Nordic Perspectives on Open Science

From mess to method: A case study about designing a workshop on research data discovery

Benjamin Derksen (0000-0001-9144-7518), Asger Væring Larsen (0000-0002-9967-2637), Pernille Nailor, Andreas Helles Pedersen (0000-0003-4492-1955), Sacha Zurcher (0000-0002-1298-9202, corr. author)

Authors work at the Royal Danish Library (ROR: https://ror.org/03vvy0v27)

Nordic Perspectives on Open Science 2025, *https://doi.org/10.7557/11.8280*

Abstract

The dream of FAIR data is to have, where possible, research data open and accessible to review and reuse. The present reality complicates data's findability in a myriad of ways not least of which is the decentralised nature of research data repositories and indexes that characterise the data publishing landscape. In an effort to improve this situation, we, a network of library research support staff in Denmark, have developed a workshop designed to enhance the competencies of researchers to search for published data. This case study summarises the genesis of this workshop, its various iterations and potential avenues for future development.

Keywords

Dataset, Discovery, FAIR principles, Open Science, Research data, Researchers, Retrieval, Reuse, Searching, Workshop

Author declarations: This case study is an extended version of a paper presented at the Creating Knowledge conference in Helsinki, Finland, June 2024.

The authors declare that the artificial intelligence tool Copilot (GPT-4) was used to assist with grammar correction and language refinements in parts of the manuscript. Copilot was used to improve coherence and not for interpretation or generation of content. The authors take full responsibility for the integrity and accuracy of the manuscript.

Introduction

Searching for research data can be a difficult and confusing experience. The landscape of data publication is diffuse and decentralised, and often you end up empty-handed and, perhaps, a bit discouraged. In the past decade, research institutions and funding agencies have spurred a gradual implementation of Open Science, calling for furthered data sharing through an agenda encouraging the sharing of knowledge through principles of transparency, collaboration, and reproducibility in research processes.

As part of Open Science, there is an increasing focus on making research data FAIR. The acronym 'FAIR' equals four principles that ought to be the solution to the challenge of searching for research data: it stands for Findable, Accessible, Interoperable, and Reusable. An ideal FAIR dataset is thus well described in its metadata, as open as possible and as closed as necessary regarding access to the data files, usable by a variety of software packages, and clearly licenced (Wilkinson *et al.*, 2016). The incentive behind such a cultural shift is to address a number of challenges in academia at present, such as positive results bias (Sterling, 1959; Song *et al.*, 2000; Nakagawa *et al.*, 2022), a reproducibility crisis (Ioannidis, 2005; Open Science Collaboration, 2015), transparency in the production and analysis of research data (Vazire, 2018), and issues in the recognition of the manifold nature of contributions to research outputs (Cronin, 2001). So far, the implementation of the FAIR principles has been focusing on the publication of research data but less on the finding and retrieval of them.

The aim of this case study as well as our workshop is to balance the uneven focus between data sharing and data discovery. Noticing a lack of formalised recommendations in data discovery in this context, we, a network of library research support staff in Denmark, have created a hands-on workshop focusing on searching for research data. If we are to facilitate data reuse (the ”R” in FAIR), we have to ensure that data actually are being discovered (the ”F” in FAIR), and that researchers know how to make qualified searches for data. By teaching researchers how to find research data, it will, simultaneously, make them aware what it takes for them to create FAIR datasets. The overall goal is to make research more transparent and increase the reuse of research data – herewith avoiding unnecessary repetition of research and, possibly, enhancing the impact of research.

Reviewing the sparse literature on data search suggests that we are on to something. According to Kathleen Gregory et al. (2019, p. 419), “data retrieval is a nascent field”, as opposed to “information retrieval [which] has been extensively studied for over 60 years”. In a more recent publication, Gregory et al. (2020, p. 461) further write that “information about [data search or data retrieval] is buried within other discussions”. Similarly, focusing on the social sciences in a German context, Thomas Krämer et al. (2021, pp. 176–177) argue that “only a few studies exist that solely focus on users' data-seeking practices”, observing that “often this research is conducted in the context of data reuse”, data sharing, as well as “user studies for particular repositories”.

Data search necessitates and deserves more attention. This is the point of departure for the development of our workshop on how to search for existing research data. The case study starts out by accounting for some background and context of the initial steps taken. Subsequently, we elaborate on how we developed several search strategies to

grapple with data search and describe how different iterations of the workshop evolved. After that, the case study describes how we have furthered the course materials in alignment with Open Science and boosted availability and reusability through Open Educational Resources (OER) and a supplementary online resource. Finally, the case study discusses potential next steps and new implementation of the course materials.

Background and context

In the spring of 2022, a researcher from Roskilde University approached a member of our group with a question about finding a specific type of research data. Our group consists of teachers and research support staff employed by the Royal Danish Library and working at the University Libraries of Aarhus, Copenhagen and Roskilde. We come from a variety of fields, including the humanities, information science, physics, biology and geography, but we all work with research data management in general and the FAIR Principles more specifically.

The researcher in question was writing a research application and needed to know whether there were figures, statistics or other data available in relation to his specific topic. It turned out that this question was not so easy to answer. We needed to apply various strategies to find some data, which in the end only gave a partial answer to his question. This experience set in motion a series of activities that ultimately led to the development of a workshop for researchers in "Searching for research data". It is the process and the development of the workshop, as well as the thoughts behind it we would like to share in this case study.

Setting out the contours of a workshop

We decided at Roskilde University Library to include our colleagues from the other university libraries at the Royal Danish Library to see whether we, together, could find an answer to the researcher's question when we approached it in a systematic way. Our starting point was the EU Horizon2020 funded project FAIRsFAIR and the developed handbook: "How to be FAIR with your data. A teaching and training handbook for higher education institutions" (Engelhardt *et al.*, 2022). We drew principally from the lesson plan "Finding and Reusing Data".

The FAIR elements that this lesson deals with focuses on F (Findable), A (Accessible) and R (Reusable) (Engelhardt *et al.*, 2022, p. 145). As stated in the "FAIR Guiding Principles for scientific data management and stewardship" (Wilkinson *et al.*, 2016), the principal goal of FAIR is to support the reuse of data. In order to be reusable, data should, on a general level, respond to all of the FAIR principles and, in particular, with regards to the "R":

> R1. (Meta)data are richly described with a plurality of accurate and relevant attributes
>
> R1.1. (Meta)data are released with a clear and accessible data usage licence
>
> R1.2. (Meta)data are associated with detailed provenance
>
> R1.3. (Meta)data meet domain-relevant community standards

The lesson plan "Finding and Reusing Data" (Engelhardt *et al.*, 2022, p. 145) aims to take participants through the core requirements for searching for and reusing data, such as identifying relevant repositories, developing a broad strategy for searching, and assessing the terms for reuse. For the workshop with our colleagues, we included some of the reading material from the lesson plan and it inspired us to think about the structure and the core learning objectives for a course in the form of a workshop on the topic.

We refined the focus, putting more emphasis on the findability of data (the "F" in FAIR) as it is our view that researchers are the most qualified to assess data and its potential for reuse (the "R" in FAIR). It was also in this area (i.e. finding data) that we identified the greatest need, namely the inspiration and know-how to search for and identify relevant data.

According to Wilkinson et al. (2016) the "F" in FAIR requires that:

> F1. (Meta)data are assigned a globally unique and persistent identifier
>
> F2. Data are described with rich metadata (defined by R1)
>
> F3. Metadata clearly and explicitly include the identifier of the data it describes
>
> F4. (Meta)data are registered or indexed in a searchable resource

We set aside one day to carry out a workshop with the aim of it being a 'train the trainer' activity, meaning that we should be prepared to act as trainers/teachers for those who were interested in carrying out such a workshop for researchers in the future. Our teacher's guide, which we discuss in the section on Open Educational Resources, is written with this in mind. The overall premise for this initial workshop was to explore the process of finding research data and, in particular, reflect upon the different strategies that could be applied in this process. The workshop was thus about the process itself – of looking for data – and not so much whether we succeeded in finding the actual data.

The purpose was:

> 1) To gain practical experience in finding data to give us a better understanding of the process from a researcher perspective; and
>
> 2) To prepare us for delivering a workshop with similar content and exercises but aimed at researchers.

The programme for this initial workshop was structured around exercises to develop and test strategies for finding research data, and to determine what one is allowed to do with the data that is found. We concluded with a joint discussion about the overall framework of the workshop and how to tailor it as an event that is aimed at researchers, Ph.D. students, and, possibly, master students.

The point of departure was the request from the Roskilde University researcher, which prompted a joint discussion on how to make the researcher's question more specific. This included generating keywords that could be relevant, and which basic questions should be answered before looking for data (e.g. geographical location of data, time frame, research data or other kinds of data).

Subsequently, we formed three groups where each group would focus on one of the following areas to search for (research) data:

a) International/national (research) repositories (e.g. Zenodo, Harvard Dataverse);

b) (Inter)governmental organisations (e.g. EU, FAO, World Bank); and

c) Public/private organisations in Denmark (e.g. Danish Agriculture & Food Council, The Danish society for Nature Conservation).

All activities were noted down in a logbook, including the searches which didn't lead to a result.

Lastly, each group selected one (or more) datasets from the previous exercise and considered the following: quality and relevance of the dataset, ethical considerations in using the dataset, and legal aspects (i.e. terms and conditions, including how to cite). During this exercise, we noted down the entire process, including all our reflections and ideas that came along the way.

We concluded the day with a joint discussion on the pros and cons of finding and reusing (research) data, what it would take to organise a similar workshop for researchers, what should be included in the workshop, and in what format.

One of the conclusions of this day was that we needed to include searching in bibliographic databases and the "Data Citation Index" of Web of Science where it is possible to search for research data and data studies in a similar way as when searching for literature. We decided that the different search strategies could be divided into a direct and indirect way of searching for research data and the workshop could be organised accordingly.

The workshop "Searching for research data" – direct and indirect strategies

Out of this initial experiment, we have developed a three-hour workshop, which is roughly divided into three equal parts: a lecture, a search task, and a follow up reflection on the task. The lecture provides a discussion of research data themselves, and the difficulty in giving a concise, always meaningful definition. We inform participants that the process of searching for research data can be difficult, full of dead ends and that a degree of serendipity could be expected. Core to the content of the workshop is a number of search strategies.

These strategies can be broadly grouped into two: direct and indirect searching. Direct searching involves searching for data 'where they live' or locating them via a bibliographic database. A number of search strategies fall under this umbrella, such as searching in subject-relevant or multidisciplinary data repositories, searching in bibliographic databases, or searching via a search engine. When searching indirectly, on the other hand, you locate data affiliated with or cited in another document. Most commonly this would occur in a research article.

Acknowledging common practice among researchers, however, we also include a direct search strategy that is supplementary, intended to show how to search for research data directly in e.g. Google, or using Artificial Intelligence. In addition, this latter strategy

helps to mitigate the tendency to give up, when no results are obtained by a search elsewhere.

Note that these strategies, at present, do not recognise the social components – e.g. contacting colleagues – in researchers' current data searching practices. This is largely because we expect these practices to occur naturally when/as needed, and what we are offering is a series of standardised search strategies as an attempt to broaden the "information horizon", to borrow a term from Diane Sonnenwald (2005).

The second and largest portion of the workshop is a data-search activity, wherein the participants make use of the strategies and resources we present and try to find either relevant data suited to a synthetic research question, or something related to their own research area. Before the participants begin their search, we hand out a list of resources, along with the strategies and a logbook.

The participants are instructed to document and reflect on their searches, specifying keywords, place of search, and whether it is fruitful. This method is drawn from the best practices of literature searches and echoes its iterative nature. The method neatly doubles as a way to evaluate our teaching material, as we are able to use them as glimpses into their search experience.

The third and final section of the workshop is a follow up, where participants are invited to reflect on the experience, both in the form of an open discussion as well as using the digital learning tool Padlet. The aim of using the latter is to pull from a broader pool of participants to share their experiences and explain in their own words how they interact with the strategies we present.

At the same time, we are curious about any other ideas and/or perspectives they might have come across during the search, and which could be added to the workshop in the future. Broad questions to provoke discussion are: What were the challenges? What worked and what did not work? Any surprises? Did you develop any of your own tricks?

This concludes the workshop. Learning to search for data takes time and so the setup is by no means an exhaustive account. Rather, our intention is to get researchers' feet wet in the world of data searching such that they are capable of continuing on their own.

Method

Given the nature of data search as a new area in information searching, our methodology has been explorative. A good example of this is that our initial thoughts of systematising data search in a manner similar to systematic reviews, turned out to be not feasible. We shortly gave up on this prospect. It is difficult to state with certainty that all possible sources have been exhausted due to the fact that data publication is diffuse and decentralised. In other words, we have learned about the nature of our subject through the teaching, communication and development of the materials.

Our method has largely been iterative and inductive. Each in-person workshop has included a small group of observers, who have followed and recorded our participants' activities. We have also used the logbook tool, which we suggest as a best practice for data searching, as a data source. It helps to reveal what our participants thought was a

hit, and which points were most important to note down in each of the prescribed fields. We did not come to the development of this course with a formal methodology for pedagogy in mind. This is in large part due to the exploratory nature of the content mentioned above. We have thus inductively refined the learning goals for the workshop over time, as we have become comfortable teaching information searching, which lacks in precision.

In the evaluation of the two workshops at the Royal Danish Library, participants have flagged some issues, including the format of the synthetic cases they worked with, an insufficient focus on Artificial Intelligence as a search strategy and addressing questions of trust and credibility in what they found. At first, the research questions posed in the cases were constructed around datasets that we had already found. The presentation of each case was essentially structured as a treasure hunt, hinting at the existence of the dataset in question and encouraging the participants to look for this resource. This is contrary to the common experience of information searching generally, and so we did away with the treasure hunt element to simulate a real-world data search.

The FAIR-by-Design methodology (Filiposka, Mishev and Leister, 2024), developed as part of the Skills4EOSC project, has been applied retroactively, and is the preferred method for further development.

Experiences

Two workshops for colleagues in Aarhus and Copenhagen

Target group: Colleagues at the Royal Danish Library (former researchers and information specialists).

As part of a preliminary trial of the course, we conducted two workshops for colleagues from Aarhus, Copenhagen and Roskilde University Libraries. The aim was to test our materials and the underlying concepts of the workshop. In the exercise section of the workshop, we asked the participants to locate an existing dataset using one or more of the techniques introduced earlier. To assess their ability to apply a specific method effectively, we directed them to search for a dataset we already knew was available. This approach was chosen to mitigate the risk of participants failing to retrieve any data — not due to incorrect application of the techniques, but simply because no relevant dataset existed.

However, this strategy inadvertently led participants to perceive the exercise as a form of treasure hunt, fostering a competitive atmosphere. When they were unable to locate the dataset, some frustration followed. Moreover, this would not mirror the actual activity of searching for data. In response to this dynamic, we revised the exercise in subsequent test runs. Participants were then asked to apply the techniques to datasets whose existence was unknown to both instructors and learners. This adjustment ensured that any failure to retrieve data could not be definitively attributed to incorrect technique or the absence of data, thereby creating a more authentic and exploratory learning experience.

Despite the uncertainty, participants were still able to engage meaningfully with the search process. In most cases, with sufficient persistence, they succeeded in identifying relevant datasets, thereby gaining practical insight into the application of the techniques.

A workshop at the 2024 NEFUS[1] conference

Target group: Colleagues from Danish university libraries (former researchers, information specialists, data management specialists).

We next took an updated version of the course to the NEFUS working group, which is a network of research support service workers working in academic libraries. Here, we implemented a series of revisions to the activities that sought to mimic the actual activity of searching for data so as to avoid the treasure hunt experience of the previous iteration. The activities were structured around research questions for which no single dataset would unlock the answer but rather many could be said to contribute to a meaningful answer. This approach was well-received and has been the format by which we have moved forward with the course. Overall, the participants were warm and encouraging about the prospect of developing a loosely standardised set of recommendations, which could be offered to researchers who were looking to search for existing data on a given topic. They noted, however, that it is, in concurrence with our own experience, a slow and often frustrating enterprise.

A workshop for researchers at Roskilde University

In the spring of 2025, the workshop was presented for researchers at Roskilde University. Initially, there were four registered participants, but unfortunately, two cancelled at the last moment. We decided to carry out the workshop anyway as we wanted to use the opportunity to present it for the first time to researchers.

Although the researchers were from the same institute, their research areas were quite different, and neither of them had much experience in searching for research data as the main objective. To avoid frustration, we emphasised that it is not guaranteed that they would find useful datasets during the workshop, as it takes time to learn and apply the different strategies, provided that relevant research data exists.

In the evaluation it was mentioned that more time was needed to search for datasets and that it still was unknown how the application of the strategies would play out in practice. The handout with a summary of the strategies and examples of resources like e.g. repositories, bibliographic databases, and search engines to find datasets, proved to be very useful.

[1] NEFUS is the Danish Network for Research Support Services.

Open Educational Resources: Enhancing Course Accessibility and Reusability

In alignment with the principles of Open Science and FAIR, we chose to publish the course materials as an Open Educational Resource (OER) in addition to delivering the physical workshops. OERs are designed to promote the sharing and reuse of educational materials and rather than simply making the slides available online, we developed a comprehensive teaching package that includes slides, a logbook, handouts, and a guide for educators on how to implement the materials effectively in their own teaching.

To maximize reusability, the materials are released under a Creative Commons Attribution-NonCommercial licence (CC BY-NC), which allows non-commercial use with proper attribution. They include slides in Danish and English, an English teacher's guide, and practical handouts — all offered under the same licence.

To further support diverse educational contexts, we developed four distinct reuse scenarios tailored to varying durations and participant knowledge levels:

> **Scenario 1**: A four-hour workshop designed for participants with no prior knowledge.
>
> **Scenario 2**: A one-hour introductory workshop for participants with no prior knowledge.
>
> **Scenario 3**: A 45-minute session for participants with no prior knowledge, excluding the practical exercise component.
>
> **Scenario 4**: A 30-minute module intended as an add-on to existing courses in Research Data Management (RDM) and FAIR data, assuming prior knowledge of these topics and focusing specifically on data search strategies.

These four scenarios enable flexible integration of the workshop materials into a variety of educational settings, depending on the needs and requirements of the user. All materials are accessible via the LearningLib platform (Pedersen *et al.*, 2025).

Further developments of the OER

In August 2025 a network of library employees who run the OER platform LearningLib met to discuss the applicability of the FAIR-by-Design methodology. They investigated how this methodology would apply to LearningLib. Our learning objects were chosen as a case study to determine the FAIR-ness of the platform and our materials.

This workshop has provided us with feedback that we will work with during the autumn 2025 to create a version that is even more FAIR. This could include splitting up the slide deck into sections that correspond to the different scenarios described above and/or inserting the Creative Commons licence logo on all the available materials (slides, teacher's guide and handouts).

Some of the feedback from the participants applied to the platform itself rather than the learning materials. This means that an improvement in this respect would necessitate using another platform for publication, like e.g. the repository Zenodo.

Further dissemination of the workshop materials

As part of the ongoing development of our educational materials, we identified a need to provide participants with a supplementary resource that could support continued learning beyond the classroom setting. To address this, we created a dedicated website (*LibGuides*, 2025) that consolidates the strategies and materials used throughout the workshop. This initiative was informed by personal experience, recognising that revisiting learning materials, like PowerPoint slides, weeks or months after completion can be challenging for participants. Our aim was thus to facilitate easy access to a concise review of the content that in addition might inspire newcomers to take the workshop.

The website includes an introduction to data discovery and the FAIR principles, guidance on where and how to search for data, methods for evaluating retrieved data and their reuse conditions, and curated recommendations for data sources organised by research discipline. Additionally, a glossary is provided to explain some of the terms that are used throughout the website. As the workshop evolved, the website was continuously updated to reflect changes, and we intend to maintain this practice going forward.

Discussion

One of the challenges in developing the workshop has been deciding how general the content should be. That is, should a workshop for researchers be tailored to a discipline-specific audience, who have a particular understanding of what constitutes data? Or should we try to stick to the generalist approach adopted thus far? If we tailor the content to a discipline-specific audience, our worry is that the course will fail to gain momentum and attract a crowd that is big enough to justify the resources. An additional aspect is that researchers are already pressed for time and, as we have highlighted with the existing literature, it is not an activity that is on very many researchers' radars (yet). An option could be to create versions of the material that are targeting specific disciplines or data types. This could cater to the specific needs of smaller groups of researchers and match their needs 'here and now'.

Conclusions and next steps

As mentioned earlier our course materials were used as a case in a workshop focused on the Skills4EOSC FAIR-by-Design methodology. The methodology describes which steps authors can take in the different phases of development that will improve the FAIR-ness of learning materials. This workshop has given us a lot of feedback to work with, and we plan to spend the autumn of 2025 improving the FAIR-ness of the OER. This will include using a dedicated publication platform that issues DOIs or other PIDs, since this is a central element of the FAIR principles. We have shared the OER with a CC BY-NC licence, which invites reuse in a teaching context, but it also enables others to develop the material further, e.g. in a local setting. We thoroughly encourage this activity and would also like to invite people to suggest improvements of the content.

While the present workshop was developed to train researchers, an idea that we are contemplating now is to create a version of the material to be used to train students in

searching for data. This was tried to modest success as a reduced course offering given through the Copenhagen University Library's Datalab. The workshop was altered in some important ways. The need for data literacy was underscored, as the target group was no longer subject specialists with extensive expertise in producing/understanding data, but rather students, who are still in the process of learning the ins and outs of a discipline. In a similar vein, source criticism and scepticism were introduced as a means to combat the idea that data are unbiased. A small number of students attended, many of which overlapped with participants in the library's Python courses, suggesting a way forward may be to hook these courses together in some way. We are currently adapting slides and teacher's guide to the level of students. Both materials will be shared on the LearningLib platform when ready.

To increase awareness of our course materials among potential reusers, we are going to develop a communication strategy. This could include using social media, blogs, newsletters, and mailing lists. A blog post (Larsen, Asger Væring, 2025) has already been posted on the OpenAIRE blog and as the course material develops, we could post updates on this or other relevant venues.

Acknowledgements

The process of developing a workshop based on a practical case has been very motivating, and the collaboration between the three university libraries of the Royal Danish Library has worked really well in spite of the physical distance between us. We would like to thank our colleagues who were part of the development of the workshop: Anne Vils Møller, Anna Slyngborg, Falco Hüser, Johanne Rübner Hansen, and Kirsten Krogh Kruuse. We would also like to thank all the participants of the workshops and their feedback. Last but not least, we would like to thank the management of the Open Science initiative of the Royal Danish Library to allow us to allocate time to the task.

References

Cronin, B. (2001) 'Hyperauthorship: A postmodern perversion or evidence of a structural shift in scholarly communication practices?', *Journal of the American Society for Information Science and Technology*, 52(7), pp. 558–569. Available at: https://doi.org/10.1002/asi.1097.

Engelhardt, C. *et al.* (2022) *D7.4 How to be FAIR with your data. A teaching and training handbook for higher education institutions*. Zenodo. Available at: https://doi.org/10.5281/zenodo.6674301.

Filiposka, S., Mishev, A. and Leister, C. (2024) 'FAIR-by-Design Microlearning'. Zenodo. Available at: https://doi.org/10.5281/ZENODO.11548061.

Gregory, K. *et al.* (2019) 'Searching Data: A Review of Observational Data Retrieval Practices in Selected Disciplines', *Journal of the Association for Information Science and Technology*, 70(5), pp. 419–432. Available at: https://doi.org/10.1002/asi.24165.

Gregory, K.M. *et al.* (2020) 'Understanding data search as a socio-technical practice', *Journal of Information Science*, 46(4), pp. 459–475. Available at: https://doi.org/10.1177/0165551519837182.

Ioannidis, J.P.A. (2005) 'Why Most Published Research Findings Are False', *PLOS Medicine*, 2(8), p. e124. Available at: https://doi.org/10.1371/journal.pmed.0020124.

Krämer, T. *et al.* (2021) 'Data-Seeking Behaviour in the Social Sciences', *International Journal on Digital Libraries*, 22(2), pp. 175–195. Available at: https://doi.org/10.1007/s00799-021-00303-0.

Larsen, Asger Væring (2025) 'Why Searching for Research Data Deserves More Attention', *OpenAIRE*, 1 September. Available at: https://www.openaire.eu/community/blogs/why-searching-for-research-data-deserves-more-attention (Accessed: 7 November 2025).

LibGuides (2025) *Roskilde University Library*. Available at: https://libguides.ruc.dk/datasearch (Accessed: 10 November 2025).

Nakagawa, S. *et al.* (2022) 'Methods for testing publication bias in ecological and evolutionary meta-analyses', *Methods in Ecology and Evolution*, 13(1), pp. 4–21. Available at: https://doi.org/10.1111/2041-210X.13724.

Open Science Collaboration (2015) 'Estimating the reproducibility of psychological science', *Science*, 349(6251), p. aac4716. Available at: https://doi.org/10.1126/science.aac4716.

Pedersen, A.H. *et al.* (2025) *Searching for research data*, *LearningLib*. Available at: https://learninglib.org/2025/05/20/sogning-efter-forskningsdata-3/ (Accessed: 7 November 2025).

Song, F. *et al.* (2000) 'Publication and related biases: a review', *Health Technology Assessment*, 4(10). Available at: https://doi.org/10.3310/hta4100.

Sonnenwald, D.H. (2005) 'Information Horizons', in K.E. Fisher, S. Erdelez, and L. McKechnie (eds) *Theories of Information Behavior*. Medford, N.J: Information Today (ASIST monograph series.), pp. 191–197.

Sterling, T.D. (1959) 'Publication Decisions and Their Possible Effects on Inferences Drawn from Tests of Significance--Or Vice Versa', *Journal of the American Statistical Association*, 54(285), pp. 30–34. Available at: https://doi.org/10.2307/2282137.

Vazire, S. (2018) 'Implications of the Credibility Revolution for Productivity, Creativity, and Progress', *Perspectives on Psychological Science*, 13(4), pp. 411–417. Available at: https://doi.org/10.1177/1745691617751884.

Wilkinson, M. *et al.* (2016) 'Comment: The FAIR Guiding Principles for scientific data management and stewardship', *SCIENTIFIC DATA*, 3. Available at: https://doi.org/10.1038/sdata.2016.18.