

# Written Evidence Submitted by The Alan Turing Institute

(RRE0088)

<b>Turing Response to Reproducibility &amp; Research Integrity Inquiry</b>	1
Executive summary	2
Submitting organisation	4
Production of the submission	5
Rationale for submission	5
The issues in academia that have led to the reproducibility crisis	6
Stakeholder roles in addressing the reproducibility crisis	8
Research funders	8
Research Institutions	9
Individual Researchers	10
Publishers	11
Governments	12
Policies to deliver a positive impact on academia’s approach to reproducible research	13
Align research incentives across equity and inclusion, responsible research and innovation, open research, and reproducibility	13
Embrace the diversity of needs across research domains & encourage skill-sharing across them	14
Carefully consider which research metrics align with incentivising reproducible and reusable research	14
Invest in open infrastructure and make it cost effective to use	15
Examples of reproducible research at The Alan Turing Institute	16
The Turing Way	16
Turing Data Stories	16
Digital Twin of the first 3D-printed bridge	17
Environmental AI Project	17
Acknowledgements	17

## Executive summary

The Alan Turing Institute is the UK's national institute for data science and artificial intelligence.

The [Tools, Practices and Systems \(TPS\) programme](#) at the Turing represents a cross-cutting set of initiatives which seek to build open source infrastructure that is accessible to all, and to empower a global, decentralised network of people who connect data with domain experts.

The TPS programme is submitting the following as part of the Science and Technology Committee's inquiry into reproducibility and research integrity:

The issues in academia that have led to the reproducibility crisis

- Academic incentives are misaligned with reproducibility and research integrity.
- Novelty prioritised over rigour for research output.
- Open research mandates are insufficient, we need reusability mandates.
- Career progression criteria does not account for reproducible work.
- Fragmentation of initiatives promoting responsible research and innovation, open research, and equality, diversity and inclusion.

Roles of different stakeholders

- As central stakeholders with the strongest position to change research culture in support of reproducibility, **funders** should re-evaluate and diversify their assessment criteria.
- **Research institutions** hiring practices could be better focused on the needs surrounding working reproducibly, including time to replicate findings and training skills.
- Although previous open research policies have focused on **individual researchers'** work and skills, it is time to consider the systemic changes that are needed for reproducibility.
- Policies should reward **publishers** who invest in infrastructure that supports reproducible outputs.
- The TPS programme sees the “reproducibility crisis” as a significant opportunity for the **UK government** to be a world-leader in supporting reproducibility and research integrity and tackling a systemic issue in academic research.

Policy suggestions to improve reproducibility and research integrity

- Align research incentives across equity and inclusion, responsible research and innovation, open research, and reproducibility.
- Embrace the diversity of needs across research domains and encourage skill-sharing across them.
- Carefully consider which research metrics align with incentivising reproducible and reusable research.
- Invest in open infrastructure and make it cost effective to use.

We include at the end of the submission evidence of reproducible research conducted at The Alan Turing Institute and a list of workshop participants who contributed to the evidence captured in this report.

## Submitting organisation

The [Alan Turing Institute](#) is the UK's national institute for data science and artificial intelligence.

At the Turing we solve real world problems by bringing together computer scientists, engineers, statisticians, mathematicians, ethicists, scientists and colleagues from the humanities. This multi-disciplinary approach allows us to innovate and develop world class research in AI and data science, generating the creation of new businesses, services and jobs. Of equal importance is Turing's work to lead the public conversation around the societal impact of AI and data science, and to ensure that research and development is carried out within an ethical framework.

The [Tools, Practices and Systems \(TPS\) programme](#) at the Turing represents a cross-cutting set of initiatives which seek to build open source infrastructure that is accessible to all, and to empower a global, decentralised network of people who connect data with domain experts.

The programme aims to:

- Build trustworthy systems
- Embed transparent reporting practices
- Promote inclusive interoperable design
- Maintain ethical integrity
- Encourage respectful co-creation

The definitions terms used in this document with regards to reproducibility and replicability are based on work by Claerbout ([Rougier et al., 2017](#)).

Reproducibility, according to Claerbout ([Rougier et al., 2017](#)) is defined as meaning “running the same software on the same input data and obtaining the same results”. Replicability on the other hand is defined as “writing and then running [new analysis] based on the description of a computational model or method provided in the original publication and obtaining results that are similar enough ...” (ibid). More detailed explanations of the definitions can be found here: [Definitions — The Turing Way \(the-turing-way.netlify.app\)](#).

We take the definition of open science from [the Organisation for Economic Co-operation and Development \(OECD\)](#), as the practice of making “the primary outputs of publicly funded research results – publications and the research data – publicly accessible in a digital format with no or minimal restriction.” However, we refer throughout this document to open research as a more inclusive term than open science given that quantitative research can be found across the arts and social sciences, as well as engineering and technology applications.

## Production of the submission

This submission to the Parliamentary Inquiry was developed via an open call workshop with different stakeholders from the Turing community. Participants were invited to reflect on their own views to the inquiry areas, to ensure diverse perspectives on this complex topic were reflected in our response and suggestions. Workshop participants are acknowledged at the end of the document.

For this response we will focus on three of the five topics outlined in the call for evidence. Specifically:

2. The issues in academia that have led to the reproducibility crisis, i.e., its causes
3. The role of the following in addressing the reproducibility crisis:
  - research funders, including public funding bodies;
  - research institutions and groups;
  - individual researchers;
  - publishers; and
  - Governments and the need for a unilateral response / action.
4. What policies or schemes could have a positive impact on academia's approach to reproducible research;

## Rationale for submission

The Turing are committed to supporting high quality, reproducible research in data science and AI, and the Tools, Practices and Systems programme supports this drive. As a charitable organisation, one of our objectives is that the work we support not only has a significant impact on fundamental research, but can also be applied in the real world to the biggest challenges facing the globe this century. To achieve real-world impact, outputs and findings need to be reproducible and replicable, to allow them to be applied and built on with confidence by policy makers, engineers, businesses, and others. Support for this critical aspect of research at a national level is welcomed.

## The issues in academia that have led to the reproducibility crisis

The TPS programme are glad that the challenges of publishing reproducible research are being explored by the Science and Technology Committee. There is a lot of work to do to improve the reusability and reliability of published research. However, we consider the phrase “reproducibility crisis” to be overly emotive. Challenges in publishing reproducible research have existed for many decades and impact every subject domain and international research community.

### **Academic incentives are misaligned with reproducibility and research integrity**

Issues in reproducible research predominantly stem from academic incentives that encourage competition between research teams working on similar questions. The system means that an individual research team - behaving rationally – is likely not to share their data, code, protocols nor experimental design expertise with researchers working in the same area as themselves. The outcomes will likely result in siloed knowledge and a lack of transparency in research methods.

### **Novelty prioritised over rigour for research output**

The current incentive structure also prioritises novelty above rigour. Research applications are usually assessed on a narrow definition of “quality”. UKRI has made “[modernising their] research evaluation policy to reward openness and diversity in research” a priority for 2020-2021 as part of their Corporate Plan ([UKRI-091020-CorporatePlan2020-21.pdf](https://www.ukri.org/corporate-plan/2020-21)). Language focusing on novelty or how a research grant is distinct from the others being submitted frequently appears within the assessment criteria for UKRI funding calls.

The TPS programme recognise the merit of exploring novel and original research directions in principle, however, we argue that balancing this with collaborative incentives would lead to better outcomes for the advancement of science. For example, sharing when an experiment did not work, or making an analysis technique easy to reuse, would speed up others’ efforts towards the same goal. Unfortunately, these activities tend not to be novel, nor are they necessarily driven by “ambition, adventure and transformation”. Rather, the activities that lead to transparent, reproducible and reusable outputs should centre around rigour and support for others’ research endeavours.

### **Open research mandates are insufficient, we need reusability mandates**

In the last ten years there has been a big push to incorporate “open science” practices into research funding and publication assessments. Specifically, the inclusion of data management plans have been required for all UKRI grants since 2018 (examples of data management plan requirements: <https://www.data.cam.ac.uk/funders> ) and publications have asked for a statement on data access - and in some cases research code access - since around 2014 (<https://journals.plos.org/plosone/s/data-availability>).

Unfortunately, making materials available does not necessarily mean there are suitable processes and checks to ensure the work is reproducible. There are significant barriers to running a reproducibility check. Sometimes data is sensitive and cannot be shared openly (applying for access to health data can take between six to and twelve months), and sometimes the algorithms that underpin findings require substantial computational time and cost.

Finally, many peer reviewers do not check the reproducibility of published papers. This is a resource (time, money and expertise) intensive activity that most peer reviewers cannot justify given that their work is not reimbursed.

### **Career progression criteria does not account for reproducible work**

There are currently no clear incentives for early career researchers to reproduce papers that are already published. If they get the same result, they will likely have little to show for this effort. If they get a different result, there are few pathways to communicate this finding. In all likelihood, these researchers would not start the work at all, as it would not likely meet the current standards of “research excellence”. Not being able to get the same findings as those published in existing papers does not - on the surface - deliver novel insights, nor actionable ones. Although for science as a whole would be in a stronger position if we learned how well findings can be reproduced (same method, same data) or replicated (same method, different data), our current career progression is not based on those needs.

### **Fragmentation of initiatives promoting responsible research and innovation, open research, and equality, diversity and inclusion.**

Researchers interact with institutional and funder policies for responsible research and innovation through their research ethics committees at the start of their projects. They interact with open research policies through requirements around data management plans or code availability statements around publication of their results. They will also engage with equality, diversity and inclusion initiatives to build a more supportive, welcoming and sustainable culture in academic research.

All of these initiatives are critical parts of improving research quality and reproducibility and are aligned with delivering high quality research. However, their development and implementation are sometimes at odds with each other, and - at the very least - siloed and disconnected. Connecting these efforts will align incentives and lead to more coherent expectations across the research ecosystem. This alignment will bring together three groups of people who may not yet see their role in improving research reproducibility. Those who have traditionally been excluded from conversations around open research ([Whitaker & Guest, 2020](#)), those who see ethical considerations as a bureaucratic barrier to innovation ([Wittrock et al., 2020](#)), and research leaders who do not see how efforts to be anti-racist can contribute to high quality science ([Forrester, 2020](#)).

The TPS programme’s suggestions in this submission of evidence are that the root causes of a lack of reproducibility are interwoven with the concerns regarding ethical management of

research projects, and inclusion of a diverse group of stakeholders, including the researchers themselves.

## Stakeholder roles in addressing the reproducibility crisis

In this section we explore the roles that research funders, research institutions, individual researchers, publishers, and government organisations can take to address the challenges associated with low levels of reproducibility in our current research outputs.

### Research funders

**As central stakeholders with the strongest position to change research culture in support of reproducibility, funders should consider re-evaluating and diversifying their assessment criteria.**

From the TPS programme's perspective, research funders appear to have the greatest influence on improving reproducible research practices because they hold the purse strings. Funders can ultimately choose which projects are considered high enough priority to fund. They do not act in isolation. Individual researchers contribute to these decisions as peer reviewers and applicants. However, these same researchers are hired by research institutes that should be considering bringing in grant funding to sustain their organisations. Publishers have some power through citation metrics as a proxy for research quality, but career progression is primarily driven by grant and fellowship funding to initiate and continue research projects.

The TPS programme suggests that funders reassess their definition of "impact" and their reliance on novelty as a measure of research quality, focusing instead on the collaborative potential or support for reproducibility. In order to incentivise reproducibility we need applications to be assessed for how well the project outcomes will help others to conduct innovative work.

An area that is severely underfunded is an investment in software projects that are widely used and that make research more efficient, robust and reproducible. Most of these projects need support primarily around [sustainability and maintenance](#). The novelty and impact come from their users who can apply the tools to a plethora of research questions. The USA based Chan Zuckerberg Initiative's programme for [Essential Open Source Software for Science](#), launched in 2019, is one of the few funding opportunities to support these projects, although the US's National Aeronautics and Space Administration (NASA) announced [funding for eight open source libraries](#) in September 2021.

Pivoting grant application assessment to focus more on how research teams have supported others breaks down a competitive framework and will likely better align with the purpose of open research. It also starts to realign the counterproductive incentives that lead to rational researchers not making it easy for others to build on their research outputs.

Alternatively, or additionally, funders could put out calls specifically for research projects that reuse materials that are already available. These projects could focus on reproducing the

original finding as a first step, publishing the output for others to learn from as a key deliverable of the grant, and then extending the work as the new research team prefers. The key requirement would be that no funding is used to collect data, nor to write new code in the first instance. This is a particularly strong possibility for pre- and post-doctoral fellowships, or fellowships with a significant training component. This could be a useful mechanism to ensure that their training follows best practices for open research and embeds rigor and reproducibility in their work.

The TPS programme suggest that funders provide guidance on how research is communicated at the end of the project. For example, the funder mandates have been a major driving factor in the adoption of open access publication pathways. The TPS programme would like to see requirements - that can be, and are, enforced - for not just the publication of data and code, but integrated testing and independent verification that the results are reproducible. This effort would be budgeted into the grant, and the results made publicly available, even if the data and/or code are too sensitive to be made public. Building on this, potential funding could be used to engage research teams that would traditionally be in competition with the original researchers to specifically incentivise collaboration within a research topic.

Another suggestion is that funders provide funding opportunities that support longer contracts for early career researchers and professional research support staff. We discuss the pressures of precarity on early career researchers in the section on research institutions below. Supporting permanent members of staff as librarians, data stewards or research software engineers, means each organisation has substantial in-house expertise to support reproducible, replicable and reusable research.

## Research Institutions

### **Research institutions hiring practices could be better focused on the needs surrounding working reproducibly, including training expertise and time to replicate findings.**

An important action that research institutions could take is to embed reproducibility into their hiring, promotion and training requirements. For hiring and promotion, The TPS programme suggest that hiring and promotion committees specifically consider how reproducible an applicant's work has been. This may be through a specific case study, or letters verifying the research can be reproduced - and potentially replicated - by independent research teams. We hope that committees are increasingly considering their candidates' open research practices - how much code or data is available for reuse, whether there is sufficient documentation associated with these artefacts - and their commitment to improving diversity, equality and inclusion. The TPS programme also suggest that they also incorporate an assessment criterion of how well their work has supported other groups' work, as described in our suggestions to funders.

The TPS programme suggest metrics of reproducibility, replicability and reusability should be carefully applied ([definitions](#)). *Reproducing* findings - defined as getting the same result with the same data and the same code - is manageable albeit rarely done. But whether or not



someone's results *replicate* - that is, whether it is possible to find the same qualitative result on an independent dataset - is not necessarily a good judgement on their individual contribution to research. In order to improve the reproducibility, replicability and rigour of research projects research teams increasingly need to share their findings regardless of whether a replication is "successful" or not. Assessment of research quality should recognise that any work conducted to a high standard represents a successful contribution to science as it improves our understanding of the conditions that can *affect* an *effect*.

As we have highlighted above, the current incentives in academia are likely a major contribution to why early career researchers likely do not choose to make their work reproducible. Another factor, that potentially compounds this misalignment of incentives, is that there are often significant technical and information governance challenges with making research reproducible. The TPS programme's suggestion - as outlined in our handbook for reproducible research [The Turing Way](#) - is that it is easier to work reproducibly from the start. That means incorporating training in version control, software testing, reproducible computational environments and research data management from undergraduate courses and through postgraduate training. Hiring faculty who can teach these skills means adjusting the standard criteria against which their applications are traditionally assessed to focus on their own skills in reproducible research and open research practices. The TPS programme also suggest welcoming professional staff from the library sciences and software engineering groups and reusing materials rather than re-creating new materials each year. There doesn't appear to be logical to create proprietary materials to teach students the importance of not reinventing the wheel in their research.

Early career researchers often have particularly precarious career pathways. They are often incentivised to "publish or perish" and the quality of those publications usually cannot be effectively assessed on a short-term basis. They may be cited well by their peers, but the citation process will likely take a minimum of three- to five-years to start to show an effect: a duration much longer than most early career researcher employment contracts. The TPS programme's final suggestion for research institutes is that they offer longer term contracts. A three-year minimum contract for a postdoctoral researcher gives them some time to check the reproducibility of their work, and hopefully to reuse their own code and data for additional research questions. Ultimately, we would likely see more reproducible and replicable research if universities gave longer contracts to their research staff, invested in their skills and the long-term sustainability of their work, and supported their career development in-house rather than requiring them to move institution, city and project every few years.

## Individual Researchers

**Although previous open research policies have focused on individual researchers' work and skills, it is time to consider the systemic changes that are needed for reproducibility.**

Much of the open research training and recommendations from the last ten years have focused on the practices of individual researchers. Although the TPS programme think that these are the practices that will ultimately lead to more reliable, reproducible, and replicable research, we consider the causes of the "reproducibility crisis" to be more systemic and principally flow from

incentives set by funders, employers and publishers. There are not many policies that we would suggest be targeted specifically at individual researchers. Rather, it would be more productive to see more training programmes and rewards for those who help others in a collaborative manner through their research activities.

There should be additional protections against individual researchers whose work is neither reproducible nor replicable. As we mentioned above, a successful attempt at replicating a finding (running the same analysis on a new dataset) is a valuable contribution to science whether or not the qualitative finding is the same. On the other hand, if a result does not reproduce - if an independent experimenter gets a different result using the same data and analysis steps - it is likely that there is an error in the analysis. It is imperative that we increasingly create a culture where those errors can be fixed, and the results updated with as little punishment or shame as possible. There is currently a gulf between purposeful fraud and human error. Almost all research papers likely have some human error in them. The TPS programme believe that we do not have a peer review system that is incentivised to find them.

## Publishers

### **Policies should reward publishers who invest in infrastructure that supports reproducible outputs.**

Trying to work reproducibly after the work has already been completed is akin to archaeology. It is orders of magnitude harder than putting in place the version control, documentation and research data management requirements as you are conducting the analyses. The TPS programme do not have many suggestions for publishers, given that they primarily impact the end of a research project. They have little power to generate new incentives, rather they should respond to mandates required by funders, librarians, and individual researchers.

We have evidence of publishers responding to initiatives led by researchers and funders. These include adoption of data and code availability statements and open access publishing requirements. Unfortunately, the vast majority of publishers still allow data to be “available on reasonable request” which is considered too vague to be implemented or tested and does not change the ecosystem’s incentives towards re-using material and monitoring the reproducibility of findings. The reliance of the publishers on unpaid peer reviewers, most of whom receive no training to undertake their assessment of the papers they read, limits how well they can likely implement changes.

On the other hand, publishers represent significant amounts of money in the academic research ecosystem and potentially have a lot of power. In the TPS programme’s workshop to capture this evidence we heard examples of journals not allowing two papers to be published as they used the same datasets. The publisher’s incentive structure is highly aligned with novelty and timeliness: both of which limit investment in reproducible research.

We refer the committee to the Invest in Open Infrastructure (IOI) report on [Furthering the Future of Open Scholarship](#) (2021). Their technical recommendations include “building for increased

modularity and compatibility of common infrastructures and designing shared service and support models to drive resources and staffing support.” They also recommend “establishing an Open Infrastructure Technology Oversight Committee, providing a foundation for sharing best practices, aligning power, and calling for broader system change including negotiating with vendors on pricing, reciprocity, and values-alignment.” The TPS programme agree with the IOI report and celebrate publishers who have invested in processes to assess reproducibility, guidance for peer reviewers, and infrastructure to make publishing more modular, cost effective and interactive. In particular, eLife’s reproducible articles are a great example of how publishing could be updated for the 21st century (<https://elifesciences.org/labs/ad58f08d/introducing-elifes-first-computationally-reproducible-article>). The TPS programme hope that policies will invest in open infrastructure for scholarly communications and incentivise a shared vision that is value-aligned with incentives incorporating responsible research and innovation and equality, diversity and inclusion.

## Governments

**We see the “reproducibility crisis” as a significant opportunity for the UK government to be a world-leader in supporting reproducibility and research integrity and tackling a systemic issue in academic research.**

While the TPS programme would suggest that the UK government take action to adjust the incentives in academia that are the root cause of the “reproducibility crisis”, we recognise that research is an international activity. If UK researchers are to be world leaders across multiple subject domains, we should work with learned societies and global organisations to coordinate expectations around the world. The TPS programme’s suggestion is to avoid a competitive framework and to invest more on a collaborative focus across borders. This is an opportunity for the UK government to demonstrate further international leadership in tackling a systemic problem in research.

Some of the challenges for reproducible research come from a lack of an economy of scale. We agree with the current investment in trusted research environments, such as those by the [Office for National Statistics](#), [NHS](#) and [Health Data Research UK](#). We need interoperable infrastructure that will allow researchers to collaborate with each other across institutions while maintaining the highest standards of information governance and participant confidentiality.

The TPS programme would like to see the UK lead the world in coordinating standards for equality, diversity and inclusion, and embedding those principles in the assessment of research integrity. To date, open research and reproducibility have been considered separately to considerations of ethics and information governance. In fact, in many research ethics committees they are held in opposition to open practices. For example, ethics boards may require that data access is limited and that research outputs are kept under lock and key. These actions protect participant privacy in the short term but cause a lack of transparency and trust in research outputs that can not be externally validated. The TPS programme also suggest that policies set by the UK government incorporate reproducibility, replicability, and reusability as key metrics for ethical research. This does not mean being lax on data protection

responsibilities. Standards need to be developed and promoted so that responsible data governance, which permits independent reproducibility checks, is incorporated into traditional research ethics considerations.

The purpose of responsible research and innovation is wholly aligned with the goal of making research fairer and more equitable. Keeping ethics, open research and EDI principles in separate policies, committees and activities mean they are misunderstood as separate goals. They are not. The purpose of ensuring that research can be reproduced and reused is so that it can deliver impact to change the world while maintaining the highest standards of data protection and consideration of the rights of the data owners and providers. As we have outlined above, treating researchers equitably is on the causal pathway to reducing the effects of our current lack of reproducibility.

## Policies to deliver a positive impact on academia's approach to reproducible research

### Align research incentives across equity and inclusion, responsible research and innovation, open research, and reproducibility

It is unlikely we can expect individual researchers and research teams to change their behaviours unless academic incentives also change. The TPS programme suggest aligning policies across open research, reproducibility, responsible research and innovation, and equity and inclusion. As we have outlined above, these initiatives are often siloed and engage different communities of researchers. However, they are all working towards the same purpose: more transparent, ethical, reliable research that can be used by others to make a positive change in the world.

Policies to align incentives should also consider what work is paid and rewarded. The lack of stable career pathways, and safe opportunities to speak out when it is hard to reproduce or replicate results, leads to the symptoms of the “reproducibility crisis”. The TPS programme suggest specifically incorporating replication studies, or independent reproductions of published findings, into doctoral training programmes. These initiatives can extend to senior undergraduates too. We commend in particular the initiatives led by Dr Kate Button at the University of Bath that “train students in open, team science, which better prepares them for challenges to come.” ([Button, 2018](#)) Through these training programmes, students learn that incorporating reproducibility, replicability and reusability are not “nice to have” additions to a traditional research output. They are equal in their importance to assessment of high-quality hypothesis creation and testing, and informed interpretation relating to theory and the published literature.

The TPS programme also suggest focusing more on rewarding the process of doing science rather than the answers at the end. Well-designed and well-conducted research should be publishable regardless of whether the finding “works” or not. As we discussed, a focus on novelty as a primary measure of research quality often leads to conditions that foster a lack of

reproducibility. The TPS programme suggest assessing the value of the work as it helps others to continue their research over individual research project's impact or uniqueness. Bringing research quality into alignment with equality, diversity and inclusion requirements - for researchers and stakeholders affected by the analyses alike - will help to incentivise better reproducibility and reusability across the research ecosystem.

## Embrace the diversity of needs across research domains and encourage skill-sharing across them

Interventions and incentives will differ across research disciplines. For example, many machine learning and computer science conferences have infrastructure to support peer review of the [reproducibility of the research submissions](#). This represents a cultural step towards more positive reproducibility, but not - necessarily - towards replication nor incorporating a diversity of research methods ([definitions](#) of reproducibility and replication). In particular, datasets that cannot be made publicly available may be some of the most important for cutting edge research questions that can improve the world, but they do not appear in these conference abstracts.

Policies that specifically invest in translating knowledge from one domain to another will need to work against the prevailing incentives that keep researchers within their domains. They should factor in how much harder it is to publish an interdisciplinary project when the reviewers are primarily used to working within their domains, and the fact that career advancement happens for the most part within specific departments that have not - to date - incorporated much interdisciplinarity in their assessment requirements. Data science and AI institutes are a likely counter to this trend, where the need to cut across statistics, computer science and domain specific applications - including ethics, philosophy and the social sciences - is being prioritised in their hiring and retention policies.

## Carefully consider which research metrics align with incentivising reproducible and reusable research

There are currently few research assessment metrics that can accurately capture whether work is reproducible or not. Availability of research code, data, and computational environments are necessary but in many cases cannot be shared if the data is sensitive or the computational environment is particularly complex. The amount of effort required to independently verify if a research output is fully reproducible will vary between research domains and whether or not the team already has the required expertise. Policies that rely on reproducibility and open research badges (or similar) risk preferentially scoring well-resourced groups highly and leaving the teams who have the greatest need behind. These policies should be considered in the context of a shared goal of making research more equitable and inclusive, within the UK and internationally.

For replication studies - where the same analysis is conducted but on a different dataset - there is no metric to know whether the replication was a "success" or not. All contributions are

valuable, and policies should incentivise a shared learning rather than a competitive framework of who is “right” and who is “wrong”.

The TPS programme suggest that all stakeholders sign up to the San Francisco [Declaration on Research Assessment](#) (DORA) and continue to eschew the use of Journal Impact Factors in their assessment of the quality of research. However, DORA contains many more themes than just avoiding journal level research assessments. The TPS programme also suggest additional investment in their recommendation to “capitalise on the opportunities provided by online publication (such as relaxing unnecessary limits on the number of words, figures, and references in articles, and exploring new indicators of significance and impact)”.

Finally, although we are generally wary of an over-reliance on metrics, the TPS programme encourage increasing the assessment of the impact of an open-source research project. For example, tracking the number of downloads and citations of a sample dataset or the number of times that a piece of code has been used by others can be used to describe how well a research output is helping others to complete their analyses.

## Invest in open infrastructure and make it cost effective to use

The hardware and software systems that we use to conduct and communicate research are important. They allow for reproducibly deployed trusted and secure research environments ([Arenas et al, 2019](#)), or for interactive exploration of the reproducibility of published analyses such as [myBinder.org](#). But these systems cannot be meaningfully deployed if we do not support the people of these infrastructure ecosystems. That means giving people the training they need to use the tools effectively, to understand how to adapt reproducibility best practices for their research projects, and to invest in systemic changes to influence long term sustainability of the ecosystem.

The TPS programme would suggest investment not just in hardware and software (such as the recent UKRI call for [Software for Research Communities](#)) but also in communities such as [The Carpentries](#) that create open source training curricula. The TPS programme’s own investments in [The Turing Way](#) as a guide for reproducible, collaborative and ethical data science is another example of open infrastructure that is designed to help others to achieve the highest impact and rigour in their work. Efforts in metadata standards and the FAIR (findable, accessible, interoperable and reusable) principles ([Wilkinson, et al., 2016](#)) are collective actions that have already improved the reusability of research data. We should consider continued investments in information governance processes for secure data analysis and research ethics review processes to connect disparate infrastructure parts into a coherent and reusable whole.

“Giving away” infrastructure by making it open source may seem like an inefficient use of funding. But there is no evidence for this position as a medium- or long-term investment. The TPS programme suggest the Invest in Open Infrastructure framework as defined in their report on the [“Costs and benefits of collective investment”](#) as a starting point for ensuring the funding from the UK tax payers is invested in an ecosystem that is modular, reusable and interoperable. It is designed to incentivise reproducibility and high-quality science from the ground up.

# Examples of reproducible research at The Alan Turing Institute

## The Turing Way

GitHub: <https://github.com/alan-turing-institute/the-turing-way>

The Turing Way is an open source and community-led project that aims to make data science comprehensible and useful for everyone. This project brings together its diverse community of researchers, educators, learners, administrators and stakeholders from within The Alan Turing Institute and beyond. The Turing Way provides a web-based book with five guides on reproducible research, project design, communication, collaboration and ethical research, as well as a community handbook.

The project is available online on GitHub (<https://github.com/alan-turing-institute/the-turing-way>) under the CC-BY license owned by The Turing Way community. The project uses Jupyter Book (<https://github.com/jupyter/jupyter-book>) to host its content online at <https://the-turing-way.netlify.com>. To date, the project has over 45 chapters across the five guides and 280 people have contributed to the project.

The Turing Way has been highly influential in promoting best practices around reproducibility, collaboration and research integrity, and is regularly cited in policy documents by government and third-sector organisations such as the [London Assembly](#) and the [UK Environmental Observation Framework](#).

## Turing Data Stories

GitHub: <https://github.com/alan-turing-institute/TuringDataStories>

“A Turing Data Story is an interactive mix of narrative, code, and visuals that derives insight from real world open data. They are written as pedagogic Jupyter notebooks that aim to spark curiosity and motivate more people to play with data.”

The peer-reviewed notebook of each Data Story walks readers through the visualisations and narrative, as well as the code and analysis used to produce them. The project is open to contributions from any interested parties and can offer informal opportunities for researchers to develop their data analysis and visualisation skills.

Recent examples of stories include: [When will the United Kingdom be Fully Vaccinated Against COVID-19?](#) and [Modelling Mail-In Votes In the 2020 US Election](#).

## Digital Twin of the first 3D-printed bridge

[Digital twin of the world's first 3D printed stainless steel bridge | The Alan Turing Institute](#)

In a world-first, a 3D-printed steel bridge has been installed in Amsterdam, covered in sensors that enable researchers to build a digital twin of the bridge, incorporating data from the sensors into real time analysis.

To enable these digital twin studies, The Turing is developing a data access platform that will integrate with Autodesk's software to enable researchers to access the large quantities of sensor data stored on cloud computing servers. The Turing is hosting the bridge data for the full two-year period covered by the bridge's current operating permit and has conducted [a thorough ethics review of the project](#) to ensure that the scientific goals of the project do not compromise the public's privacy. Using a custom data platform, the Turing supports researchers who require access to the sensor data stored in its secure cloud. This work demonstrates how even sensitive data can be made available for reuse and replication of work.

## Environmental AI Project

[Welcome — The Environmental AI book \(acocac.github.io\)](#)

Developed as part of ongoing work on environment and sustainability as part of the AI for Science and Government programme at The Turing, Environmental AI is an open-source community project producing guides to data analysis for environmental data scientists. The project produces detailed, step-by-step guides to exploring and visualising publicly accessible datasets relevant for analysing the impact of climate change and other environmental challenges. The project is still in its early stages but seeks to support the replicability and reproducibility of environmental data science research.

## Acknowledgements

The Tools, Practices & Systems programme is grateful to the following participants of the open call workshop for sharing their views on the inquiry topics:

- Malvika Sharan, Tools, Practices and Systems, The Alan Turing Institute
- Georgia Aitkenhead, The Alan Turing Institute
- Aida Mehonic, Tools, Practices and Systems, The Alan Turing Institute
- Ann Gledson, University of Manchester (Research IT)
- Nadia Papamichail, The University of Manchester/The Alan Turing Institute
- Louise Bowler, Research Engineering Group, The Alan Turing Institute
- Martin J. Turner, The University of Manchester (Research IT)
- Kirstie Whitaker, Tools, Practices and Systems, The Alan Turing Institute



- Emma Karoune, Tools, Practices and Systems, The Alan Turing Institute
- Fiona Grimm, The Health Foundation
- Vashti Galpin, School of Informatics, University of Edinburgh
- Arielle Bennett, Tools, Practices and Systems, The Alan Turing Institute

***(October 2021)***