=0&t=1698317086&hash=eb8a9e1d288524d7352ef8a552e04d775c00e501&file=/fileadmin/be_user/CEO_Unit/MO_FORA/MOFORUM_Eval_PFR__II_/Publications/mof_evaluation_final.pdf

Elina Griniece, Jelena Angelis, Alasdair Reid, Silvia Vignetti, Jessica Catalano, Ana Helman, Matias Barberis Rami, Henning Kroll. (2020). Guidebook for Socio-Economic Impact Assessment of Research Infrastructures. Zenodo. https://doi.org/10.5281/zenodo.3950043

Klebel, Thomas, Cole, Nicki Lisa, Tsipouri, Lena, Kormann, Eva, Karasz, Istvan, Liarti, Sofia, Stoy, Lennart, Traag, Vincent, Vignetti, Silvia, & Ross-Hellauer, Tony. (2023). PathOS - D1.2 Scoping Review of Open Science Impact. Zenodo. https://doi.org/10.5281/zenodo.7883699

Patricia J. Rogers, Carol H. Weiss (2007), Theory-based evaluation: Reflections ten years on: Theory-based evaluation: Past, present, and future, https://doi.org/10.1002/ev.225.

Tilley, Nick. (2000). Realistic Evaluation: An Overview, https://www.researchgate.net/publication/252160435_Realistic_Evaluation_An_Overview .

Annexes

Indicators & Metrics

We distinguish between indicators, metrics and measurement:

- An indicator for the concept that we are trying to say something about, for example uptake of Open Access.
- A metric is an operationalisation of the indicator, for example: number of Green Open Access publications.
- A measurement of the proposed metric is the source for the metric, for example use data collected by Unpaywall to determine Green Open Access status for publications.

Open Science templates, academic impact templates, societal impact templates, economic impact templates and reproducibility templates were made by WP2 to collect indicators and metrics on those subjects.

Case Studies Template

Table 5 Case studies template

	CASE STUDY	ADDITIONAL DETAILS
Name of case study		
Short description Please make clear what is the object of your case study. For instance, it deals with the provision of an open software/repository/platform/library/databases, a combination of open resources (e.g. databases, cloud computing, etc), a research collaboration/partnership, etc.?		
HYPOTHESIS / Impact Pathway **(i.e. what effects do we want to examine, what are the confounding factors, etc)** **Pathway logic for (that follows the logic of D1.1, in line with the project glossary)** **The definition of clear pathways will be taken up by WP4 to define CBAs.**		

General (e.g., effect of OS on Gender)		
Specific OS Instruments/Treatment (e.g., open access mandate, open access resources, policy)		
Objectives (What was the goal of this policy/OS instrument?)	-	
Enabling and confounding factors (e.g., gender equality improving over time in any case)	-	
Stakeholder groups affected and/or included (e.g., funders, scientists, businesses, public institutions, students, governmental agencies, citizens, general public, etc.)	**STAKEHOLDER GROUP (CHOOSE FROM D1.1)**	**DETAILS (AFFECTED DIRECTLY/INDIRECTLY, WHY, ETC)**
Geographical Coverage (e.g., Europe, Portugal)		
Domain Coverage (e.g., life sciences)		

Time Range Please specify (i) when the intervention was developed (e.g. when the platform/repository was set up), (ii) when it started its operation (e.g. the platform was launched and became public available), (iii) when it started/will start to produce impact, (iv) for how long do you expect impact will keep on materialising (n. of years):	Setting up (Start/End Year): Operation (Start Year): Impact (Start Year): Expected impact period (N. of years):	

Impact Targets (INDICATOR THEMES) *(e.g., publication output, engagement in social media, uptake of innovations)*	**SHORT-TERM**	**MEDIUM-TERM**	**LONG-TERM**
SCIENCE	[in WP2 spreadsheet, economic impact tab]		
ECONOMY/INDUSTRY	- [in WP2 spreadsheet, reproducibility tab] - [in WP2 spreadsheet, economic impact tab]	- [in WP2 spreadsheet, economic impact tab]]; [in WP2 spreadsheet, economic impact tab]] - [in WP2 spreadsheet, economic impact tab]	The below indicators will most likely to be done via cost-benefit analysis (with WP4) [in WP2 spreadsheet, economic impact tab]: Jobs creation in the industry sector Creation of start-ups/spin-offs Increase of turnover Increase in tax revenues Regional Increase in productivity (e.g. through cost savings, increase in profit margins

To be refined

	SOCIETY			Socially relevant products and processes (e.g. bioinformatics applications of societal benefit in health, food security, the environment, and more broadly the Sustainable Development Goals) [most likely to be done via cost-benefit analysis (with WP4) [in WP2 spreadsheet, economic impact tab]
Causality Narrative (First thoughts on proving that the correlations seen in the data have a causal relationship)				
Causality Methodology ***To be refined***				
DATA				

Quantitative or Qualitative Data sources (e.g., interviews, surveys, OpenAIRE, user data, etc.): **make sure to add information on financial data***

*For financial data we refer to (i) costs related to the set up/development of your intervention as well (ii) costs borne to operate it. Specifically, we are interested in knowing the following information:

(i) are these costs available or shall be reconstructed?

(ii) Are these costs covered by one institution or shared between different ones?

(ii) Can this information be shared for the purpose of CBA analysis or are they restricted to circulation?

To be refined

DATA SOURCE	DESCRIPTION	AVAILABILITY (TO BE COLLECTED?), PRIVACY, AND OTHER COMMENTS

OPERATIONALISATION

Methods and Tools to be used (e.g., classification, data cleaning, author disambiguation, etc.)

To be refined

METHOD / TOOL	DESCRIPTION	COMMENTS

RELATED LITERATURE

Are you aware of any related literature/studies on the same topic or case study?		
FOCUS GROUPS AND MEMBER ENGAGEMENT		
Goals for first round Focus Group (What do you want to learn from them?) **View Focus Group tentative plan (what we do in each round) here.**		
Who should be involved? (looking for 8-10 people ideally, less if not possible)	PLEASE FILL THIS OUT: CASE STUDY CONTACTS.XLSX	
Questions and topics for discussion in first round	• General feedback on the case study, to refine it • General feedback on selected pathway logics for resource types	
Medium and long-term expectations for group engagement (how you envision them being involved across the lifespan of the project)	• At the start of PathOS, the Focus Group will be used as a sounding board for the case study, so as to refine it. • As PathOS progresses, the case study's Focus Group will be given the opportunity to provide feedback on how the case study is developing, initial results, as well as help with arising questions. • The Focus Group will also be used to provide participants for the project-wide events documented in "Deliverables that include input from Experts & Case Studies". • Towards the end of PathOS, the Focus Group will be used to disseminate, across ELIXIR, the knowledge created by PathOS,	

	to build their understanding of the needs of the industry sector, the pathways to impact, and enabling factors.	
COMMUNICATIONS		
Additional Contacts that could participate in case study-related workshops	PLEASE FILL THIS OUT: CASE STUDY CONTACTS.XLSX	

This systematic and transparent approach also gives opportunities to scale up and organise large teams for the literature review and case studies.

Case Studies Checklist

Table 6 Case studies checklist

Steps	Check
Step 0*: Familiarisation with the Concept of the Theory of Change	Based on D1.1 and in MS6 description, to adapt the standard Theory of Change to the case study, using the related glossary.
Step 1: Understand the Case Study Objectives	Have the objectives of the case study been clearly identified?
Step 2*: Contextualise the case study and the objectives	What is the context? Is it an ex-post or and ex-ante evaluation? Is there any potential initial triggering or hindering factor? Can we draw a clear hypothesis and/or a research question?
Step 3*: Stakeholder analysis	Which are the key stakeholders with power of impactful intervention? Which ones are relevant to define the key impact pathways?
Step 4: Draw a tailored, hypothetical key impact pathway	Based on MS6 description, D1.1 and D1.2, draw a KIP tailored to the case study that supposed to answer the research question or that is along with the hypothesis. This will also indicate potential causal links to be assessed according to the evidence.
Step 5*: Choose Relevant Indicators	Have relevant indicators for the case study been selected from the works of WP2 based on the objectives and the context? Indicate their potential places within the hypothetical tailored KIP.
Step 6: Define Measurement Metrics	Have the right (clear and quantifiable) metrics been selected for each indicator according to the works of WP2?
Step 7: Set Baselines	Have baseline values been set for each metric?
Step 8*: Gather Data	Is there a clear process in place for data collection? Is there a clear workflow set up with direct communication towards WP1, WP2 and WP4?
Step 9*: Assess Causality	Is there a systematic process to assess and rigorously validate the established causal links using collected data?
Step 10*: Analyse Data	Is there a standardized framework for data analysis across all case studies and other WPs will be able to assess and validate causality links? WP4 will be able to assess CBA, based on collected data?
Step 11: Interpret Results	Are there clear criteria for interpreting the results of each indicator?

Step 12: Adjust Based on Causal Findings	Are adjustments made based on the findings from the causality assessment? Is there adjustment needed to the overall KIPs, made by WP1?
Step 13: Implement Changes	Based on the results, are there steps outlined for implementing changes, if necessary?
Step 14: Repeat Measurement	Is there a plan for when and how to repeat the measurement of these indicators?
Step 15: Review and Refine Process	Is there a plan for reviewing and refining the indicator implementation process based on the experiences of the case studies?
Step 16: Cross-checking with other case studies	Is there an agreement between case study leads on commonalities and clear differences that can highlight different triggering and disabling factors? Can those contextual factors be quantifiable? If yes, please define.

Data chart

Table 7 Data chart (used in D1.2)

HEADING	DESCRIPTION
Author	Name of author/s
Date	Date article sourced
Title of study	Title of the article or study
Publication year	Year that the article was published
Publication type	Journal, website, conference, etc.
DOI/URL	Unique identifier
Exclusion	Out of scope, non-English, duplicate
Justification	If a study was deemed to be out of scope, a justification had to be provided.
Study details and design (if applicable)	Type of study, empirical or review, etc. Notes on methods used in study (whether qualitative or quantitative, which population demographics studied, etc.)

Types of data sources included	Detail the data sources
Study aims	Overview of the main objectives of the study
Relevance to which aspect of Open Science	Open Access, Open/FAIR Data, Open Methods, Citizen Science, Open Evaluation, Open Science General
Relevance to which aspect of impact	**Academic:** (Provisional list: Quality, Citations, Integrity, Equity, Collaboration, Trust, Efficiency, Productivity, Reuse) **Societal:** (Provisional list: Engagement, Participation, Education, Trust, Policy, Sustainable development goals, Gender, Diversity, Health, Climate/Environment, COVID-19, Equity, Empowerment) **Economic:** (Provisional list: Economic impact, Financial/monetary impact, Costs, Cost-Benefit Analysis, Input-output, Return on investment, Patenting, Innovation, Productivity, Saving)
Key findings	Noteworthy results of the study that contribute to the scoping review question(s)
Coverage	Optional field to note any relevant information about the level of coverage of the study, e.g., only specific countries, disciplines, demographics covered
Confidence assessment	Optional field to note any concerns about reliability/generalisability of findings (e.g., conflict of interest, potential biases, small sample sizes, or other methodological issues) within the study

Source: D1.2 PATHOS project

4.1. Project terminology/glossary

For the common understanding, through internal and external consultations and workshops, we started producing the project glossary to be able to have a common basis of understanding on the terminology that we are using. As the project consists of professionals from a wide-range of disciplines and sectors, this is essential to be able to effectively and efficiently cooperate and collaborate (Table 11.).

Table 8 Project glossary

Expression	Description	Source WP
Activity	Individual action that aims to successfully implement policy initiatives, related to the evaluated objectives.	WP1
Contribution or additionality	How much of the observed 'outcomes' are being caused by the intervention in the content of theory of change.	WP1
Data Source	A data source is a ready-to-use database that can be used to construct a metric or a variable. For example, there are multiple data sources to measure the number of citations to a scientific publication or the open access status of publications.	WP2
Effectiveness	The effectiveness of the inputs in converting to the right outputs and outcomes that help in achieving the intended objectives of the intervention initiator.	WP1
Efficiency	The degree of conversion from inputs to outputs and outcomes.	WP1
Evaluation	Is an evidence-based assessment of the extent to which an intervention is: effective in fulfilling expectations and meeting its objectives; is efficient in terms of cost-effectiveness and proportionality of actual costs to benefits; is relevant to current and emerging needs; is coherent (internally and externally with other EU interventions or international agreements; has added value from the policy makers' point of view	WP1
Fitness check	Is a comprehensive evaluation of two or more interventions usually in the same policy area that are related in some way (normally by sharing the same objectives or specific procedures) thus justifying a joint analysis.	WP1

Impact	Long-lasting, elementary and wide-spread change that effects different segments of the society, the economy, the academia, therefore the environment in general. All stakeholders are effected and therefore only horizontal and multiple stakeholder interventions can change the effects.	WP1
Indicator	When a metric is thought to represent some underlying concept, we say the metric is an indicator for that underlying concept. We sometimes say that the metric is a proxy for the underlying concept. For example, if the number of downloads of a software package is the metric, we may think it is a proxy for the usage of a software package, and hence the number of downloads is an indicator for the usage of a software package. There may be multiple indicators for the same concept. For example, the number of dependencies of a software package can also be an indicator for the usage of a software package. Similarly, the same metric can possibly represent multiple concepts and hence be an indicator for multiple concepts. For example, the number of downloads may also be an indicator for the ease-of-use of a software package.	WP2
Input	The actual collective resources that are allocated (by design or unintentionally) to the activities related to the policy initiative. For example: finance, organisation, legal framework	WP1
Intervention	The act of interfering (of key decision making stakeholders) with the outcome or course especially of a condition or process (as to prevent harm or improve functioning) according to the available contextual background assessment.	WP1
Intervention initiator	A key decision-making stakeholder that consciously initiates any kind of activity that aims in resolving a problem. Activities need to be initiated with a clearly defined objective and dedicated resources/inputs, originated by the intervention initiator.	WP1
Intervention Logic	The INTERVENTION LOGIC (or Logic Model) of an intervention explains and visualises the overall concept of the intervention and explains how change happens by starting from the identified needs for action to the anticipated impacts that would reduce the needs. To design a Logic Model helps you in thinking through your act systematically. This includes identifying the relationships of the intervention with relevant strategies, inputs, activities, outputs, outcomes and impacts. It provides a justification	WP1

	why your intervention should be undertaken, what it seeks to achieve and how it will do this and to what effect.	
Key decision making stakeholder	Those stakeholders (authority, institutional, corporate) that have the decision making competence on initiating policies with major impact on the observed thematic area.	WP1
Key Impact Pathways	Key Impact Pathways (KIPs) are a structured framework used to identify and describe the series of causal relationships or logical steps that connect a specific set of actions, decisions, or events (drivers) to their resulting impacts on society, academia, or the economy. KIPs are designed to provide a clear and visual representation of how various factors interact and contribute to specific outcomes.	WP1
Metric	We call a measurement of some outcome a metric. For example, the number of downloads of a software package, the number of downloads of a dataset, or the number of Mendeley readers of a publication are all called a metric. They should be differentiated from an indicator, which says something about the concept that a metric tries to represent. A metric can possibly be constructed from a data source, but if this is not the case, data might need to be gathered in alternative ways, for example by some automated techniques (e.g. text-mining of full-text publications), or via surveys, interviews or document analysis.	WP2
Object	Any object of interest in the context of this study. This can be for example a publication, a software package, a data set, a pre-registration, a patent, an infrastructure, a repository, OA policies, et cetera.	WP2
Objective	The goal that a public policy initiative wants to inherently achieve.	WP1
Outcome	Directly related, mostly measurable social, economic and academic changes that can be somewhat controlled by the intervention initiator. It needs a wider scope and engagement from the stakeholders of the intervention.	WP1
Output	Directly related measurable technical effects of activities that can be directly controlled by the initiators of the public intervention. For example Services, infrastructure, subsidies	WP1
Pilot activities	Individual actions in order to test programme or project ideas.	WP1

Policymaking cycle	The ideal policymaking is designed in a cycle manner that consists of Issue identification, Issue framing / agenda setting, Policy measure identification/ policy formulation, Ex-ante evaluation, Policy measure development & adoption, Implementation, Interim evaluation, Final/ Ex-post evaluation, Feedback.	WP1
Problem	Defined by the intervention initiator with wide-consensus within related actors.	WP1
Programme	Often a set of organised but often varied activities (a programme may encompass several different projects, measures and processes) directed towards the achievements of specific objectives. Some projects have evolved into programmes.	WP1
RACER framework	The framework that helps providing a guideline in case of an impact assessment. R= Relevant Is there a clear link between the indicator and the objective to be reached? Where change is being assessed, has baseline data been made available? A = Accepted Have specialists/stakeholders been consulted on the construction of indicators? Is the indicator actively used in connection with the intervention itself? C= Credible Are there definitions for each indicator including statements of what the indicator shows and how data are collected? E = Easy Are indicators updated to reflect changes to interventions? R= Robust Are there established criteria or characteristics for assessing the quality?	WP1
Rationale	The reason behind the public policy initiative.	WP1
Relevance	The direct connection of the needs and problems to the intervention objective.	WP1
Sustainability	Longevity of an intervention and its rationale.	WP1
Theory of change	Logical links (hypotheses) between a project or programme and the rationale for its funding.	WP1
Key impact pathway	A high-level description and logical depiction of the major outcomes and impacts of a specific intervention.	WP1