Publishing Reproducible Research Output

Literature findings

8th April 2021

https://doi.org/10.5281/zenodo.4647697

Knowledge Exchange

Report outline

Navigation

The following icon is used throughout this slide deck to bring key findings to the reader's attention:

Key finding

Introduction

Section 1

About this activity

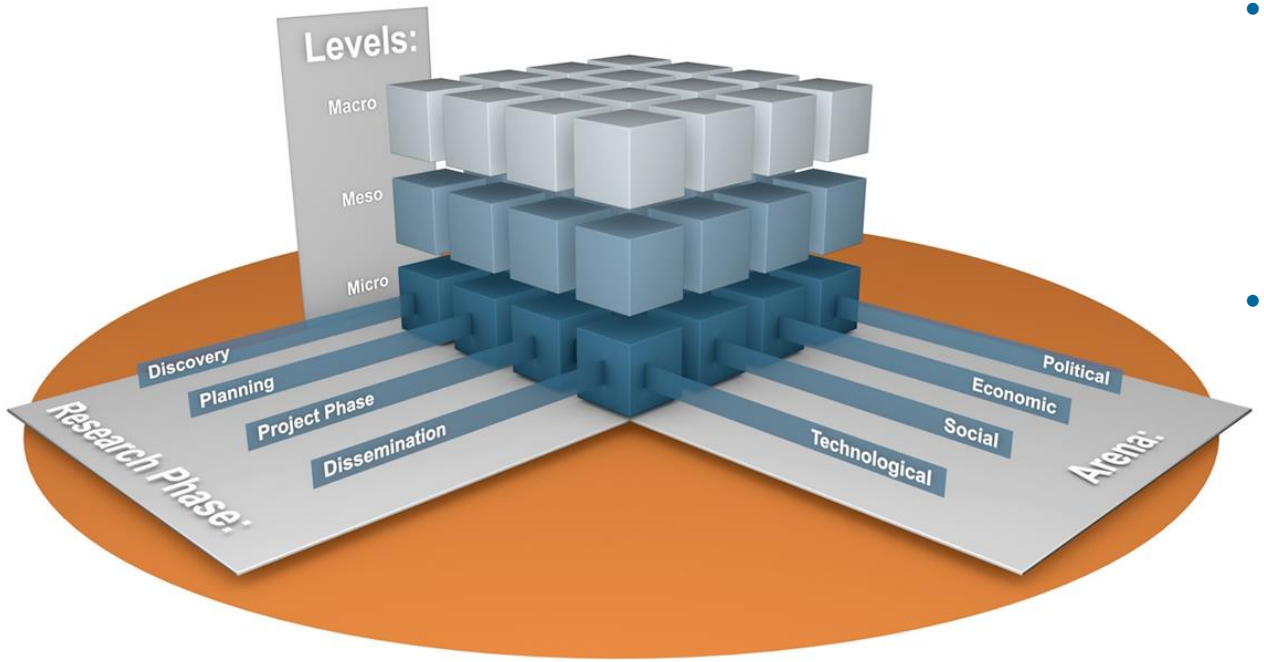
Aims and objectives

- This activity is funded by Knowledge Exchange (KE), an international partnership of organisations in the UK, Netherlands, Germany, Finland, Denmark, and France, as part of their work on Open Access.
- The purpose of this project is to conduct a gap analysis and investigate researchers' needs in order to make their published research outputs more reproducible. In addition, we seek to investigate how infrastructures (both technical and social) can support researchers in their reproducibility efforts.
- As a sub-goal, KE seeks to explore disciplinary differences and map different research areas on a spectrum of reproducibility.

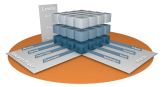
Task & Finish Group

- The activity is led by Juliane Kant (DFG) and Anna Mette Morthorst (DeiC).
- The Task & Finish Group for this activity consists of the following researchers and infrastructure experts:
  - Birgit Schmidt, Göttingen State and University Library
  - Jeroen Sondervan, Utrecht University
  - Birte Christensen-Dalsgaard, Aarhus University
  - Daniel Nüst, University of Münster
  - Matthew Jaquiery, University of Oxford
  - Pierre Carl Langlais, Paris Sorbonne-CELSA
  - Saskia Woutersen, Leiden University Library
  - Verena Heise, University of Oxford / Hanse-Wissenschaftskolleg
  - Yrsa Neuman, Åbo Akademi University

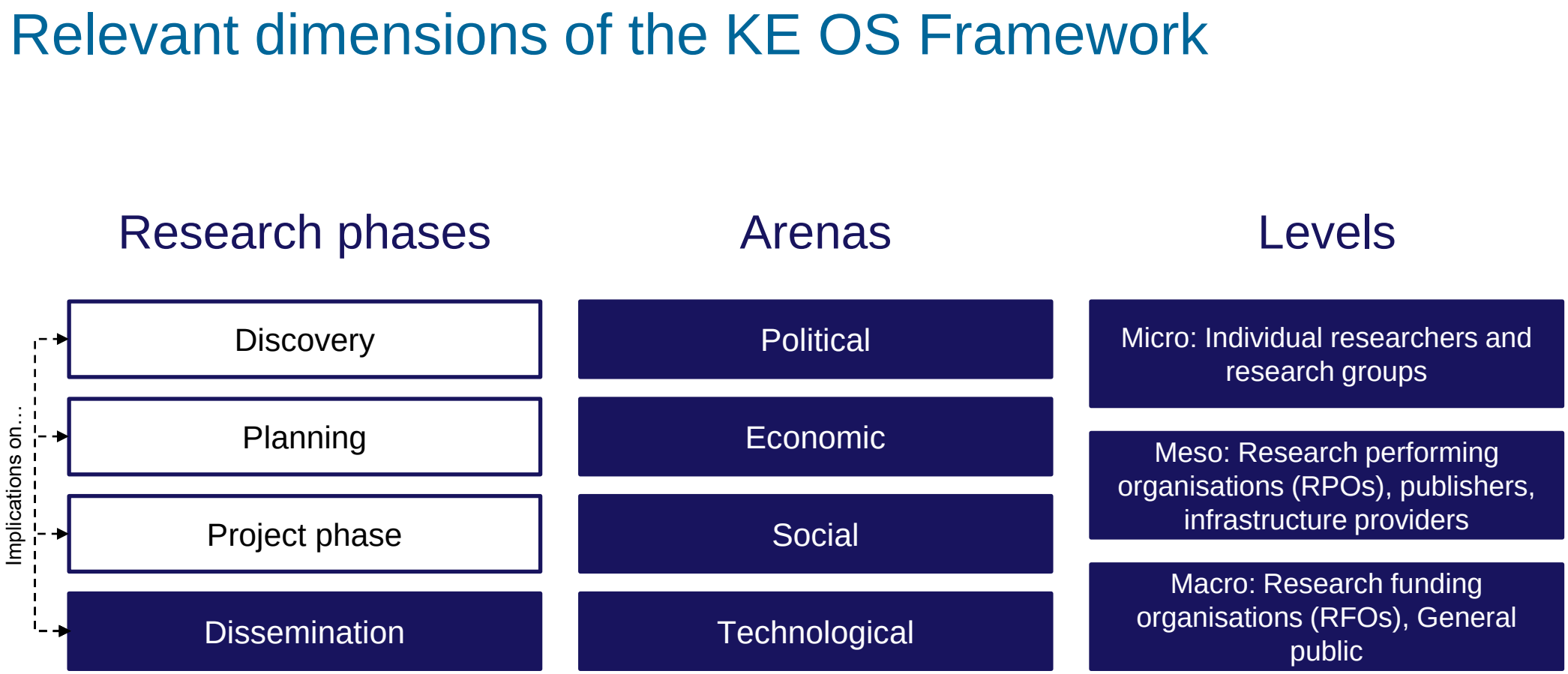
Applying the KE Open Scholarship (OS) Framework



- In this project, we used the KE OS Framework as a lens to study research reproducibility, focusing on the “dissemination” end of the spectrum.
- The framework is helpful in identifying the appropriate stakeholders (at different levels), arenas and research phases.
  - A smaller version of this chart is used for signposting in slides where the KE OS Framework is applied.



Relevant dimensions of the KE OS Framework



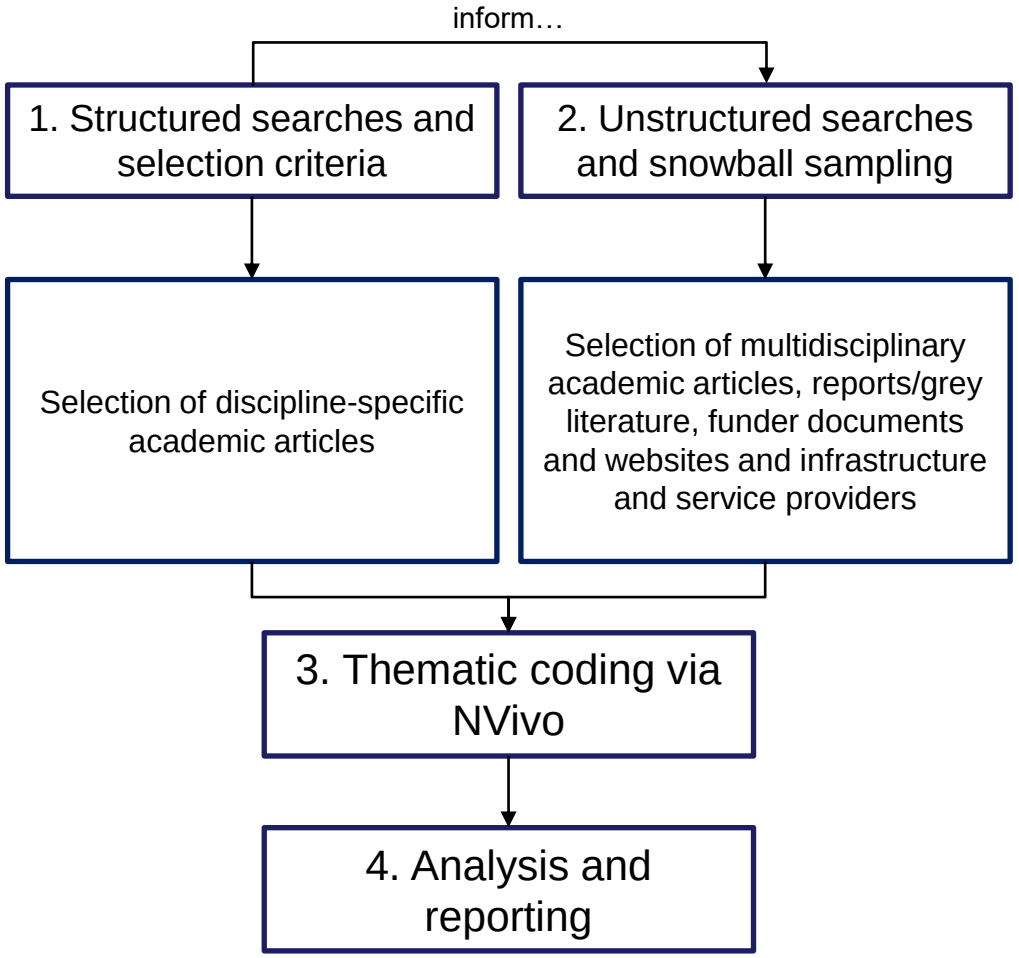
Legend

- Main focus of this activity

Methodology

Section 2

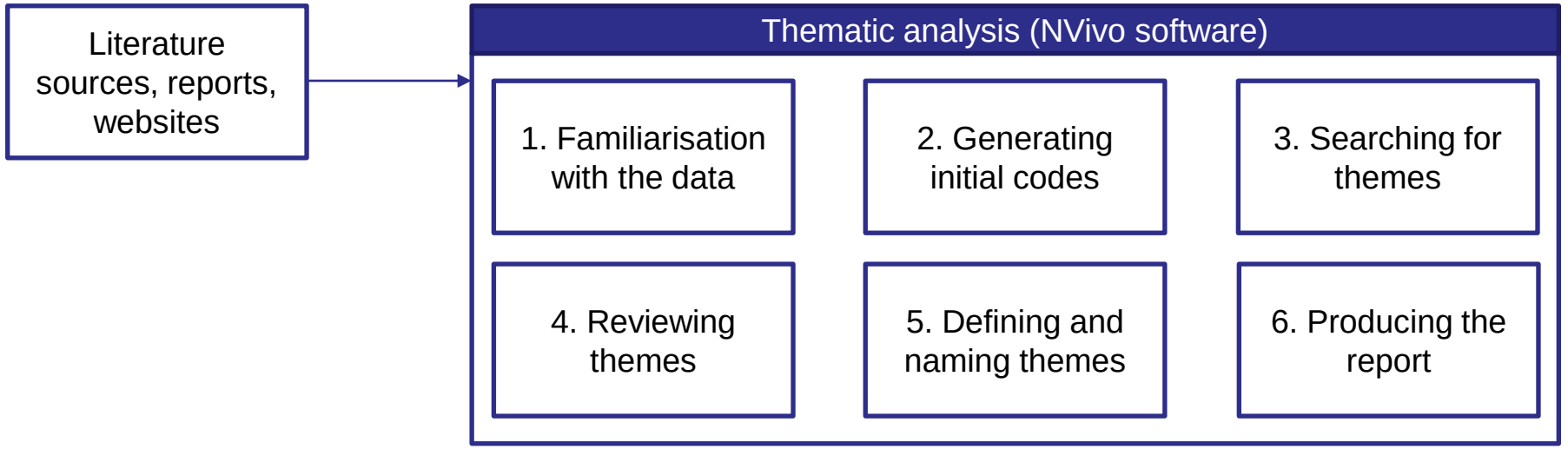
Literature review approach



- This project took an iterative approach to literature selection, considering the two pathways shown on this slide.
- Our research sought to develop a breadth of understanding across disciplines and stakeholder groups.
- The purpose of the literature review was to build an overview of:
  - definitions and language;
  - problems and issues;
  - recommendations and proposals put forward by the literature; and
  - ongoing initiatives, technical solutions and infrastructure.

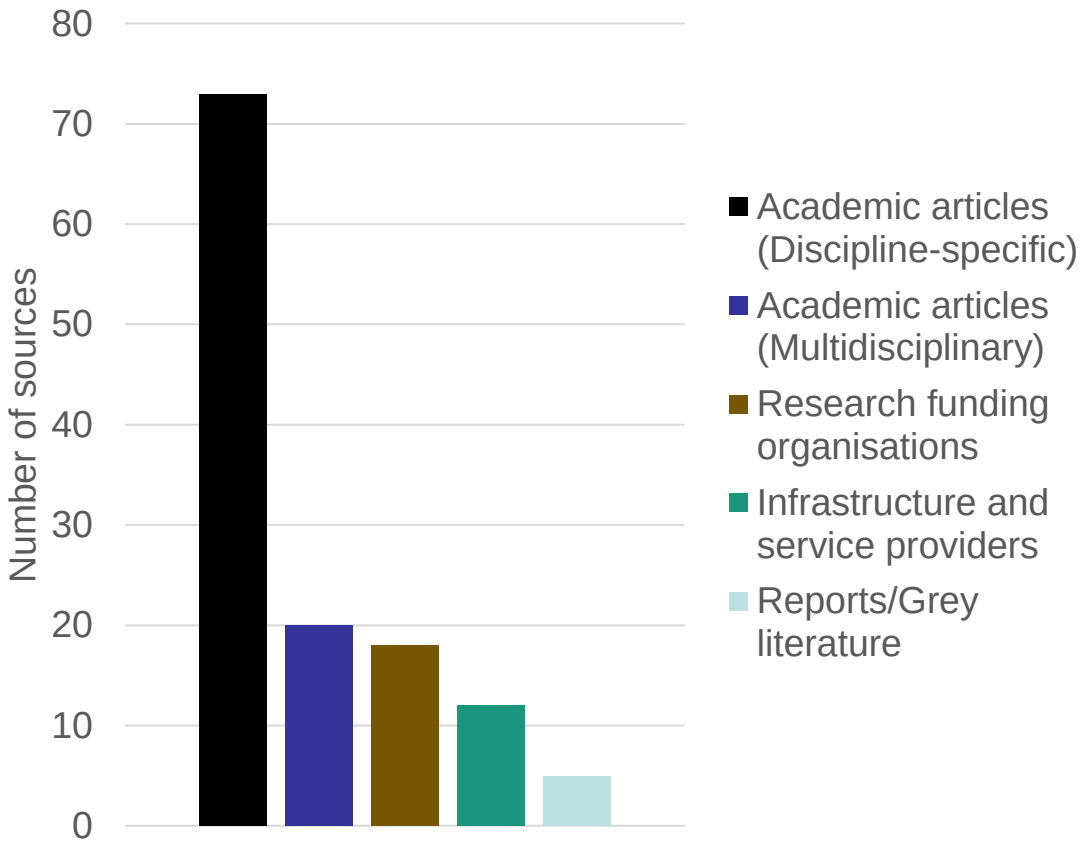
Approach to analysis

- This report was prepared following a process of thematic coding and analysis:



- The qualitative data analysis required us to prioritise coded findings and text extracts for the purposes of reporting. The prioritisation was based on the frequency of findings in the dataset, their relevance to the project's objectives and the individual judgement of the research team.
- All findings reported in this report arise from our literature review. Where applicable, the report uses the convention "n=#" to describe the number of sources that mentioned a given finding or insight.

Overview of sources considered



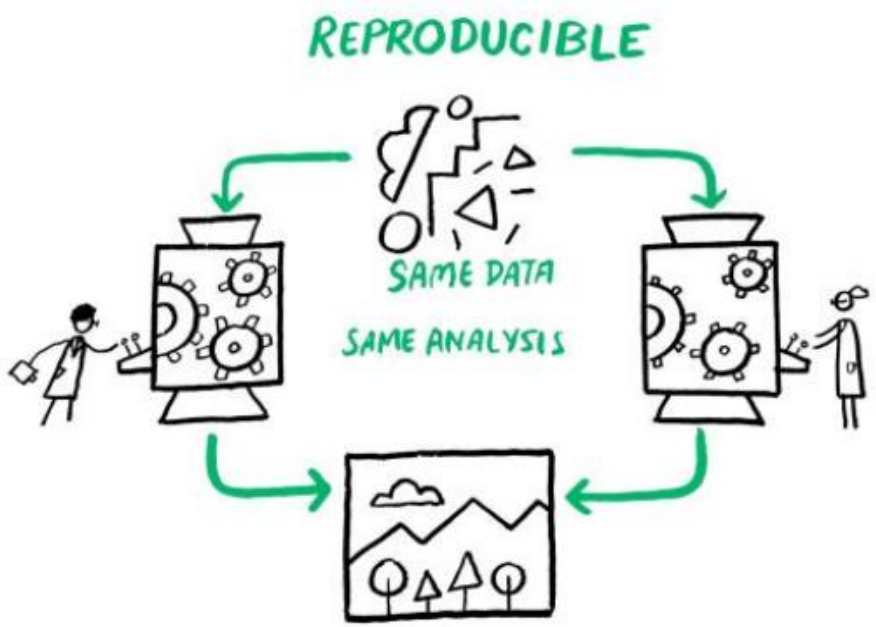
- This literature review considered 128 sources, focusing mainly on academic articles but also including websites and reports (particularly in the cases of research funding organisations and infrastructure providers).
- The taxonomy of disciplines considered in this study was created by building on and simplifying Scopus disciplines, in consultation with the project's Task & Finish Group, as follows:
  - Humanities
  - Life Sciences
  - Mathematics, Computer Sciences, and Statistics
  - Medical Sciences
  - Physical Sciences
  - Psychology
  - Social Sciences
- It should be noted that Psychology was considered as a standalone discipline due to its prominence in the research reproducibility discourse and because it may involve elements of medical, social and life sciences.

Definitions and language

Section 3

Our definition of «reproducibility»

In this KE activity, research reproducibility is defined as cases where researchers use the same data and procedures (e.g. code) shared by others to obtain the same results as in the original study.

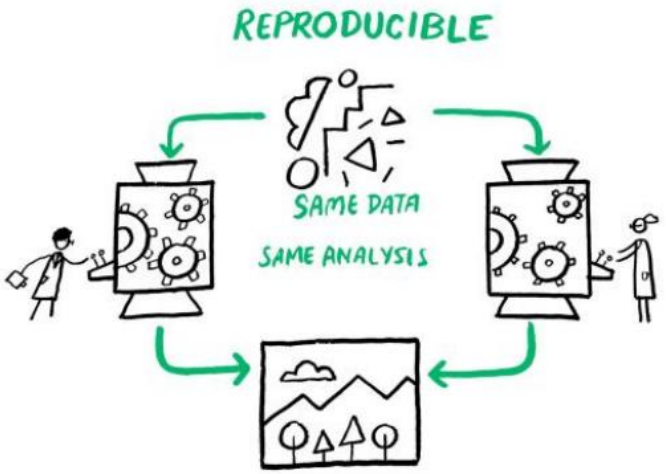


Definitions from the literature

	Views of reproducibility as defined in this activity	Other interpretations and usage
Humanities	"Reproducibility, which is the ability to compute the same result…"	"[A **replication** study] can be carried out in three forms: **reanalysis of existing data sets**, collection of new data with the same study protocol, or collection of new data with a modified study protocol."
Life Sciences	"…providing a Web-based environment in which users can perform computational analyses and have all of the details automatically tracked for later inspection, publication, or reuse."	"In some fields, **replication** may be accomplished by **reanalyzing a published data set."**
Mathematics, Computer Sciences, and Statistics	"Precisely speaking, we used a mix of replication and reproduction, i.e., we used both artifacts provided by the authors and our own artifacts."	""Reproducibility" describes the case in which similar ideas lead to similar experimental results given similar evaluations and scenarios, where "**similar results**" are results that allow the same conclusions to be drawn."
Medical Sciences	"The entire analysis workflow would be completely automated in a workflow engine and packaged in a software container or virtual machine to ensure computational reproducibility. All data sets and results would be assigned version numbers…"	""Results reproducibility, **or replicability**, "refers to obtaining the same results from the conduct of an independent study"."
Social Sciences	"Results reproducibility—the ability of others to reach the same results as the original paper using the data provided by the authors."	"Results reproducibility is conceptually analogous to **reliability** because it is about consistency. [...] Inferential reproducibility is conceptually analogous to **validity** because it is about making similar inferences based on the results."

Definitions from the literature

- As noted in the previous slide, there is no clear agreement around definitions. Different terms, such as “reproducibility”, “replication” and “replicability”, are sometimes used interchangeably, both within and across disciplines.
- We therefore acknowledge that it is difficult to find coherent terminology that speaks consistently to all stakeholder groups and disciplines. Key finding
- The remainder of this deck will only follow the definition on slide 12 to avoid ambiguity.

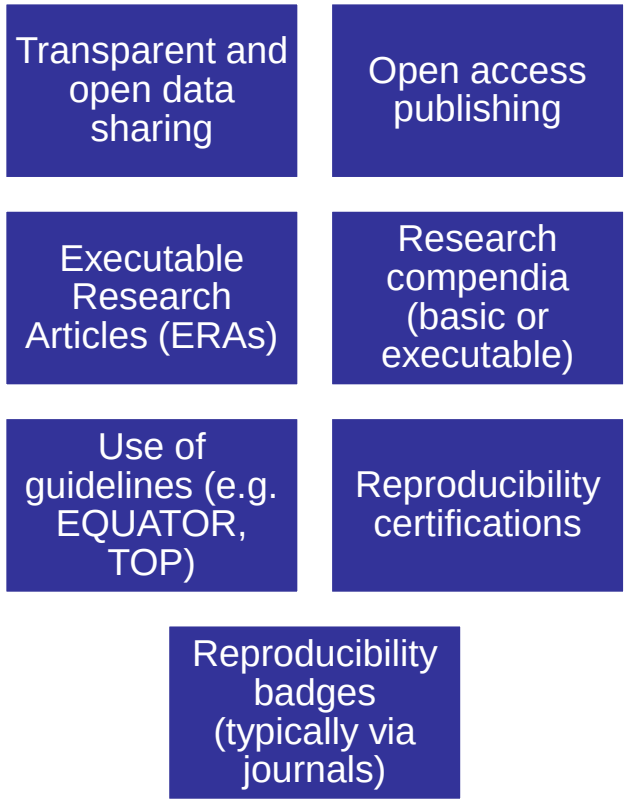


Forms and benefits of publishing reproducible research outputs

Section 4

Forms of publishing reproducible research outputs

- At the publication and dissemination stage, reproducible research practices can be achieved and/or supported via a range of pathways (see chart). There is currently no widespread preference for any of these, and individual researchers tend to follow disciplinary standards, where any are available.
- Although research reproducibility is desirable as a general principle, the literature makes some important points:
  - Reproducibility is an enabler of trust in and reuse of research, but it is not (nor should it be) the ultimate aim of scientific endeavours. Key finding
  - Open access and open data do not guarantee reproducibility by default. Some additional activities are key, such as the sharing of code and detailed methods or procedures. Key finding



Benefits of publishing reproducible research outputs

Micro level

- Reproducible practices can enhance transparency in the peer-review process.
- Reproducibility enables continuity of a researcher's work, as the approach is carefully documented, and the data is curated.
- Reproducibility helps to build a researcher's reputation by showcasing transparency and openness.

Meso level

- Journals can enhance their reputation, readership and prestige.
- There is more potential for RPOs to start new collaborations, including across traditional disciplinary boundaries.

Macro level

- Research funding organisations can ensure the continuity of funded work, as materials, data and code are shared by researchers.
- The general public can access a complete history of published research, strengthening perceptions of rigour, integrity and transparency of scientific findings.

- Improving publishing and reporting standards can improve the overall transparency of scientific endeavours.

Barriers

Section 5

Key barriers

- Our literature review identified a range of barriers to the publication of reproducible research outputs. We have grouped and categorised these based on their frequency in our evidence base:
  - **Key barriers** mentioned by a large range of sources are shown in blue boxes;
  - **Additional barriers** are shown in brown boxes and were less frequently mentioned in the literature we reviewed. However, they were deemed significant due to their implications across the research lifecycle and prominence in the broader open scholarship discourse.

- In the next slides, we map these barriers to the levels in the KE OS Framework (Macro, Meso and Micro). It should be noted that several barriers have implications across multiple stakeholder groups.

Key barriers

Rewards and incentives

- Current incentives are not conducive to reproducible science (n=24)
- Current evaluative metrics focus on quantity and impact rather than on quality and reproducibility (n=11)

Publishing and reporting standards

- Code, data and methods are not always available or complete (n=14)
- Lack of detail in methods (n=12) and poor documentation due to limitations on word/page counts (n=11)
- Need for better reporting standards for authors and peer reviewers (n=6)

Technical and analytical skills

- Lack of training and mentoring, including in computing analysis, coding and novel techniques (n=22)
- Lack of familiarity with computing and software for reproducibility (n=3)

Macro and Meso ⟷ Meso and micro

Additional barriers

Technical infrastructure

- The connectivity of scholarly communication infrastructure with researchers' workflow and research tools can be improved (n=2)
- There is room for enhanced collaboration and communication between research repositories and between repositories and journals(n=1)

Practical challenges in data sharing

- Authors may be unwilling to share data (n=5)
- Data sharing may give rise to significant issues (n=4) around:
  - Data ownership
  - Data sensitivity
  - Ethics and confidentiality
  - Intellectual property
  - Misuse of shared data

Meso and micro

Implications across the research lifecycle

- Given the scope of this activity, our literature search strategy focused on the publication and dissemination phase (see section 2).
- However, all the barriers we identified have implications across the entire research lifecycle: this indicates that phases cannot be considered separately when it comes to research reproducibility.

Key finding

Implications on…

	Rewards and incentives	Publishing and reporting standards	Technical and analytical skills
Discovery	✓	✓	✓
Planning	✓		✓
Project phase	✓		✓
Dissemination	✓	✓	✓

Threats to progress

System level

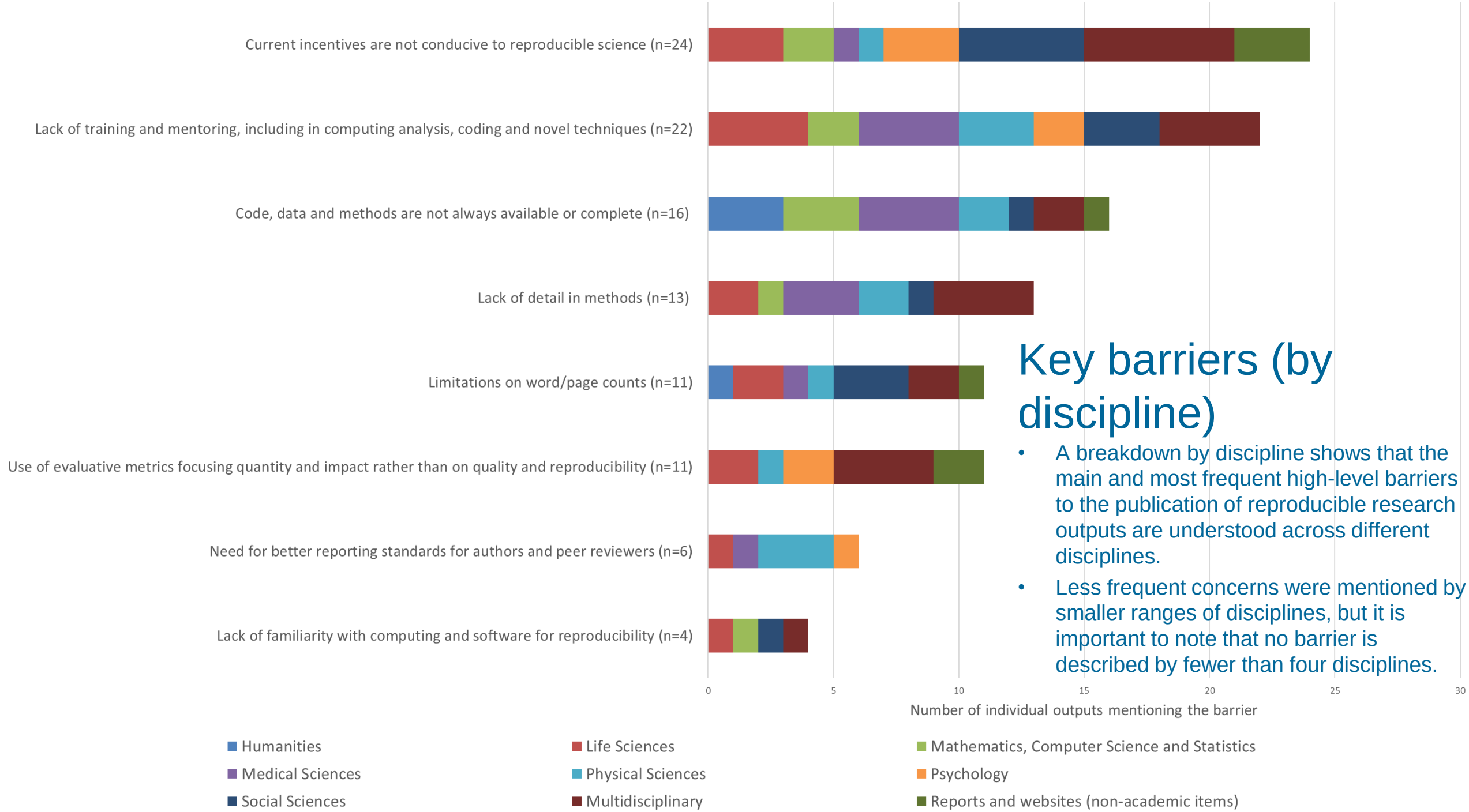
- There is no holistic approach to culture change at the system level when it comes to research reproducibility.
- There is a shared understanding that the key barriers to research reproducibility cannot be solved by a single player in higher education and research, and that there is a need for collective action. Key finding
- The “publish or perish” culture encourages quantity over quality in the publication of research findings.

Study/Project level

- There are some cases where publishing reproducible research outputs will likely remain difficult in practice, e.g. in cases where:
  - authors and/or their organisations may not be able or willing to share sensitive or commercial research data;
  - a study is based on specific media that cannot be effectively captured and packaged in a format suitable for sharing alongside an article (e.g. tactile and/or olfactory data obtained in the field).

Disciplinary differences

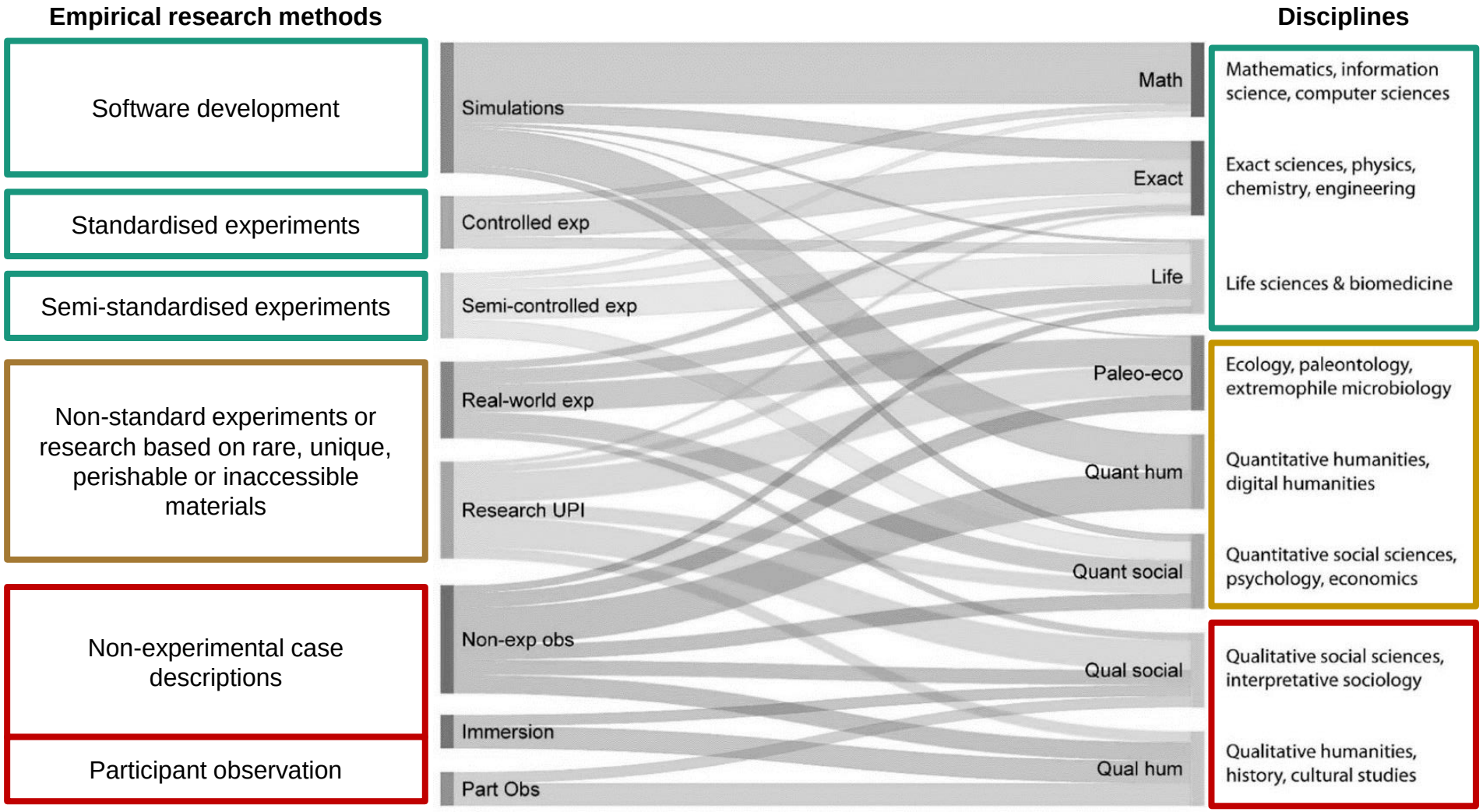
Section 6



Key barriers (by discipline)

- A breakdown by discipline shows that the main and most frequent high-level barriers to the publication of reproducible research outputs are understood across different disciplines.
- Less frequent concerns were mentioned by smaller ranges of disciplines, but it is important to note that no barrier is described by fewer than four disciplines.

Research methods affect the publication of reproducible research outputs



- Based on a researcher's discipline and the typical research methods used, publishing reproducible research outputs may present more or less significant challenges. Key finding
- On average, a **data table** (typical of mainly quantitative disciplines) is easier to curate and share compared to **participant observations** or **thematic coding** (typical of mainly qualitative disciplines, where personal data protection and ethical considerations may play a role, too).

Role of research funding organisations and infrastructures

Section 7

Knowledge Exchange

The role of research funding organisations

- Some research funding organisations across the globe have developed statements or guidance with regard to research reproducibility, and this is a positive step towards increased publication of reproducible research outputs.
  - Positions vary widely, and we have identified four main categories as exemplified on this slide. Key finding
- A majority of research funding organisations are yet to implement reproducibility policies, which means that the landscape is fragmented.

Requirement in grant applications

"By having a statement on data and research materials, NERC is looking to ensure that the research it funds is transparent and ***reproducible****, to allow others to confirm or challenge the research."* NERC (UK)

Project or initiative

"HEC Paris, University of Orléans, and the French National Center for Scientific Research (CNRS) are launching cascad, Certification Agency for Scientific Code and Data, the world's first public laboratory specialized in the certification of the ***reproducibility*** *of scientific research."* CNRS (France)

Research funding

"With the pilot programme Replication Studies, NWO wants to encourage researchers to carry out replication research. […] There are three types of replication research: ***Reproduction -*** *replication* ***with existing data: repeated analysis of the datasets from the original study****. [...]."* NWO (Netherlands)

High-level commitment

"[…], open availability of research results, free of charge for the user, must be ensured, improving by the same means transparency, ***reproducibility****, visibility and democratisation of research."* CHIST-ERA (network of funding organisations in Europe)

Underlying themes in the reproducibility discourse

- A review of definitions in the literature leads to an important realisation: the publication of reproducible research outputs is closely related with other, higher-level, themes, which have historically been discussed by a wide range of stakeholders in higher education and research. For example, research funding organisations often cover aspects related to reproducibility as part of broader discussions. Key finding
- In addition, open scholarship practices are seen as creating an enabling environment for reproducible research, as they provide tools and principles that complement specific solutions around reproducibility.
- The following extracts reflect the above and exemplify shared views in the reproducibility discourse:
  - "**Transparency**, **openness**, verification and reproducibility are important features of research and innovation. Open research helps to support and uphold these features across the whole lifecycle of research – improving public value, research **integrity**, re-use and innovation." (UKRI, n.d.)
  - "In this sense, work to increase reproducibility conveys the idea that methodological **rigour** and **transparency** do (and must) go hand in hand with greater **openness**, with research assessment and with career progression." (European Commission, 2020)

Rigour

Integrity

Transparency

Openness

Tools and infrastructures (technical and social)

- A wide range of tools and infrastructures supporting research reproducibility are already available in today's landscape. Some solutions have a **broad remit** and focus on open scholarship practices such as sharing data and code, while a minority have a **specific mission** around reproducibility. Key finding
- Among the tools and infrastructures with a specific purpose, we highlight:
  - Research compendia (e.g. ReproZip – tool, Whole Tale – infrastructure, RC – Zenodo Community)
  - Executable research articles (e.g. eLife – infrastructure)
  - Badging systems (e.g. cascad – infrastructure)
- Infrastructures may serve research approaches/methodologies asymmetrically, in response to user needs or established disciplinary norms (e.g. the sharing of computer code is more common compared to the sharing of thematic coding).
  - Our literature review has not identified the publication of reproducible research outputs as a priority in disciplines that are based on non-digital research items. Key finding

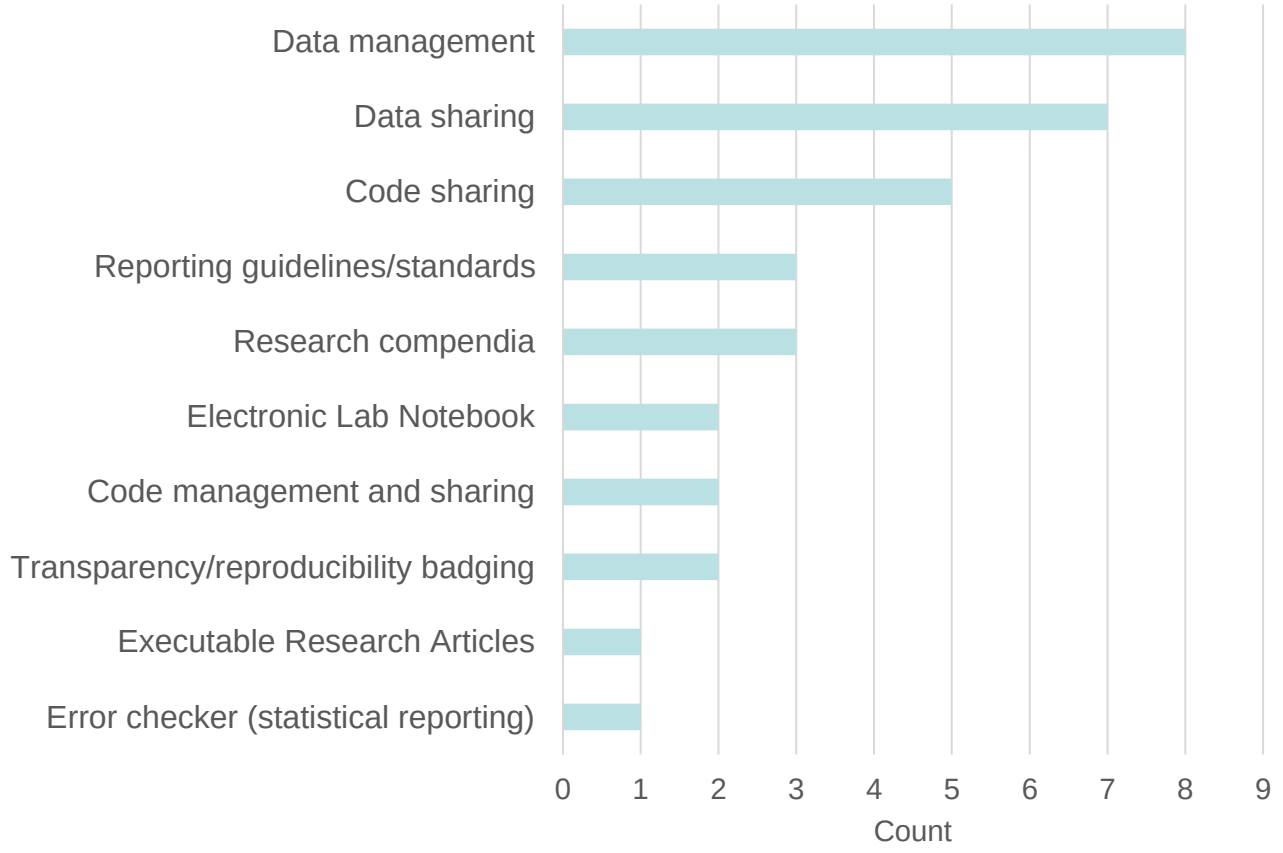
General purpose (e.g. open scholarship, data sharing)

zenodo, EUROPEAN OPEN SCIENCE CLOUD, jupyter, GitLab

Specific purpose

ReproZip The Reproducibility Packer!, WHOLE TALE, cascad, eLife, RC

Tools and infrastructures (technical and social)



- In our literature review, we have identified 34 examples of infrastructures that are relevant to the publication of reproducible research outputs.
  - In this slide, we have taken a broad view of what may be considered as an infrastructure, to consider both technical and social infrastructures. For example, under reporting guidelines and standards we have included both TOP guidelines (an initiative) and the EQUATOR network (mainly a database).
- Although the list is likely not to be comprehensive, it shows that infrastructures can play a role at any stage during the research reproducibility lifecycle, from discovery (e.g. data sharing) to dissemination (e.g. research compendia, executable research articles).

Note: This slide presents a summary overview. A full list of infrastructures we identified in the literature review is available in Appendix A.

Addressing barriers

Section 8

Potential solutions to the barriers identified

Developing incentives for rigorous, transparent, and reproducible research (n=15)
Delivering training at all levels of seniority or experience (n=14)
Increasing awareness, not shame and blame (n=11)
Using journal and funder policies on data sharing to lead behavioural change (n=11)
Improving reporting standards to enhance transparency of published research (n=10)
Encouraging data and code sharing (n=8)
Using badges to motivate reproducible and transparent research (n=8)
Building clear guidance and practical instructions for researchers (e.g. guidelines) (n=7)
Reprioritising data curation in the research process (n=5)

- This section discusses the potential solutions identified in the literature, alongside practical and cultural threats to their implementation.
- Our literature review has identified tens of different opportunities, which refer to research reproducibility across all phases in the research lifecycle.
  - In this section, we focus on potential solutions with regard to the publication and dissemination phase, in line with the scope of the project.
  - It should be noted that most opportunities have been highlighted by a minority of documents/authors: this top 10 list seeks to highlight areas where the literature shows an extent of agreement.

Knowledge Exchange

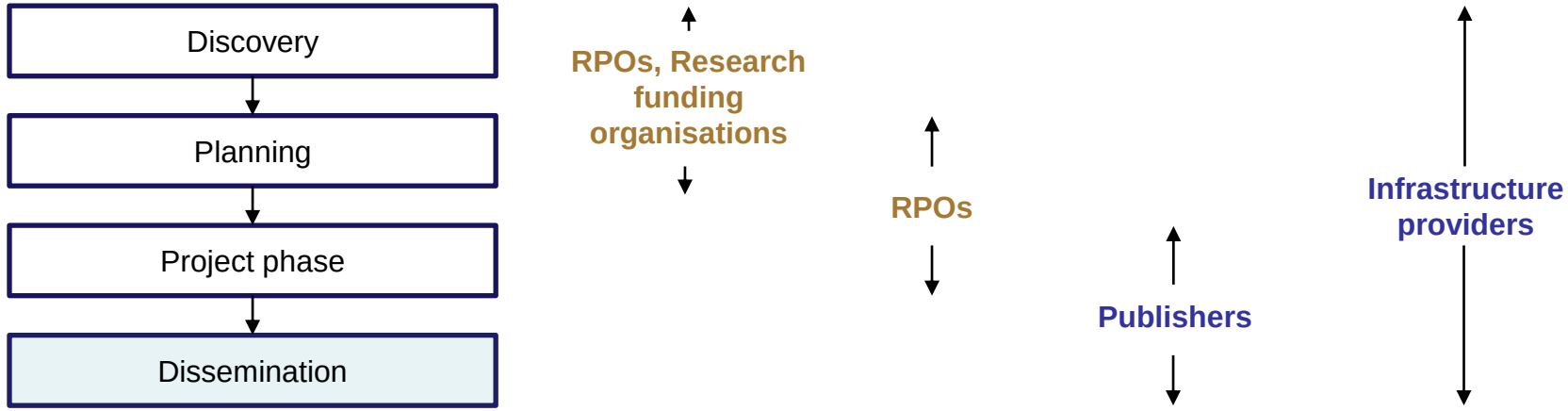
Barriers:
- Rewards and incentives
- Publishing and reporting standards
- Technical and analytical skills

Potential solutions to the barriers identified

Developing incentives for rigorous, transparent, and reproducible research	Delivering training at all levels of seniority or experience	Increasing awareness, not shame and blame
Using journal and funder policies on data sharing to lead behavioural change	Improving reporting standards to enhance transparency of published research	Encouraging data and code sharing
Using badges to motivate reproducible and transparent research	Building clear guidance and practical instructions for researchers (e.g. guidelines)	Reprioritising data curation in the research process

Roles and responsibilities across the research lifecycle

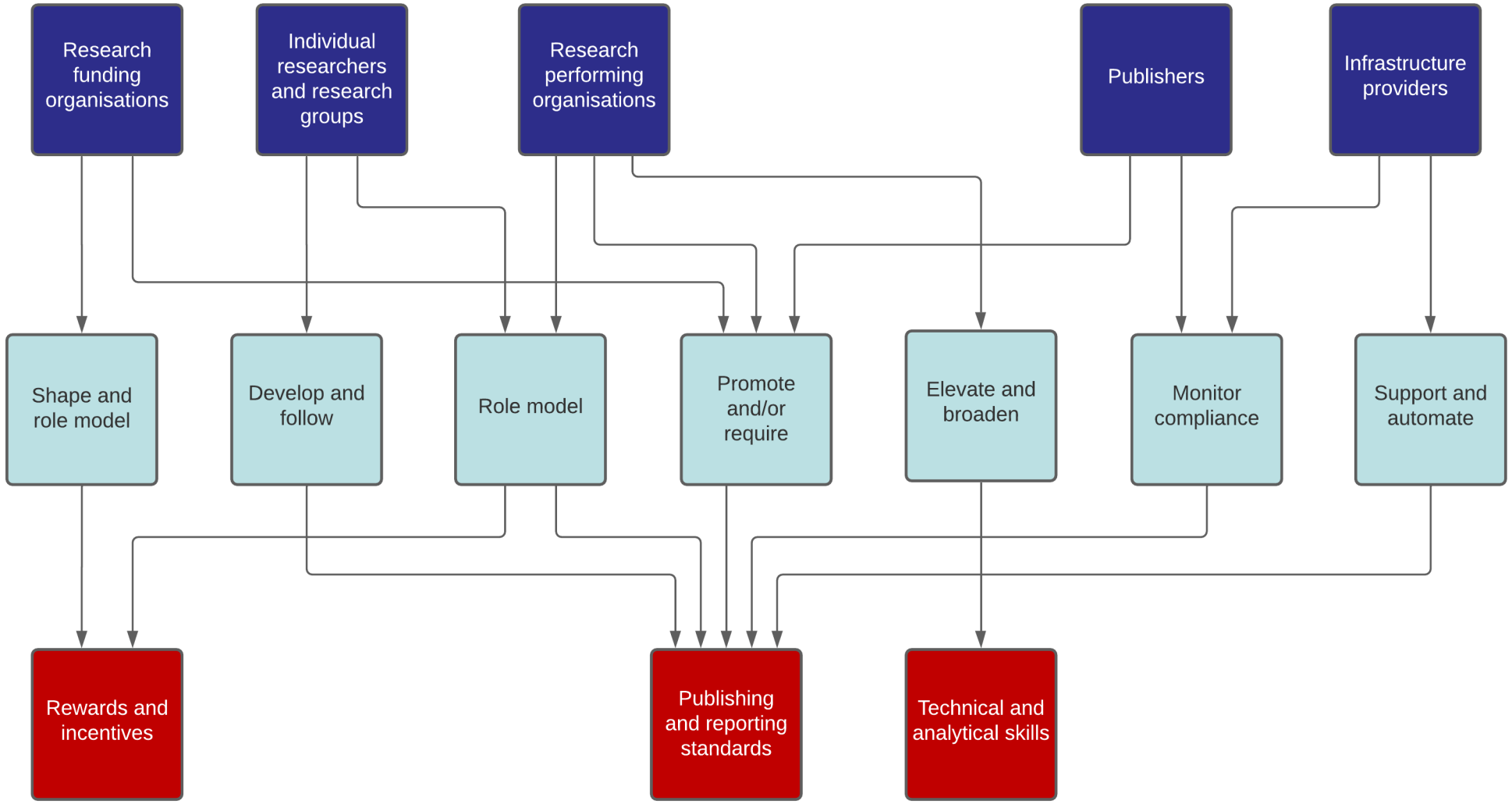
- To address the barriers identified, actors at different levels will have to share responsibilities across a continuum of research phases.
- Although this activity focuses on the dissemination phase, we acknowledge:
  - the importance of **policy mandates** or **institutional requirements** in setting both RPOs and researchers on the right path; and
  - the role of infrastructure providers and publishers in setting **additional requirements and expectations** (e.g. data sharing) across the research lifecycle.



Knowledge Exchange

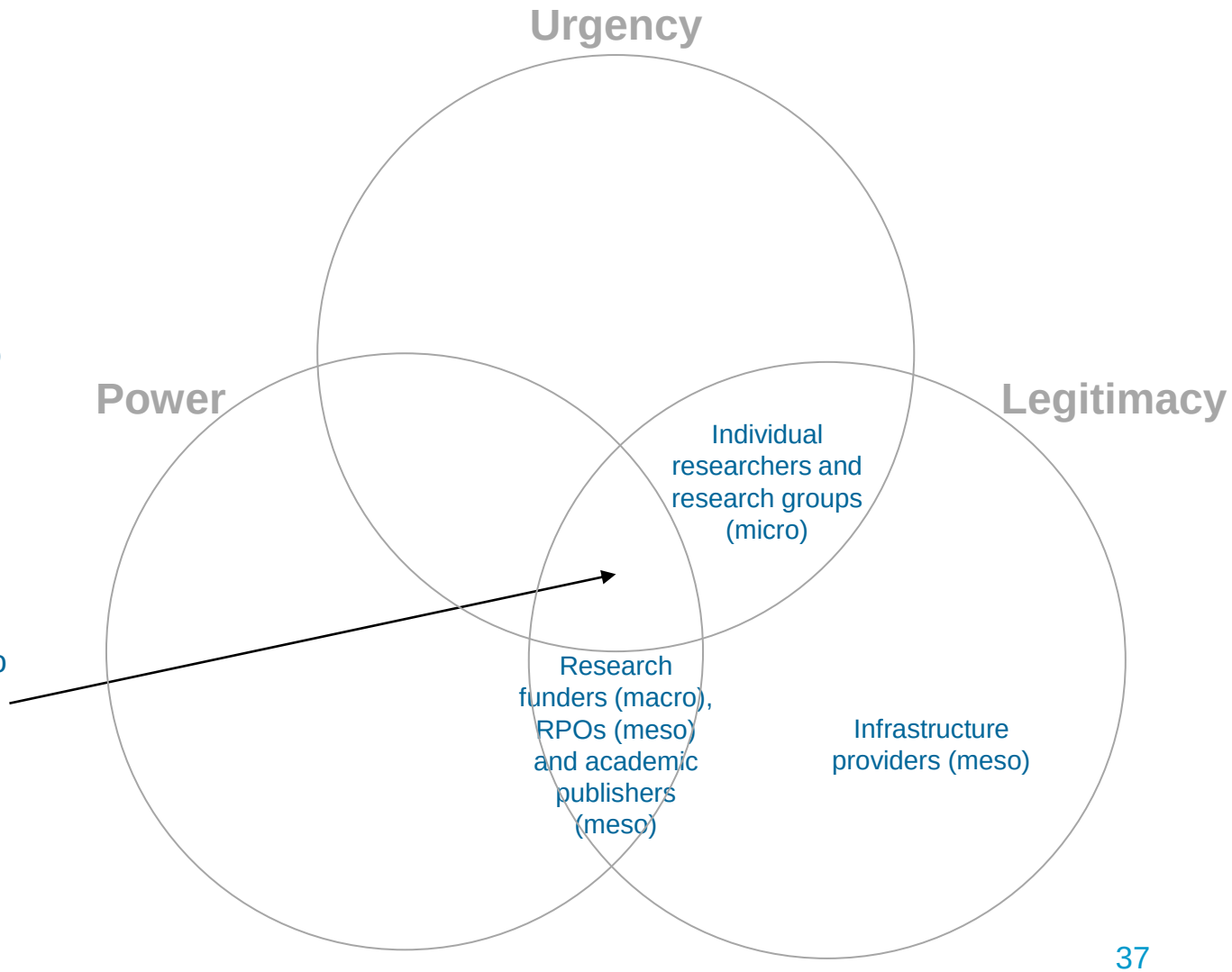
Legend:
- Stakeholder groups
- High-level action(s) required
- Key barriers

Identifying ways forward



Leading change

- The literature recognises research reproducibility as a significant and widespread concern: several stakeholder groups, projects and initiatives are working on this topic. However, beyond the claims made by individual researchers and research groups, other stakeholders do not convey a significant sense of urgency. We acknowledge that some stakeholders (e.g. research funding organisations) have made urgent claims with regard to broader topics such as open science or research integrity, which in some cases also include reproducibility.
- Therefore, we would argue that there are no "**definitive stakeholders**" when it comes to publishing reproducible research outputs: this makes change a more complex and politicised multi-stakeholder process.



Open questions for further research

Section 9

Reproducibility in practice

Our literature review highlighted a range of questions around the practical implications of publishing reproducible research outputs.

Who will carry out reproducibility checks?

- For example, checks for reproducibility may be carried out by a professional such as a data scientist or via the funding and development of automatic systems of compliance for reproducibility before publication.

Who will monitor compliance?

- At the level of individual articles/findings, monitoring and compliance with reproducibility practices and guidelines (where any exist) are often achieved via ex-post studies.
- At the system level, should reproducibility become part of more funder and publisher policies, there will also be a need for monitoring and compliance checks. Key finding

Who will foot the bill?

- The economic cost of monitoring and compliance efforts will need to be mitigated, potentially by acknowledging that there is a link between the ex-ante (helping before the submission of a grant proposal) and the ex-post (checking published evidence). Key finding

What happens when all research is reproducible?

- The research we reviewed tends to comment on the publication of reproducible research outputs in the short and medium terms. This leaves an important question for future consideration: what is the end point of this process of culture change? In particular, what is going to change in a hypothetical long-term scenario where all, or most, research is published in a reproducible way?

Summary of research questions for further investigation

Understanding the problem

- Are the barriers in our mind map (slide 35) a realistic representation of the barriers perceived by individuals and organisations in the research reproducibility landscape?

Stakeholders and collaboration

- To what extent do stakeholder groups feel that collaboration is needed to deliver on the responsibilities in our mind map?
- What specific inputs and/or support does each stakeholder group need from others?
- What are the key roles and responsibilities in terms of reproducibility checks and compliance monitoring?

Practical steps

- To what extent do stakeholders in our mind map have the means to deliver on the actions attributed to them?
- What factors are hindering progress?
- To what extent are today's technical infrastructure and tools suitable for the publication of reproducible research outputs? What improvements or new infrastructures may be needed going forward?
- To what extent are the barriers and issues we identified attributable to (sub)disciplines as opposed to the overall research culture?
- How can the publication of reproducible research outputs be funded in a sustainable way, recognising a range of roles and responsibilities across different players?

The future of research reproducibility

- What happens if all research is eventually published in a reproducible way? In other words, are we going through a transition stage that will lead to the reproducibility discourse being dropped once research practices improve across the board?

Summary

Section 10

Summary of key findings

Political	Economic	Social	Technological

It is difficult to find coherent terminology that speaks consistently to all stakeholder groups involved.

The barriers we identified are multi-stakeholder in nature, even if specific issues may be attributed to select stakeholder groups. Addressing them will require collective action.

A minority of research funding organisations address reproducibility via different models. Many are yet to address research reproducibility explicitly, but do refer to rigour, integrity, transparency and openness in their policies.

Reproducibility is not (nor should it be) the ultimate aim of scientific endeavours.

When addressing reproducibility at the publication and dissemination stage, implications across the entire research life cycle should also be considered.

The economic cost of monitoring and compliance efforts will need to be mitigated, potentially by providing support before submitting a grant proposal or by checking published evidence. Specialist staff may be needed for this.

Based on a researcher's discipline and the typical research methods used, publishing reproducible research outputs may present more or less significant challenges. As (sub)disciplines have very specific requirements, tiered interventions and solutions at the macro, meso and micro levels are likely to be beneficial.

Our literature review has not identified the publication of reproducible research outputs as a priority in disciplines that are based on non-digital research items.

Open access and open data do not guarantee reproducibility by default. Some additional activities are key, such as the sharing of code and detailed methods or procedures.

The publication of reproducible research outputs is closely related with other themes that have historically been discussed by a wide range of stakeholders in higher education and research: rigour, integrity, transparency and openness.

A range of infrastructures supporting open scholarship practices are broadly available. A small range of infrastructures with a specific mission around reproducibility are also in place.

Report commissioned by *Knowledge Exchange*
https://www.knowledge-exchange.info/

Report authors:
Andrea Chiarelli, Lucia Loffreda, Rob Johnson
www.research-consulting.com
Contact:
andrea.chiarelli@research-consulting.com

Contributors:
Noémie Aubert Bonn, Hasselt University (Qualitative coding)
Laura Fortunato, University of Oxford (Review and quality assurance)

Find out more about the project:

https://zenodo.org/communities/ke-prro

https://doi.org/10.5281/zenodo.4647697

Knowledge Exchange

Appendix A – Typology of infrastructures

Name	Typology
Git	Code management and sharing
Mercurial	Code management and sharing
Code Ocean	Code sharing
Gigantum	Code sharing
GitHub	Code sharing
GitLab	Code sharing
Project Jupyter	Code sharing
European Open Science Cloud	Data management
Inquisite	Data management
Open Science Framework	Data management
COINS	Data management (disciplinary)
INFRAFRONTIER	Data management (disciplinary)
Psychological Science Accelerator	Data management (disciplinary)
Resource Identification Initiative (FORCE11)	Data management (disciplinary)
FAIRsharing	Data management + Reporting guidelines/standards
CERN Open Data portal	Data sharing
Figshare	Data sharing
Registry of Research Data Repositories	Data sharing
Zenodo	Data sharing
IRIS repository	Data sharing (disciplinary)
ISPS Data Archive	Data sharing (disciplinary)
LONI image data archive	Data sharing (disciplinary)

Name	Typology
LabArchives	Electronic Lab Notebook
LabFolder	Electronic Lab Notebook
Statcheck	Error checker (statistical reporting)
eLife	Executable Research Articles (ERA)
Equator Network	Reporting guidelines/standards
TOP guidelines	Reporting guidelines/standards
ReproZip	Reproducibility packaging/bundling solution (research compendia)
Whole Tale	Reproducibility packaging/bundling solution (research compendia)
Zenodo RC Community	Reproducibility packaging/bundling solution (research compendia)
CASCaD Certification	Transparency/reproducibility badging
Curate Science	Transparency/reproducibility badging