

Reproducibility literature analysis - a federal information professional perspective

Erin Antognoli¹, Regina Avila², Jonathan Sears³, Leighton Christiansen⁴, Jessica Tieman⁵, Jacquelyn Hart⁶

Abstract

This article examines a cross-section of literature and other resources to reveal common reproducibility issues faced by stakeholders regardless of subject area or focus. We identify a variety of issues named as reproducibility barriers, the solutions to such barriers, and reflect on how researchers and information professionals can act to address the ‘reproducibility crisis.’ The finished products of this work include an annotated list of 122 published resources and a primer that identifies and defines key concepts from the resources that contribute to the crisis.

Keywords

reproducibility, reproducibility crisis, replicability, research data, landscape analysis, culture shift

Introduction

Over the last number of years, the terms ‘reproducibility’ and ‘replicability’ have left the realm of science and become widely discussed in mainstream media. Books, blogs, and television news shows have talked about the emergence of a ‘reproducibility crisis’ which brings the validity of scientific research into question. Numerous studies and reports identifying the status of research reproducibility reveal problems at all levels of study across multiple research disciplines. Even published research from prominent journals and institutions suffer from reproducibility issues (Weir, 2015). While some question the idea that the issues surrounding reproducibility constitute a ‘crisis’ (Baker, 2016c), evidence points to a widespread difficulty to reproduce published scientific results.

As data managers and information professionals in U.S. federal libraries working in a variety of disciplines and backgrounds, we understand the difficulties surrounding this topic. In order to address these concerns, we formed a team and embarked on a project to identify the ‘crisis’ and the ongoing challenge it creates for ourselves and our stakeholders.

Our operating definitions of ‘reproducibility’ and ‘replication’ were as follows:

- *Reproducibility* measures whether a study or experiment can be reproduced in its entirety. To achieve adequate reproducibility, studies implement measures to support verification of research, including, for example, sharing data and methods. No single factor or method alone achieves reproducibility in a study, and likewise, many factors can result in a study with poor reproducibility (Munafò et al., 2017).

- *Replication* is the attempt to recreate the conditions believed sufficient for obtaining a previously observed finding and is the means of establishing reproducibility of a finding with new data (Open Science Collaboration, 2015).

It should be noted that formal definitions of the two terms ‘reproducibility’ and ‘replicability’ were offered in a report by the National Academies of Sciences, Engineering, and Medicine (2019). The distinct differences between the two definitions did not play a key part in this project. The Academies’ definitions were published after most of our own investigation was complete, and the terms were deemed somewhat interchangeable throughout this exercise.

To fully understand the reproducibility crisis, we must understand the multitude of contributing factors influencing reproducibility (Open Science Collaboration, 2015). We gathered a number of resources and began studying. Our findings resulted in two products discussed in this paper: an annotated list of 122 resources reviewed to understand the ‘crisis,’ and a primer that lists the top issues and solutions defined within the resources. These items provided our team a common language and understanding of the problem which we can use in our profession moving forward.

Project origins

Our journey started with a proposal brought up within CENDI⁷, a U.S. federal scientific and technical information managers group. CENDI is ‘a volunteer-powered membership organization that serves the federal information community - that is, all those who create, manage, aggregate, organize, and provide access to federally-funded data and publications’ within federal scientific and technical information agencies. Its member organizations represent a cross-section of federal data and publication stakeholders—including libraries, data centers, aggregators, information technology developers, and content management providers. CENDI’s mission is to ‘increase the impact of federally funded science and technology by improving the management and dissemination of data and information’ (CENDI, 2019).

CENDI is home to a small number of working groups, including the Data Curation Discussion Group (DCDG). While principle CENDI members are the managers of federal scientific and technological information libraries, the DCDG members are, in the main, hands-on data management and curation staff within the libraries and home agencies. The DCDG’s goal is ‘to collaborate across agencies, employ data curation best practices, tools, and workflows, promote efficiencies and consistency, work through challenges, and avoid ‘reinvention.’’ (Christiansen, 2017). In late 2017, CENDI leadership proposed that DCDG develop tools to assess and address the ‘reproducibility crisis’ with possible outcomes being:

- developing or populating a website with content that puts reproducibility challenges in context and identifies both real issues and spurious concerns
- sharing information about approaches to reproducibility among the CENDI members
- disseminating information about best practices within the respective agencies, based on consensus findings in CENDI and/or noted elsewhere

Project goals and methodology

The DCDG began the project over the first half of 2018 as an effort to familiarize CENDI members and interested parties on the topics and issues surrounding reproducibility. A subset of DCDG members formed a team that collected and annotated resources on various aspects of scientific reproducibility and replicability. The resources the team reviewed were primarily articles appearing in top results from a Google Scholar search, as well as resources cited within those initial results, dating from 2005 to the present. A number of other relevant resources such as books, presentations, and websites were also included.

The team divided the list of resources and each member was assigned as a reader who offered an annotation or abstract of the resource. They also identified the top three issues and/or solutions shared within each. These issues and solutions populated separate spreadsheets and given definitions based on the literature.

The goals of this exercise were to:

1. Identify the variety of issues named as barriers to reproducibility
2. Identify solutions to such barriers
3. Reflect on how researchers, information professionals, and librarians perceive the 'reproducibility crisis'

Although this was not a formal analysis of the literature, or a refined scientific experiment, this information gathering exercise served to inform our group about the reproducibility crisis. The result of DCDG's work included here is an annotated list of 122 resources which were reviewed, and a primer which identifies and defines key terms that surfaced in this exercise. These terms are a snapshot of our interpretation of the resources when it was undertaken in early 2018.

When reviewing this list and the readers' annotations, a formal rubric was not developed or required for participation. Nor did members attempt to agree on definitions or classification. As each member brought their own professional background to bear on their assessment of the themes in each resource, we attempted to achieve a group understanding of a very broad issue with each participant contributing their own perspective. Reporting these results is our desire to share our findings without any attempt to filter understanding of each participant. Additionally, while most resources in this list acknowledged a reproducibility problem on some level, we did not differentiate between 'pro-crisis' or 'no-crisis' authors. Group members read through each resource, identifying the primary problems or issues raised, as well as any solutions proposed.

The final products the DCDG produced (as of October 2019) from this effort are publicly available at <https://doi.org/10.18434/M32150> They include:

- *Reproducibility Resources*: an annotated list of resources that were evaluated, with annotations penned by DCDG members, with resource citations
- *Issues*: a list of terms deemed as 'issues' or problems relating to reproducibility, with definitions

- *Solutions*: a list of terms deemed as ‘solutions’ to the reproducibility issues, with definitions
- *Metrics*: tallies of how often each term was selected as a primary theme of the resource, the resource dates, the scientific disciplines represented, and the types of resources reviewed

Resource metrics

Source information for this project consists of a variety of material types from numerous research disciplines. In total, we reviewed 122 resources relating to reproducibility in research data. To provide perspective on our source material, we include metrics for the resources selected for this project. The bulk of our research was derived from scholarly journal articles (59 percent), but we also included other sources such as books (4.1percent), presentations (0.8 percent), and websites (7.4 percent).

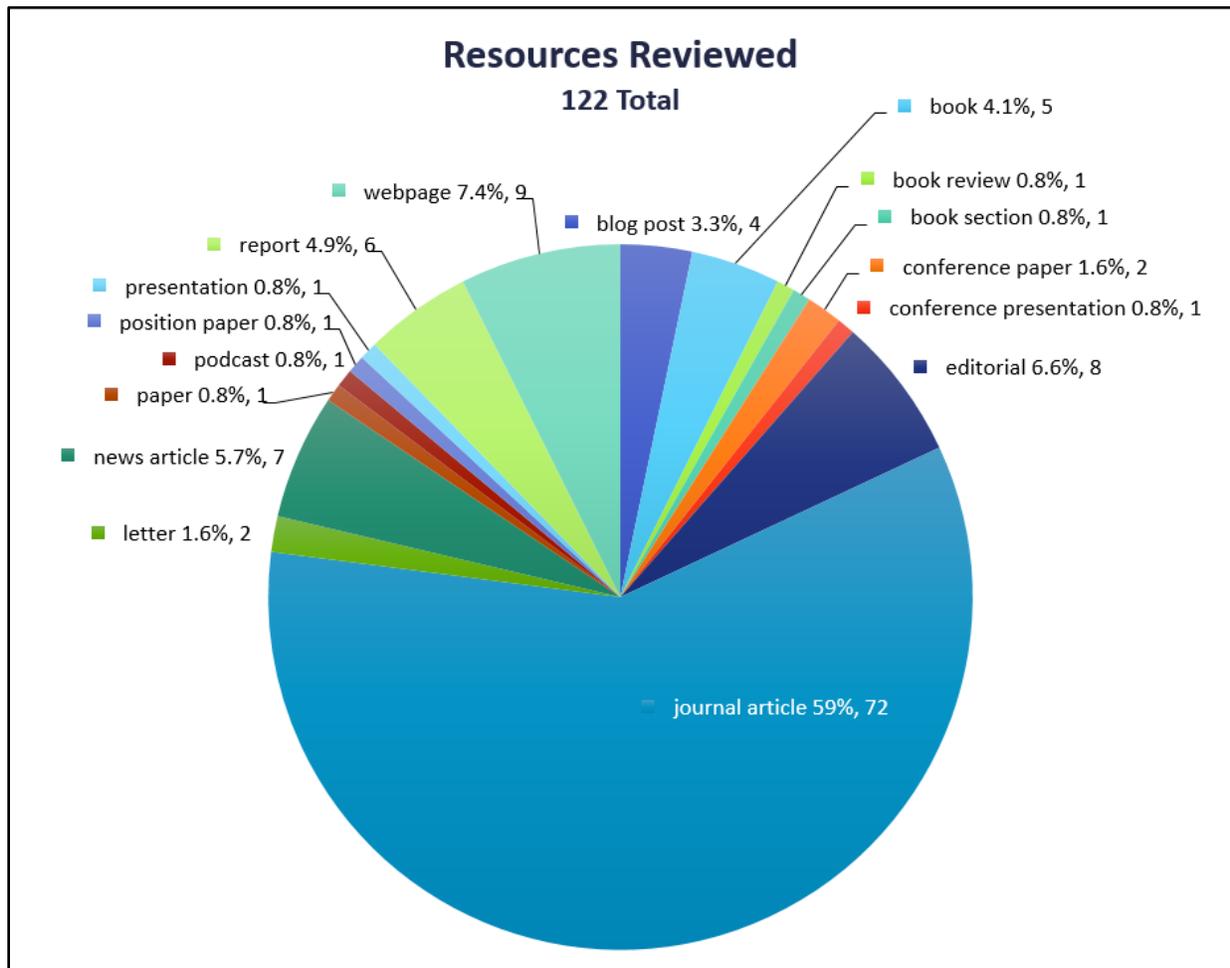


Figure 1: Resource types reviewed by the group varied, but consisted overwhelmingly of journal articles.

Most of the resources, about 93 percent, discussed reproducibility and pointed to or offered solutions and best practices. A few, roughly 7 percent, pointed out reproducibility issues but did not speculate on causes and/or offered no solutions.

The compiled resources represent a broad sampling of items written for, or about, researchers within a variety of disciplines. The majority of the resources address science in general, but the total number reflects an impressive breadth of disciplines. They range from biomedicine to economics to psychology. In all, the materials target over 20 different disciplines, all discussing or analyzing the issue of reproducibility from their respective viewpoints.

While numerous, this collection does not represent an exhaustive list of resources. Therefore, this compilation of resources, while annotated, does not constitute a proper 'literature review' in the typical sense. There was no comprehensive selection of papers from any single discipline, nor were all scientific disciplines represented. This analysis did not intend to favor any one discipline over another, as we intended to gather more general information about reproducibility wherever the topic appeared in various resources.

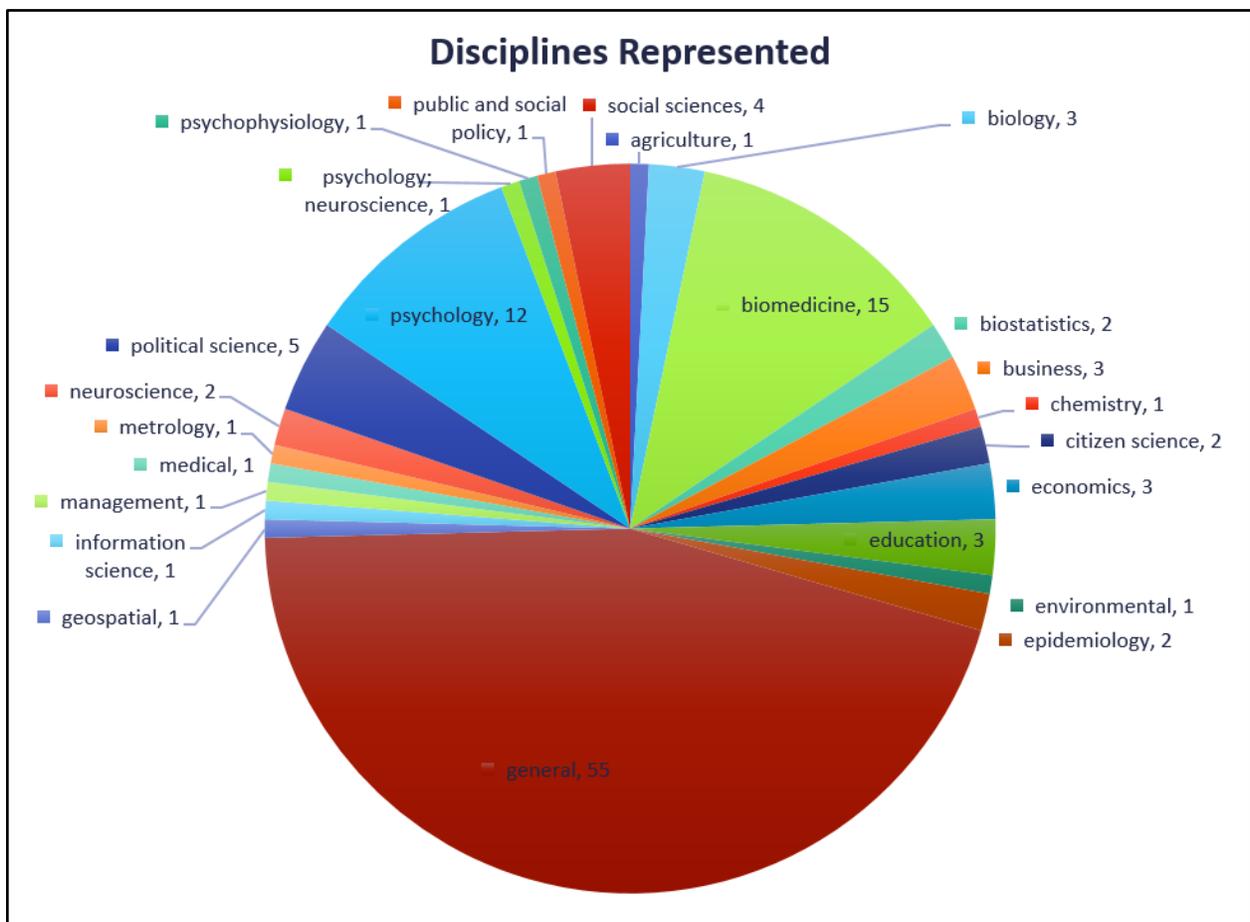
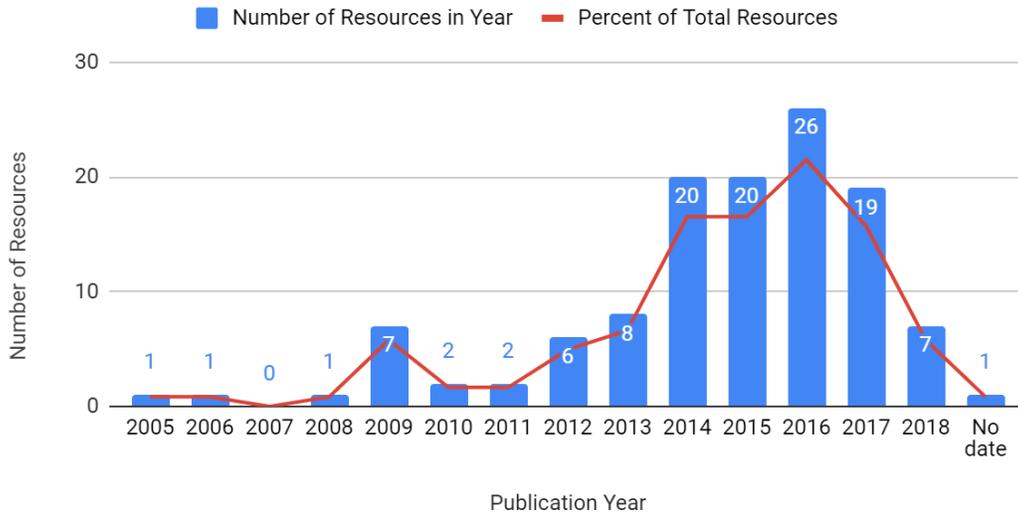


Figure 2: Disciplines represented in the reviewed resources varied greatly - over 20 are represented in this exercise.

The resources reviewed date from 2005 to the present. Over 70 percent of these resources were published between 2014 to 2017. The lack of articles for 2018 occurred because the group compiled most of these resources in early 2018, at the start of the project.

Number of Resources in Year and Percent of Total Resources (n=122)



Number of Publications in Other Databases

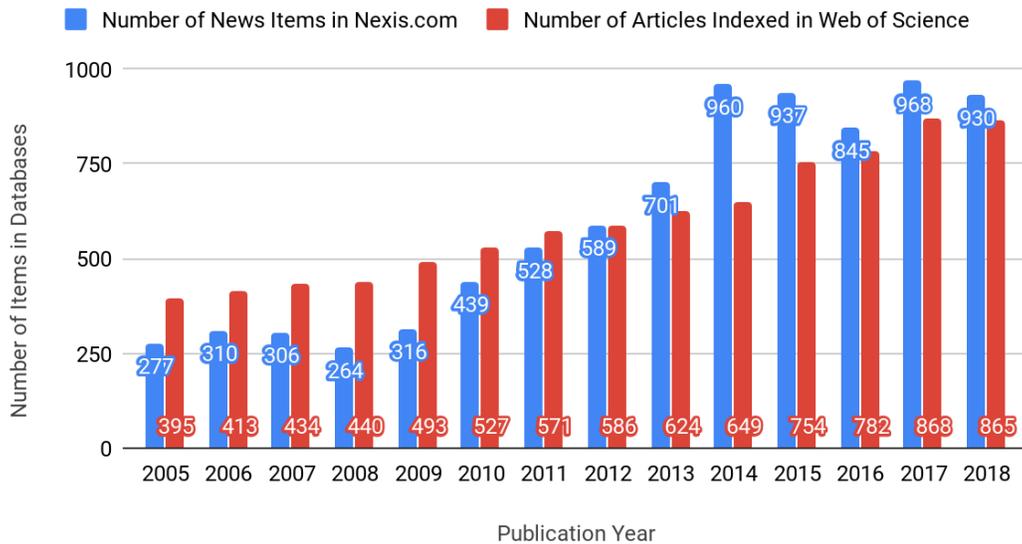


Figure 3. Charts showing the number of resources published in years 2005-2018. The top chart shows the numbers from the DCDG resource list. The bottom charts publication counts from Nexis.com and Web of Science that contained 'reproducibility' or 'replicability' in the title, headline, or lead paragraph. The overall increase in publications on the topic from 2014-2017 is similar both in the DCDG resources list, and publications indexed in the other two databases.

While the resources reviewed comprise only a fraction of what was published during that time period, they do appear to be representative of that period. Results from other databases reflect a similar

increase in publications on the topic of reproducibility during that same period. A search of Nexis.com reveals the number of news items mentioning 'reproducibility' or 'replicability' in headlines or lead paragraphs follow the same peak, from 2014 to 2017. Results from the science database Web of Science (<https://apps.webofknowledge.com/>) show a steady increase in the publications indexed from 2005-2018, with the highest being the latter years, 2017 and 2018.

Only one publication was added to our collection long after the others were gathered: the aforementioned report by the National Academies of Science, which provides recommendations to improve reproducibility and replicability in science.

Variety of issues and solutions

Many resources centered on the general problem of reproducibility and replicability, while others focused more on specific causes of the reproducibility crisis. Of the 57 issues identified, the most common problems or issues discussed in our cross-section of resources include: replicability (in 23 resources); reproducibility (20); reproducibility crisis (20); bias (10); cherry picking (10); publish or perish (10); data sharing (8); data quality (7), and, researcher misconduct (6). The remaining 48 primary topic issues presented in five or fewer of the reviewed resources.

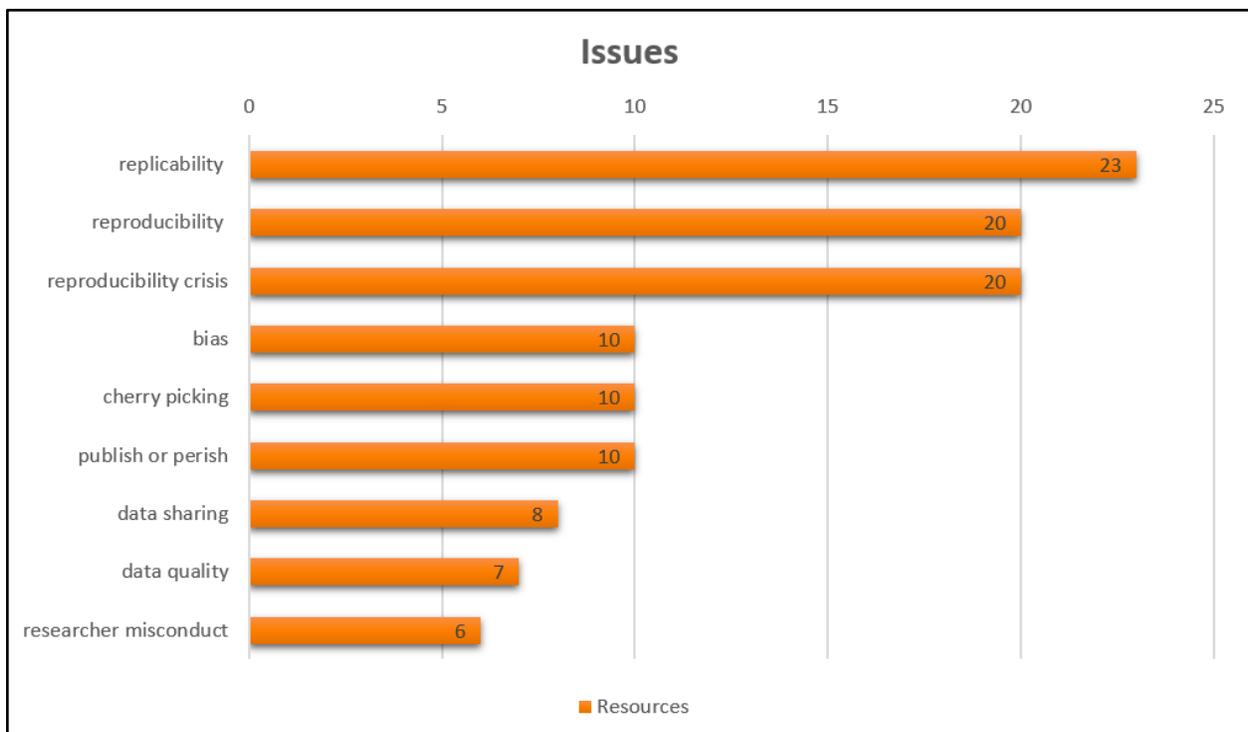


Figure 4: The nine (9) issues most frequently selected as a primary topic in the reviewed resources.

Solutions presented in the resources typically fell into three categories: best practices/standards; transparency/sharing; and, culture. Solutions in the best practices category included general

recommendations on how to conduct data management at different stages of the research. Solutions in the transparency category covered topics about sharing methods and data. The culture category comprised a broader look at research community behaviors and proposed methods for turning best practices into regular practice. Many of the terms and definitions overlap, and many issues were also selected as solutions. The scientific research community is diverse and complex. Ergo, the issues often overlap and require a comprehensive view when considering solutions.

Of the 55 solutions identified, the most common solutions discussed in our resource list were: experimental design (in 16 resources); transparency (16); code sharing (13); data sharing (13); replication studies (13); publication policy (12); standards (12); best practices (11); methods sharing (11); training (11); and, quality assurance (11). The remaining 44 solutions were selected as a primary topic in ten or fewer of the reviewed resources.

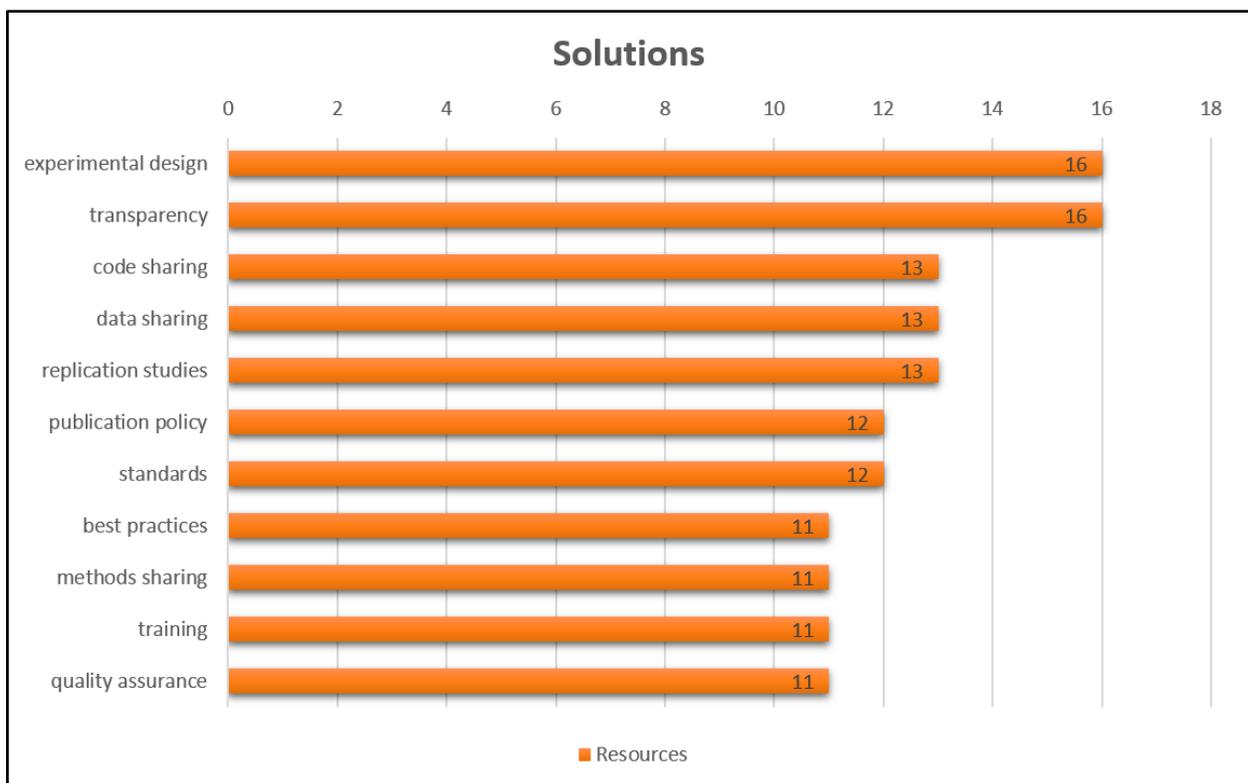


Figure 5: The eleven (11) most frequent solutions selected as a primary topic in the reviewed resources.

Primer of terminology and findings

Reproducibility issues or challenges

This exercise revealed many concepts surrounding research reproducibility, though several ideas appeared much more frequently across the board. This section highlights some of the terms derived from the reviewed materials and most commonly identified as reproducibility issues or challenges.

Many of the reported issues relate to or feed off one another. In the terms defined below we note other terms from the list that are related, where applicable. For a full list of the selected issues and their definitions, review the related data file at <https://doi.org/10.18434/M32150>.

Replicability

Replication is the attempt to recreate the conditions believed sufficient for obtaining a previously observed finding and is the means of establishing reproducibility of a finding with new data (Open Science Collaboration, 2015). [Related terms: reproducibility; repeatability]

Reproducibility

Reproducibility measures whether a study or experiment can be reproduced in its entirety. To achieve adequate reproducibility, studies implement measures to support verification of research, including, for example, sharing data and methods. No single factor or method alone achieves reproducibility in a study, and likewise, many factors can result in a study with poor reproducibility (Munafò et al., 2017). [Related terms: replicability; repeatability]

Reproducibility crisis

The reproducibility crisis is defined as widespread failure to replicate the results of experiments and studies (Weir, 2015). While many acknowledge the problem and see a need for research and experimental reform, many people debate the reproducibility problem as exaggerated (Baker, 2016c).

Bias

Bias includes prejudice in favor of or against one thing, person, or group compared with another, usually in a way considered unfair. Two common types of bias in research studies are confirmation bias and hindsight bias. Confirmation bias promotes the tendency to focus on evidence that is in line with our expectations or favored explanation. Hindsight bias is the tendency to see an event as having been predictable only after it has occurred (Munafò et al., 2017). [Related terms: cherry picking; replication studies (solution)]

Cherry picking

Cherry picking data includes suppressing evidence, or the fallacy of incomplete evidence by pointing to individual cases or data that seem to confirm a particular position or statistical significance, while ignoring a significant portion of related cases or data that may contradict that position (Baker, 2016a). [Related terms: bias; P-hacking]

Publish or perish

'Publish or perish' is a phrase coined to describe the pressure in academia to rapidly and continuously publish academic work to sustain or further one's career. Frequent publication is one of the few methods at scholars' disposal to demonstrate academic talent. The desire or need to publish work at a near-constant rate can lead to problems in reproducibility, and can lead to issues concerning selective reporting, also known as 'cherry picking' (Baker, 2016a). [Related terms: incentives; publication policy (solution); culture shift (solution)]

Data sharing

Definition is included alongside 'Code sharing' in the Solutions section below, as it appeared as both an issue and solution in this analysis.

Data quality

Many definitions and factors determine data quality. However, data ‘fit for their intended uses in operations, decision making and planning’ are generally considered high quality. Incorrect or incomplete data used to influence decision-making minimizes accuracy and strategic advantage (Redman, 2008). Bad or low-quality data can result from many avenues, including negative cultural influences such as ‘publish or perish’ as well as researcher misconduct, bias, P-hacking, or cherry picking data, among others. [Related terms: validation; trust; quality assurance (solution)]

Researcher misconduct

‘The National Science Foundation (2001) defined scientific misconduct as fabrication, falsification, or plagiarism in proposing, performing, or reviewing research or in reporting research results. Such misconduct is committed intentionally, knowingly, or in disregard of accepted practices. Fabrication of data involves totally inventing a data set, while falsification refers to manipulation of equipment or changing data such that the research is not accurately represented in the research report’ (Stroebe, Postmes and Spears, 2012). [Related terms: trust; data quality]

Potential solutions to the reproducibility crisis

While many potential solutions to reproducibility issues appeared throughout the reviewed resources, here we highlight some of the terms appearing the most frequently, or which we felt were important. As with the reproducibility issues discussed above, many of these reported solutions intersect. It is worthwhile to reiterate that many of the terms and ideas encountered during our research appeared as both issues and solutions—such as publication policy, incentives, training, data sharing, and data quality. For example, lack of ‘data sharing’ is cited as a hindrance to reproducibility. Others cite ‘data sharing’ as a potential solution. Again, where applicable, we append other terms to these definitions that are related to those we highlight here. For a full list of the selected solutions and their definitions, review the related data file available at <https://doi.org/10.18434/M32150>.

Experimental design

Experimental design aims to describe or explain the variation of information under conditions that are hypothesized to reflect the variation, and use this knowledge to collect more accurate data (Aceves-Bueno et al., 2017). A framework for a systematic process to guide researchers and reviewers in assessing, documenting, and mitigating the sources of uncertainty in a study enhance comparability and reproducibility (Plant et al., 2018). Experimental design features should enhance, or facilitate inference about, the reproducibility and generalizability of the expected results (Würbel, 2017). [Related terms: pre-registration of results; case study]

Transparency

Simply put, transparency means ‘provable to the outside’ (Bartling and Fecher, 2015). Transparency is the basis of open science, which refers to the process of making the content and process of producing evidence and claims clear and accessible to others. Transparency is a scientific ideal, and adding ‘open’ should therefore be redundant (Munafò et al., 2017). [Related terms: open review; open science; data sharing; code sharing; methods sharing; trust (issue)]

Data sharing, code sharing

Studies and research that implement data sharing support verification of research and conduct alternative analysis. Sharing data in public repositories offers field-wide advantages in terms of accountability, data longevity, efficiency and quality (Peng, Dominici and Zeger, 2006). Likewise, code sharing provides discovery and access to the details of computational analysis including programming code and data (Gezelter, 2015). These terms appear as both issues and solutions, since lack of sharing creates reproducibility issues. [Related terms: Data Discovery Index; methods sharing; open science; transparency]

Replication studies

Somewhat related to experimental design, replication is a term referring to the repetition of a research study, generally with different situations and different subjects, to determine if the basic findings of the original study can be applied to other participants and circumstances (Gezelter, 2015). Replication studies also help identify potential biases in the original study and serve as a basis for confirming or disconfirming prior findings (Spector, Johnson and Young, 2014) (Camerer et al., 2016). [Related terms: collaborative replication; replication files; reproducible research standard (RRS); bias (issue)]

Publication policy

Journals have power to enforce transparency and reproducibility through their review and publication policies. This could help establish and enforce best practices. For instance, journals could require authors to register reports in advance so that the study protocol and analysis plan is locked in place before data collection even begins, and scientists should be encouraged to store methods, data, and code in repositories to help other groups reproduce experiments. One source suggested 5 to 10 percent of research funding should be spent on replication studies, and journals should devote more space to replication studies and null results (McNutt, 2014) (Begley and Ioannidis, 2015). [Related terms: data citation; funding agency requirements; incentives (issue); open review; public access; publish or perish (issue)]

Standards

Standards comprise the fundamental reference for a system of weights and measures, against which all other measuring devices are compared. Standards contribute to improved research practices and promote positive change (Capes-Davis and Neve, 2016). [Related terms: best practices; metrological standards; reporting guidelines; reproducible research standard (RRS)]

Training

Open Science, the movement to make scientific products and processes accessible to, and reusable by all, relies on culture and knowledge as much as it does on technologies and services. Convincing researchers of the benefits of changing their practices, and equipping them with the skills and knowledge needed to do so can happen through training and education. A recommendation from citizen science states that "Training, in particular, has been shown elsewhere to enhance accuracy and credibility (Freitag, 2016, Kosmala, 2016). [Related terms: best practices; standards; trust (issue)]

Quality assurance

Quality is an encompassing term comprising utility, objectivity, and integrity (National Institute of Standards and Technology, 2009). Whether in a laboratory setting or defining a quality system, quality assurance is the “system of activities whose purpose is to provide to the producer or user of a product or a service the assurance that it meets defined standards of quality with a stated level of confidence” (Taylor, 1987). Implementing quality assurance may involve a variety of checks for data completeness, validity, consistency, precision, and accuracy, among other aspects of the data (Wiggins et al., 2011). Research units often take an ad-hoc approach to methods and workflow, but standardizing operations and following certified protocols increases confidence in research results (Baker, 2016b). [Related terms: standards; best practices; data quality (issue); trust (issue); validation (issue)]

Incentives

Rewards for publishing, often tied to showcasing certain results, comprise the biggest challenge to widespread adoption of open data. Conversely, well-conceived incentives may also provide solutions for increased reproducibility. If journals in particular regulate and highlight incentives for research practices promoting reproducibility, researchers will more widely adopt these positive practices (Gezelter, 2015) (Begley and Ioannidis, 2015). [Related terms: funding agency requirements; publication policy; publish or perish (issue)]

Culture shift

An overarching theme with regard to reproducibility solutions boils down to a culture shift throughout the research endeavor and data gathering. Culture shift encompasses changing beliefs, behaviors, and outcomes. Industries must broadly address their practices at all stages of data collection, processing, publication, dissemination, and preservation to make reproducibility commonplace (Baker, 2016b).

Opportunities for future study

The resource list and primer presented here are an introduction to the landscape of the reproducibility crisis. The terms are defined broadly and remain at surface-level with regard to the topics described. While the final resource list contains annotations with key terms and definitions, more targeted research will uncover more nuances of specific problems and solutions. Subject-specific analysis of reproducibility, as well as further and more honed examination of any of the issues and/or solutions may produce additional understanding.

The work of this group is only the beginning for DCDG and other data and information managers. Numerous opportunities for future research studies remain. Such work could underpin the formation of other resources for those pursuing study of this topic, and for those who wish to improve reproducibility as a means to increase confidence in science within their own organizations.

This collection of annotated resources spans the past fourteen years, with the bulk published in the last decade when digital methods have been the norm for scientific research. While digital data and modern computing and modeling practices certainly may cause their own unique reproducibility issues, an analysis of research practice in earlier literature may reveal more clarity into the scope and depth of these problems (Bastian, 2016). As new research data insights, trends, and studies emerge, this resource list and primer should be updated to reflect the latest information that pertains to the reproducibility

crisis. In addition to adding more literature to the existing compilation of resources, performing more in-depth exercises with these resources—such as textual analysis—may uncover new insights that can accurately inform training or education strategies to increase reproducibility.

Conclusion

The issues surrounding reproducibility present great challenges. Comprehending the numerous, nuanced causes within this exercise sometimes seemed insurmountable. While potentially overwhelming, our group made strides to understand the issues, root causes, and history of the reproducibility crisis in order to create a guide for ourselves. Many organizations and individuals who recognize research reproducibility as an issue may not currently have the knowledge or resources to effect significant change. However, in summarizing a portion of the available literature on this topic, it is our hope that information professionals now have a better starting point to begin incremental change in promoting reproducible research.

Ultimately, we concluded that because the reproducibility crisis stemmed from such a wide variety of causes, stakeholders must take a multi-pronged approach to tackling the problem. A culture shift across all branches of research must occur to reverse the distrust this crisis has engendered. In moving forward, we must also realize our limitations. As data managers and librarians, many reproducibility issues stem from actions that occur prior to or after our typical involvement with the research. We recognize that we can facilitate reproducibility from within our own roles. Given our positions as information professionals reflecting on the nature and scope of these reproducibility issues, our objective should involve education, training, and building awareness.

For example, as data managers, we may not write data management plans, but we do share information and guidelines about how to write, curate, and archive them. We do not generate the data that results from research, but we can assist with organizing data, finding documentation standards for data, creating proper metadata, and assist in building and managing trusted repositories for the data. We do not format or publish the data, but we can share best practices for FAIR data, thereby contributing to research data that is Findable, Accessible, Interoperable, and Reusable (Wilkinson, 2016). Even from our set positions we can take concrete steps to aid and influence activities that move toward the goal of reproducibility. Increasing education and awareness within our fields helps the cause.

Common past practice may have seen librarians contributing to the scientific endeavor in very limited ways, such as assisting with initial literature access and reviews, or as cataloging and preserving reported scientific results. However, the evolution of modern scientific research has, as discussed above, opened up a number of roles for library, information, and data professionals throughout the entire scientific research lifecycle. Our participation can positively impact scientific reproducibility and replicability. First, we must understand the issues, and this paper is one contribution to develop that understanding. Next we should apply that comprehension to aid our colleagues across research disciplines.

References

- Aceves-Bueno, E., Adeleye, A., Feraud, M., Huang, Y., Tao, M., Yang, Y. and Anderson, S. (2017). The Accuracy of Citizen Science Data: A Quantitative Review. *The Bulletin of the Ecological Society of America*, 98(4), pp.278-290. <https://doi.org/10.1002/bes2.1336>
- Baker, M. (2016a). 1,500 scientists lift the lid on reproducibility. *Nature*, 533(7604), pp.452-454. <https://doi.org/10.1038/533452a>
- Baker, M. (2016b). How quality control could save your science. *Nature*, 529(7587), pp.456-458. <https://doi.org/10.1038/529456a>
- Baker, M. (2016c). Psychology's reproducibility problem is exaggerated – say psychologists. *Nature*. <https://doi.org/10.1038/nature.2016.19498>
- Bartling, S. and Fecher, B. (2015). Could Blockchain provide the technical fix to solve science's reproducibility crisis?. *LSE Impact of Social Sciences*. [online] Available at: http://eprints.lse.ac.uk/67354/1/Could_Blockchain_provide_technical_fix.pdf [Accessed 10 Sep. 2019].
- Bastian, H. (2016). Reproducibility crisis timeline—milestones in tackling research reliability. [online] Phys.org. Available at: <https://phys.org/news/2016-12-crisis-timelinemilestones-tackling-reliability.html> [Accessed 10 Sep. 2019].
- Begley, C. and Ioannidis, J. (2015). Reproducibility in Science. *Circulation Research*, 116(1), pp.116-126. <https://doi.org/10.1161/circresaha.114.303819>
- Camerer, C., Dreber, A., Forsell, E., Ho, T., Huber, J., Johannesson, M., Kirchler, M., Almenberg, J., Altmejd, A., Chan, T., Heikensten, E., Holzmeister, F., Imai, T., Isaksson, S., Nave, G., Pfeiffer, T., Razen, M. and Wu, H. (2016). Evaluating replicability of laboratory experiments in economics. *Science*, 351(6280), pp.1433-1436. <https://doi.org/10.1126/science.aaf0918>
- Capes-Davis, A. and Neve, R. (2016). Authentication: A Standard Problem or a Problem of Standards? *PLOS Biology*, 14(6), p.e1002477. <https://doi.org/10.1371/journal.pbio.1002477>
- Cendi.gov. (2019). *CENDI*. [online] Available at: <https://cendi.gov/> [Accessed 10 Sep. 2019].
- Christiansen, L. (2017). *Expanding the U.S. Federal Data Curation Community: Year 01 at the National Transportation Library*. Research Data Alliance Tenth Plenary. Montreal, Quebec, Canada. <https://doi.org/10.21949/1504432>
- Crane, H. (2017). *Why "Redefining Statistical Significance" Will Not Improve Reproducibility and Could Make the Replication Crisis Worse*. [online] arXiv.org. Available at: <http://arxiv.org/abs/1711.07801> [Accessed 10 Sep. 2019].
- Freitag, A., Meyer, R. and Whiteman, L. (2016). Strategies Employed by Citizen Science Programs to Increase the Credibility of Their Data. *Citizen Science: Theory and Practice.*, 1(1), p.2. <https://doi.org/10.5334/cstp.6>

- Gezelter, J. (2015). Open Source and Open Data Should Be Standard Practices. *The Journal of Physical Chemistry Letters*, 6(7), pp.1168-1169. <https://doi.org/10.1021/acs.jpcllett.5b00285>
- Kosmala, M., Wiggins, A., Swanson, A. and Simmons, B. (2016). Assessing data quality in citizen science. *Frontiers in Ecology and the Environment*, 14(10), pp.551-560. <https://doi.org/10.1002/fee.1436>
- McNutt, M. (2014). Journals unite for reproducibility. *Science*, 346(6210), pp.679-679. <https://doi.org/10.1126/science.aaa1724>
- Munafò, M., Nosek, B., Bishop, D., Button, K., Chambers, C., Percie du Sert, N., Simonsohn, U., Wagenmakers, E., Ware, J. and Ioannidis, J. (2017). A manifesto for reproducible science. *Nature Human Behaviour*, 1(1). <https://doi.org/10.1038/s41562-016-0021>
- National Academies of Sciences, Engineering, and Medicine (2019). *Reproducibility and Replicability in Science*. Washington, DC: The National Academies Press. <https://doi.org/10.17226/25303>
- National Institute of Standards and Technology (2009). NIST Information Quality Standards. [online] Available at: <https://www.nist.gov/nist-information-quality-standards> [Accessed 18 Oct. 2019]
- Open Science Collaboration (2015). Estimating the reproducibility of psychological science. *Science*, 349(6251), aac4716. <https://doi.org/10.1126/science.aac4716>
- Peng, R., Dominici, F. and Zeger, S. (2006). Reproducible Epidemiologic Research. *American Journal of Epidemiology*, 163(9), pp.783-789. <https://doi.org/10.1093/aje/kwj093>
- Plant, A., Becker, C., Hanisch, R., Boisvert, R., Possolo, A. and Elliott, J. (2018). How measurement science can improve confidence in research results. *PLOS Biology*, 16(4), p.e2004299. <https://doi.org/10.1371/journal.pbio.2004299>
- Redman, T. (2008). *Data Driven: Profiting from Your Most Important Business Asset*. Boston: Harvard Business Review Press.
- Spector, J., Johnson, T. and Young, P. (2014). An editorial on replication studies and scaling up efforts. *Educational Technology Research and Development*, 63(1), pp.1-4. <https://doi.org/10.1007/s11423-014-9364-3>
- Stroebe, W., Postmes, T. and Spears, R. (2012). Scientific Misconduct and the Myth of Self-Correction in Science. *Perspectives on Psychological Science*, 7(6), pp.670-688. <https://doi.org/10.1177/1745691612460687>
- Taylor, J. (1987). *Quality assurance of chemical measurements*. Chelsea: Lewis Publ.
- Weir, K. (2015). A reproducibility crisis?. *Monitor On Psychology*, [online] 46(9), p.39. Available at: <https://www.apa.org/monitor/2015/10/share-reproducibility> [Accessed 10 Sep. 2019].
- Wiggins, A., Newman, G., Stevenson, R. and Crowston, K. (2011). Mechanisms for Data Quality and Validation in Citizen Science. *2011 IEEE Seventh International Conference on e-Science Workshops*. <https://doi.org/10.1109/eScienceW.2011.27>

Wilkinson, M.D., Dumontier, M., Aalbersberg, I.J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J.W., da Silva Santos, L.B., Bourne, P.E. and Bouwman, J., 2016. The FAIR Guiding Principles for scientific data management and stewardship. *Scientific data*, 3. <https://doi.org/10.1038/sdata.2016.18>

Würbel, H. (2017). More than 3Rs: the importance of scientific validity for harm-benefit analysis of animal research. *Lab Animal*, 46(4), pp.164-166. <https://doi.org/10.1038/lab.1220>

Appendix A

List of Resources Evaluated

Aceves-Bueno, E., Adeleye, A., Feraud, M., Huang, Y., Tao, M., Yang, Y. and Anderson, S. (2017). The Accuracy of Citizen Science Data: A Quantitative Review. *The Bulletin of the Ecological Society of America*, 98(4), pp.278-290. <https://doi.org/10.1002/bes2.1336>

Aguinis, H., Cascio, W. and Ramani, R. (2017). Science's reproducibility and replicability crisis: International business is not immune. *Journal of International Business Studies*, 48(6), pp.653-663. <https://doi.org/10.1057/s41267-017-0081-0>

Akpan, N. (2017). Why bad science is plaguing health research—and how to fix it. [online] PBS NewsHour. Available at: <https://www.pbs.org/newshour/science/why-bad-science-is-plaguing-health-research-rigor-mortis-richard-harris> [Accessed 10 Sep. 2019].

Anderson, C., Anderson, J., Assen, M., Attridge, P., Attwood, A., Axt, J., Babel, M., Bahník, Š., Baranski, E. and Barnett-Cowan, M. (2019). Reproducibility Project: Psychology. [online] OSF. Available at: <https://osf.io/ezcuj/> [Accessed 10 Sep. 2019]. <https://doi.org/10.17605/OSF.IO/EZCUJ>

Aschwanden, C. (2015). Science Isn't Broken. (online). FiveThirtyEight. Available at: <https://fivethirtyeight.com/features/science-isnt-broken/> [Accessed 10 September 2019]

Asendorpf, J., Conner, M., De Fruyt, F., De Houwer, J., Denissen, J., Fiedler, K., Fiedler, S., Funder, D., Kliegl, R., Nosek, B., Perugini, M., Roberts, B., Schmitt, M., van Aken, M., Weber, H. and Wicherts, J. (2013). Recommendations for Increasing Replicability in Psychology. *European Journal of Personality*, 27(2), pp.108-119. <https://doi.org/10.1002/per.1919>

Baker, M. (2015). Over half of psychology studies fail reproducibility test. *Nature*. <https://doi.org/10.1038/nature.2015.18248>

Baker, M. (2016). 1,500 scientists lift the lid on reproducibility. *Nature*, 533(7604), pp.452-454. <https://doi.org/10.1038/533452a>

Baker, M. (2016). How quality control could save your science. *Nature*, 529(7587), pp.456-458. <https://doi.org/10.1038/529456a>

Baker, M. (2016). Muddled meanings hamper efforts to fix reproducibility crisis. *Nature*. <https://doi.org/10.1038/nature.2016.20076>

Baker, M. (2016). Psychology's reproducibility problem is exaggerated – say psychologists. *Nature*. <https://doi.org/10.1038/nature.2016.19498>

Bal, L. (2015). Is science broken? The reproducibility crisis. [Blog] On Biology. Available at: <https://blogs.biomedcentral.com/on-biology/2015/03/20/is-science-broken-a-reproducibility-crisis/> [Accessed 10 September 2019].

Bartling, S. and Fecher, B. (2015). Could Blockchain provide the technical fix to solve science's reproducibility crisis?. LSE Impact of Social Sciences. [online] Available at: http://eprints.lse.ac.uk/67354/1/Could_Blockchain_provide_technical_fix.pdf [Accessed 13 Aug. 2019].

Bastian, H. (2016). Reproducibility crisis timeline—milestones in tackling research reliability. [online] Phys.org. Available at: <https://phys.org/news/2016-12-crisis-timelinemilestones-tackling-reliability.html> [Accessed 10 Sep. 2019].

Bauer, H. (2015). How Medical Practice Has Gone Wrong: Causes of the Lack-of-Reproducibility Crisis in Medical Research. *Journal of Controversies in Biomedical Research*, 1(1). <https://doi.org/10.15586/jcbmr.2015.8>

Begley, C. (2013). Six red flags for suspect work. *Nature*, 497(7450), pp.433-434. <https://doi.org/10.1038/497433a>

Begley, C. and Ioannidis, J. (2015). Reproducibility in Science. *Circulation Research*, 116(1), pp.116-126. <https://doi.org/10.1161/CIRCRESAHA.114.303819>

Bergh, D., Sharp, B., Aguinis, H. and Li, M. (2017). Is there a credibility crisis in strategic management research? Evidence on the reproducibility of study findings. *Strategic Organization*, 15(3), pp.423-436. <https://doi.org/10.1177/1476127017701076>

Bissell, M. (2013). Reproducibility: The risks of the replication drive. *Nature*, 503(7476), pp.333-334. <https://doi.org/10.1038/503333a>

Bollen, K., Cacioppo, J., Kaplan, R., Krosnick, J., Olds, J. and Dean, H. (2015). Social, Behavioral, and Economic Sciences Perspectives on Robust and Reliable Science. [online] Nsf.gov. Available at: https://www.nsf.gov/sbe/AC_Materials/SBE_Robust_and_Reliable_Research_Report.pdf [Accessed 10 Sep. 2019].

Boos, D. and Stefanski, L. (2011). P-Value Precision and Reproducibility. *The American Statistician*, 65(4), pp.213-221. <https://doi.org/10.1198/tas.2011.10129>

Borgman, C. (2019). Research Data, Reproducibility, and Curation. *Digital Social Research: A Forum for Policy and Practice* — Oxford Internet Institute. [online] Oii.ox.ac.uk. Available at: <http://www.oii.ox.ac.uk/events/?id=487> [Accessed 10 Sep. 2019].

Camerer, C., Dreber, A., Forsell, E., Ho, T., Huber, J., Johannesson, M., Kirchler, M., Almenberg, J., Altmejd, A., Chan, T., Heikensten, E., Holzmeister, F., Imai, T., Isaksson, S., Nave, G., Pfeiffer, T., Razen, M. and Wu, H. (2016). Evaluating replicability of laboratory experiments in economics. *Science*, 351(6280), pp.1433-1436. <https://doi.org/10.1126/science.aaf0918>

Camerer, C., Dreber, A., Holzmeister, F., Ho, T., Huber, J., Johannesson, M., Kirchler, M., Nave, G., Nosek, B., Pfeiffer, T., Altmejd, A., Buttrick, N., Chan, T., Chen, Y., Forsell, E., Gampa, A., Heikensten, E., Hummer, L., Imai, T., Isaksson, S., Manfredi, D., Rose, J., Wagenmakers, E. and Wu, H. (2018). Evaluating the replicability of social science experiments in *Nature* and *Science* between 2010 and 2015. *Nature Human Behaviour*, 2(9), pp.637-644. <https://doi.org/10.1038/s41562-018-0399-z>

- Capes-Davis, A. and Neve, R. (2016). Authentication: A Standard Problem or a Problem of Standards?. PLOS Biology, 14(6), p.e1002477. <https://doi.org/10.1371/journal.pbio.1002477>
- Chang, A. C., and Li, P. (2015). Is Economics Research Replicable? Sixty Published Papers from Thirteen Journals Say "Usually Not". Finance and Economics Discussion Series 2015-083. Washington: Board of Governors of the Federal Reserve System. <https://doi.org/10.17016/FEDS.2015.083>
- Collins, F. and Tabak, L. (2014). Policy: NIH plans to enhance reproducibility. Nature, 505(7485), pp.612-613. <https://doi.org/10.1038/505612a>
- Cook, B. (2014). A Call for Examining Replication and Bias in Special Education Research. Remedial and Special Education, 35(4), pp.233-246. <https://doi.org/10.1177/0741932514528995>
- Couchman, J. (2013). Peer Review and Reproducibility. Crisis or Time for Course Correction?. Journal of Histochemistry & Cytochemistry, 62(1), pp.9-10. <https://doi.org/10.1369/0022155413513462>
- Crane, H. (2017). Why "Redefining Statistical Significance" Will Not Improve Reproducibility and Could Make the Replication Crisis Worse. [online] arXiv.org. Available at: <https://arxiv.org/abs/1711.07801> [Accessed 10 Sep. 2019].
- Dafoe, A. (2013). Science Deserves Better: The Imperative to Share Complete Replication Files. PS: Political Science & Politics, 47(01), pp.60-66. <https://doi.org/10.1017/S104909651300173X>
- Donoho, D. (2010). An invitation to reproducible computational research. Biostatistics, 11(3), pp.385-388. <https://doi.org/10.1093/biostatistics/kxq028>
- Donoho, D., Maleki, A., Rahman, I., Shahram, M. and Stodden, V. (2009). Reproducible Research in Computational Harmonic Analysis. Computing in Science & Engineering, 11(1), pp.8-18. <https://doi.org/10.1109/MCSE.2009.15>
- Elman, C. and Kapiszewski, D. (2013). Data Access and Research Transparency in the Qualitative Tradition. PS: Political Science & Politics, 47(01), pp.43-47. <https://doi.org/10.1017/S1049096513001777>
- EPA Press Office (2018). EPA Administrator Pruitt Proposes Rule To Strengthen Science Used In EPA Regulations | US EPA. [online] US EPA. Available at: <https://www.epa.gov/newsreleases/epa-administrator-pruitt-proposes-rule-strengthen-science-used-epa-regulations> [Accessed 10 Sep. 2019].
- Estimating the reproducibility of psychological science. (2015). Science, 349(6251), pp.aac4716-aac4716. <https://doi.org/10.1126/science.aac4716>
- Evanschitzky, H. and Armstrong, J. (2013). Research with In-built replications: Comment and further suggestions for replication research. Journal of Business Research, 66(9), pp.1406-1408. <https://doi.org/10.1016/j.jbusres.2012.05.006>
- Fidler, F. & Gordon, A. (2013). Science is in a reproducibility crisis: How do we resolve it? [online] Phys.org. Available at: <https://phys.org/news/2013-09-science-crisis.html> [Accessed 10 Sep. 2019].

Firestein, S. (2016). Op-Ed: Why failure to replicate findings can actually be good for science. [online] Los Angeles Times. Available at: <http://www.latimes.com/opinion/op-ed/la-oe-0214-firestein-science-replication-failure-20160214-story.html> [Accessed 10 Sep. 2019].

FitzJohn, R., Pennell, M., Amy Zanne, A., & Cornwell, W. (2014). Reproducible research is still a challenge. [Blog] rOpenSci. Available at: <https://ropensci.org/blog/2014/06/09/reproducibility/>

Fomel, S. and Claerbout, J. (2009). Guest Editors' Introduction: Reproducible Research. *Computing in Science & Engineering*, 11(1), pp.5-7. <https://doi.org/10.1109/MCSE.2009.14>

Gezelter, J. (2015). Open Source and Open Data Should Be Standard Practices. *The Journal of Physical Chemistry Letters*, 6(7), pp.1168-1169. <https://doi.org/10.1021/acs.jpcclett.5b00285>

Goecks, J., Nekrutenko, A., Taylor, J. and Galaxy Team, T. (2010). Galaxy: a comprehensive approach for supporting accessible, reproducible, and transparent computational research in the life sciences. *Genome Biology*, 11(8), p.R86. <https://doi.org/10.1186/gb-2010-11-8-r86>

Goodman, S., Fanelli, D. and Ioannidis, J. (2016). What does research reproducibility mean?. *Science Translational Medicine*, 8(341), pp.341ps12-341ps12. <https://doi.org/10.1126/scitranslmed.aaf5027>

Greenland, S., Senn, S., Rothman, K., Carlin, J., Poole, C., Goodman, S. and Altman, D. (2016). Statistical tests, P values, confidence intervals, and power: a guide to misinterpretations. *European Journal of Epidemiology*, 31(4), pp.337-350. <https://doi.org/10.1007/s10654-016-0149-3>

Harris, R. (2017). *Rigor Mortis*. 1st ed. Basic Books, pp.1-279.

Hoffman, J. (2016). Archive computer code with raw data. *Nature*, 534(7607), pp.326-326. <https://doi.org/10.1038/534326d>

Höller, Y., Uhl, A., Bathke, A., Thomschewski, A., Butz, K., Nardone, R., Fell, J. and Trinka, E. (2017). Reliability of EEG Measures of Interaction: A Paradigm Shift Is Needed to Fight the Reproducibility Crisis. *Frontiers in Human Neuroscience*, 11. <https://doi.org/10.3389/fnhum.2017.00441>

Hossenfelder, S. (2017). Science needs reason to be trusted. *Nature Physics*, 13(4), pp.316-317. <https://doi.org/10.1038/nphys4079>

Ioannidis, J. (2005). Why Most Published Research Findings Are False. *PLoS Medicine*, 2(8), p.e124. <https://doi.org/10.1371/journal.pmed.0020124>

Ioannidis, J. (2008). Why Most Discovered True Associations Are Inflated. *Epidemiology*, 19(5), pp.640-648. <https://doi.org/10.1097/EDE.0b013e31818131e7>

Ioannidis, J. (2012). Why Science Is Not Necessarily Self-Correcting. *Perspectives on Psychological Science*, 7(6), pp.645-654. <https://doi.org/10.1177/1745691612464056>

Ioannidis, J. (2014). How to Make More Published Research True. *PLoS Medicine*, 11(10), p.e1001747. <https://doi.org/10.1371/journal.pmed.1001747>

lorns, E. and Chong, C. (2014). New forms of checks and balances are needed to improve research integrity. *F1000Research*, 3, p.119. <https://doi.org/10.12688/f1000research.3714.1>

Ishiyama, J. (2013). Replication, Research Transparency, and Journal Publications: Individualism, Community Models, and the Future of Replication Studies. *PS: Political Science & Politics*, 47(01), pp.78-83. <https://doi.org/10.1017/S1049096513001765>

Jarvis, M. and Williams, M. (2016). Irreproducibility in Preclinical Biomedical Research: Perceptions, Uncertainties, and Knowledge Gaps. *Trends in Pharmacological Sciences*, 37(4), pp.290-302. <https://doi.org/10.1016/j.tips.2015.12.001>

Jasny, B., Wigginton, N., McNutt, M., Bubela, T., Buck, S., Cook-Deegan, R., Gardner, T., Hanson, B., Hustad, C., Kiermer, V., Lazer, D., Lupia, A., Manrai, A., McConnell, L., Noonan, K., Phimister, E., Simon, B., Strandburg, K., Summers, Z. and Watts, D. (2017). Fostering reproducibility in industry-academia research. *Science*, 357(6353), pp.759-761. <https://doi.org/10.1126/science.aan4906>

Jones, A. and Kemp, A. (2016). Why is so much research dodgy? Blame the Research Excellence Framework. [online] *The Guardian*. Available at: <http://www.theguardian.com/higher-education-network/2016/oct/17/why-is-so-much-research-dodgy-blame-the-research-excellence-framework> [Accessed 10 Sep. 2019].

Jordan, P. (2014). Give young scientists a level playing field. *Science*, 346(6210), pp.711-711. <https://doi.org/10.1126/science.346.6210.711-a>

Kappenman, E. and Keil, A. (2016). Introduction to the special issue on recentring science: Replication, robustness, and reproducibility in psychophysiology. *Psychophysiology*, 54(1), pp.3-5. <https://doi.org/10.1111/psyp.12787>

Lawton, J. (2016). Reproducibility and replicability of science and thoracic surgery. *The Journal of Thoracic and Cardiovascular Surgery*, 152(6), pp.1489-1491. <https://doi.org/10.1016/j.jtcvs.2016.08.044>

Leek, J. and Peng, R. (2015). Opinion: Reproducible research can still be wrong: Adopting a prevention approach: Fig. 1. *Proceedings of the National Academy of Sciences*, 112(6), pp.1645-1646. <https://doi.org/10.1073/pnas.1421412111>

LeVeque, R. (2009). Python Tools for Reproducible Research on Hyperbolic Problems. *Computing in Science & Engineering*, 11(1), pp.19-27. <https://doi.org/10.1109/MCSE.2009.13>

Lupia, A. and Alter, G. (2013). Data Access and Research Transparency in the Quantitative Tradition. *PS: Political Science & Politics*, 47(01), pp.54-59. <https://doi.org/10.1017/S1049096513001728>

Makel, M. and Plucker, J. (2014). Facts Are More Important Than Novelty. *Educational Researcher*, 43(6), pp.304-316. <https://doi.org/10.3102/0013189X14545513>

Maniadis, Z., Tufano, F. and List, J. (2017). To Replicate or Not to Replicate? Exploring Reproducibility in Economics through the Lens of a Model and a Pilot Study. *The Economic Journal*, 127(605), pp.F209-F235. <https://doi.org/10.1111/eoj.12527>

McDermott, R. (2013). Research Transparency and Data Archiving for Experiments. *PS: Political Science & Politics*, 47(01), pp.67-71. <https://doi.org/10.1017/S1049096513001741>

McLellan, M., Brannon, P., Campa, A., Daley-Laursen, S., Kannan, G., Olsen, N., Taylor, R. and Thilmany, D. (2016). Reproducibility and Rigor in REE's Portfolio of Research. Technical Report. 9pp. <https://doi.org/10.13140/RG.2.2.24369.17761>

McNutt, M. (2014). Journals unite for reproducibility. *Science*, 346(6210), pp.679-679. <https://doi.org/10.1126/science.aaa1724>

McNutt, M. (2014). Reproducibility. *Science*, 343(6168), pp.229-229. <https://doi.org/10.1126/science.1250475>

Monroe, D. (2015). When data is not enough. *Communications of the ACM*, 58(12), pp.12-14. <https://doi.org/10.1145/2833138>

Morrell, K. and Lucas, J. (2012). The replication problem and its implications for policy studies. *Critical Policy Studies*, 6(2), pp.182-200. <https://doi.org/10.1080/19460171.2012.689738>

Mullane, K. and Williams, M. (2017). Enhancing reproducibility: Failures from Reproducibility Initiatives underline core challenges. *Biochemical Pharmacology*, 138, pp.7-18. <https://doi.org/10.1016/j.bcp.2017.04.008>

Mullane, K., Curtis, M. and Williams, M. (2018). Reproducibility in Biomedical Research. *Research in the Biomedical Sciences*, pp.1-66. <https://doi.org/10.1016/B978-0-12-804725-5.00001-X>

Munafò, M. (2017). Metascience: Reproducibility blues. *Nature*, 543(7647), pp.619-620. <https://doi.org/10.1038/543619a>

Munafò, M., Nosek, B., Bishop, D., Button, K., Chambers, C., Percie du Sert, N., Simonsohn, U., Wagenmakers, E., Ware, J. and Ioannidis, J. (2017). A manifesto for reproducible science. *Nature Human Behaviour*, 1(1). <https://doi.org/10.1038/s41562-016-0021>

National Academies of Sciences, Engineering, and Medicine (2016). *Statistical Challenges in Assessing and Fostering the Reproducibility of Scientific Results: Summary of a Workshop*. Washington, D.C.: National Academies Press. <https://doi.org/10.17226/21915>

National Academies of Sciences, Engineering, and Medicine (2017). *Fostering Integrity in Research*. Washington, DC: The National Academies Press. 326pp. ISBN 978-0-309-39125-2. <https://doi.org/10.17226/21896>

National Academies of Sciences, Engineering, and Medicine (2019). *Reproducibility and Replicability in Science*. Washington, DC: The National Academies Press. <https://doi.org/10.17226/25303>

National Institutes of Health (NIH) (2019). *Rigor and Reproducibility*. [online] National Institutes of Health (NIH). Available at: <https://www.nih.gov/research-training/rigor-reproducibility> [Accessed 10 Sep. 2019].

Neuroskeptic (2015). Psychology Should Aim For 100% Reproducibility [Blog] Neuroskeptic Available at: <http://blogs.discovermagazine.com/neuroskeptic/2015/09/07/100-percent-reproducibility/#.WqBYo2rwZaQ>. [Accessed 10 September 2019]

Pellizzari, E., Lohr, K., Blatecky, A. and Creel, D. (2017). Reproducibility. Research Triangle Park, NC: RTI Press. <https://doi.org/10.3768/rtipress.2017.bk.0020.1708>

Peng, R. (2015). The reproducibility crisis in science: A statistical counterattack. Significance, 12(3), pp.30-32. <https://doi.org/10.1111/j.1740-9713.2015.00827.x>

Peng, R. and Eckel, S. (2009). Distributed Reproducible Research Using Cached Computations. Computing in Science & Engineering, 11(1), pp.28-34. <https://doi.org/10.1109/MCSE.2009.6>

Peng, R., Dominici, F. and Zeger, S. (2006). Reproducible Epidemiologic Research. American Journal of Epidemiology, 163(9), pp.783-789. <https://doi.org/10.1093/aje/kwj093>

Plant, A., Becker, C., Hanisch, R., Boisvert, R., Possolo, A. and Elliott, J. (2018). How measurement science can improve confidence in research results. PLOS Biology, 16(4), p.e2004299. <https://doi.org/10.1371/journal.pbio.2004299>

Plant, A., Locascio, L., May, W. and Gallagher, P. (2014). Improved reproducibility by assuring confidence in measurements in biomedical research. Nature Methods, 11(9), pp.895-898. <https://doi.org/10.1038/nmeth.3076>

Pröll, S. and Rauber, A. (2014). A Scalable Framework for Dynamic Data Citation of Arbitrary Structured Data. Proceedings of 3rd International Conference on Data Management Technologies and Applications. <https://doi.org/10.5220/0004991802230230>

Rauber, A. and Pröll, S. (2015). Scalable Dynamic Data Citation - RDA-WG-DC Position Paper. [online] Research Data Alliance. Available at: <https://www.rd-alliance.org/groups/data-citation-wg/wiki/scalable-dynamic-data-citation-rda-wg-dc-position-paper.html> [Accessed 10 Sep. 2019].

Resnik, D. and Shamoo, A. (2016). Reproducibility and Research Integrity. Accountability in Research, 24(2), pp.116-123. <https://doi.org/10.1080/08989621.2016.1257387>

Richardson, D., Kwan, M., Alter, G. and McKendry, J. (2015). Replication of scientific research: addressing geoprivacy, confidentiality, and data sharing challenges in geospatial research. Annals of GIS, 21(2), pp.101-110. <https://doi.org/10.1080/19475683.2015.1027792>

Sarewitz, D. (2015). Reproducibility will not cure what ails science. Nature, 525(7568), pp.159-159. <https://doi.org/10.1038/525159a>

Sarewitz, D. (2015). Reproducibility will not cure what ails science. Nature, 525(7568), pp.159-159. <https://doi.org/10.1038/525159a>

Scannell, J. and Bosley, J. (2016). When Quality Beats Quantity: Decision Theory, Drug Discovery, and the Reproducibility Crisis. PLOS ONE, 11(2), p.e0147215. <https://doi.org/10.1371/journal.pone.0147215>

Schulz, J., Cookson, M. and Hausmann, L. (2016). The impact of fraudulent and irreproducible data to the translational research crisis - solutions and implementation. *Journal of Neurochemistry*, 139, pp.253-270. <https://doi.org/10.1111/jnc.13844>

Sené, M., Gilmore, I. and Janssen, J. (2017). Metrology is key to reproducing results. *Nature*, 547(7664), pp.397-399. <https://doi.org/10.1038/547397a>

Shaw, S. and D'Intino, J. (2017). Evidence-Based Practice and the Reproducibility Crisis in Psychology. *National Association of School Psychologists*, 45(5). Available at: https://www.mcgill.ca/connections-lab/files/connections-lab/ebp_nasp_cq.pdf [Accessed 10 Sep. 2019].

Spector, J., Johnson, T. and Young, P. (2014). An editorial on replication studies and scaling up efforts. *Educational Technology Research and Development*, 63(1), pp.1-4. <https://doi.org/10.1007/s11423-014-9364-3>

Stanford University (2019). Meta-Research Innovation Center at Stanford | METRICS. [online] [Metrics.stanford.edu](https://metrics.stanford.edu/). Available at: <https://metrics.stanford.edu/> [Accessed 10 Sep. 2019].

Steward, O. (2016). A Rhumba of "R's": Replication, Reproducibility, Rigor, Robustness: What Does a Failure to Replicate Mean?. *eneuro*, 3(4), pp.ENEURO.0072-16.2016. <https://doi.org/10.1523/ENEURO.0072-16.2016>

Stodden, V. (2009). The Legal Framework for Reproducible Scientific Research: Licensing and Copyright. *Computing in Science & Engineering*, 11(1), pp.35-40. <https://doi.org/10.1109/MCSE.2009.19>

Stodden, V. (2009). The Reproducible Research Standard: Reducing Legal Barriers to Scientific Knowledge and Innovation. Presented at: Communia Conference 2009: Global Science & Economics of Knowledge-Sharing Institutions, Torino, Italy, June 2009. [Powerpoint]

Stodden, V. (2014). 2014 : What Scientific Idea is Ready for Retirement? Reproducibility. [online] [Edge.org](https://www.edge.org/response-detail/25340) Available at: <https://www.edge.org/response-detail/25340>

Stodden, V. (2015). Reproducing Statistical Results. *Annual Review of Statistics and Its Application*, 2(1), pp.1-19. <https://doi.org/10.1146/annurev-statistics-010814-020127>

Stodden, V., Bailey, D., Borwein, J., LeVeque, R., Rider, W. and Stein, W. (2019). Setting the Default to Reproducible. [online] [Stodden.net](http://stodden.net/icerm_report.pdf). Available at: http://stodden.net/icerm_report.pdf [Accessed 10 Sep. 2019].

Stodden, V., Borwein, J. M., & Bailey, D. H. (2013). "Setting the Default to Reproducible" in *Computational Science Research*. *SIAM News*, 46(05).

Stodden, V., Guo, P. and Ma, Z. (2013). Toward Reproducible Computational Research: An Empirical Analysis of Data and Code Policy Adoption by Journals. *PLoS ONE*, 8(6), p.e67111. <https://doi.org/10.1371/journal.pone.0067111>

Stodden, V., Leisch, F. and Peng, R. (2014). *Implementing reproducible research*. Boca Raton, Florida: CRC Press.

Stodden, Victoria, Enabling Reproducible Research: Open Licensing for Scientific Innovation (March 3, 2009). International Journal of Communications Law and Policy, Forthcoming. Available at SSRN: <https://ssrn.com/abstract=1362040>

Stroebe, W. and Strack, F. (2014). The Alleged Crisis and the Illusion of Exact Replication. Perspectives on Psychological Science, 9(1), pp.59-71. <https://doi.org/10.1177/1745691613514450>

The National Academy of Sciences Arthur M. Sackler Colloquium on Reproducibility of Research: Issues and Proposed Remedies (2017). Reproducibility of Research: Issues and Proposed Remedies - YouTube. [online] YouTube. Available at: <http://www.youtube.com/playlist?list=PLGJm1x3XQeK0FeRdgKcyvyBH8TKUBtwFv> [Accessed 10 Sep. 2019].

The National Science Foundation (2016). Dear Colleague Letter: Encouraging Reproducibility in Computing and Communications Research. [online] Nsf.gov. Available at: <https://www.nsf.gov/pubs/2017/nsf17022/nsf17022.jsp> [Accessed 10 Sep. 2019].

Träger, U. (2019). Going beyond impact factors—reforming scientific publishing to value integrity. [online] Phys.org. Available at: <https://phys.org/news/2016-08-impact-factors-reforming-scientific-publishing.html> [Accessed 10 Sep. 2019].

Uncles, M. and Kwok, S. (2013). Designing research with in-built differentiated replication. Journal of Business Research, 66(9), pp.1398-1405. <https://doi.org/10.1016/j.jbusres.2012.05.005>

Vedantam, S. and Penman, M. (2016). When Great Minds Think Unalike: Inside Science's 'Replication Crisis'. [online] Npr.org. Available at: <https://www.npr.org/2016/05/24/477921050/when-great-minds-think-unlike-inside-sciences-replication-crisis> [Accessed 10 Sep. 2019].

Voelkl, B. and Würbel, H. (2016). Reproducibility Crisis: Are We Ignoring Reaction Norms?. Trends in Pharmacological Sciences, 37(7), pp.509-510. <https://doi.org/10.1016/j.tips.2016.05.003>

Warren, M. (2018). Make replication studies 'a normal and essential part of science,' Dutch science academy says. Science. <https://doi.org/10.1126/science.aat0224>

Weir, K. (2015). A reproducibility crisis? Monitor on Psychology: American Psychological Association, 46(9), p.39

Wiggins, A., Newman, G., Stevenson, R. and Crowston, K. (2011). Mechanisms for Data Quality and Validation in Citizen Science. 2011 IEEE Seventh International Conference on e-Science Workshops. <https://doi.org/10.1109/eScienceW.2011.27>

Würbel, H. (2017). More than 3Rs: the importance of scientific validity for harm-benefit analysis of animal research. Lab Animal, 46, pp.164–166 <https://doi.org/10.1038/labon.1220>

Yousefi, M. and Dougherty, E. (2012). Performance reproducibility index for classification. Bioinformatics, 28(21), pp.2824-2833. <https://doi.org/10.1093/bioinformatics/bts509>

Endnotes

¹ Erin Antognoli <https://orcid.org/0000-0003-0569-0808> is the Metadata Librarian and Data Curator, at the National Agricultural Library, United States Department of Agriculture, and can be reached by email: erin.antognoli@usda.gov

² Regina Avila <https://orcid.org/0000-0002-4340-2558> is the Digital Services Librarian, at the National Institute of Standards & Technology, and can be reached by email: regina.avila@nist.gov

³ Jonathan Sears <https://orcid.org/0000-0002-9045-713X> is the Data Scientist, LAC Group, on assignment at the National Agricultural Library, United States Department of Agriculture.

⁴ Leighton L. Christiansen <https://orcid.org/0000-0002-0543-4268> is the Data Curator at the National Transportation Library, in the Bureau of Transportation Statistics, at the United States Department of Transportation.

⁵ Jessica Tieman <https://orcid.org/0000-0002-9547-0448> is the Digital Preservation Librarian, at the U.S. Government Publishing Office.

⁶ Jacquelyn Hart <https://orcid.org/0000-0002-5408-6853> is the Acquisitions & Cataloging Librarian, Canada & Oceania Section, at the Library of Congress.

⁷ The name CENDI was derived from original membership from Departments of Commerce, Energy, NASA, and the Defense Information Managers group. Current membership includes several other federal agencies.