Ger J Exerc Sport Res
https://doi.org/10.1007/s12662-025-01076-3
Received: 29 August 2024
Accepted: 21 October 2025

Tanja Eberhardt · Katja Keller · Hannah Zimmermann · Lars Schlenker · Klaus Bös · Claudia Niessner
Institute of Sports and Sports Science, Karlsruhe Institute of Technology, Karlsruhe, Germany

The potential of data pooling through open data infrastructure in motor performance research

A discussion using the example of MO|RE data

The rise of open science has revolutionised research practices, offering new opportunities for collaboration, transparency, and innovation. Alongside open access and open methodology, *open data* has become one of the core principles of the open science movement. It promotes not only the publication of research findings but also the accessibility of the underlying data by making them publicly available. This practice increases the value of the data by extending the research data lifecycle and ensuring their reuse for further innovative analyses. One particularly promising opportunity is *data pooling*, i.e. the merging of individual datasets from multiple studies or sources into a single, large-scale database to enable more powerful, robust, and comprehensive analyses (Adhikari et al., 2021).

To ensure transparency and good scientific practice, the implementation of open data should follow the FAIR principles, which stipulate that data should be findable, accessible, interoperable, and reusable (Wilkinson et al., 2016). These principles serve not only as guidelines for data management but also as foundational criteria for the development of supporting tools and infrastructure (Betancort Cabrera et al., 2020).

A holistic research data lifecycle as envisioned by the FAIR approach requires adequate infrastructure (Lopes et al., 2020; Rümpel, 2011). In recent years, various disciplines have made significant progress in building such infrastructure (e.g. Aschenbrenner & Neuroth, 2011; Huschka, Oeller, Notburga, & Wagner, 2011), supporting both the technical handling of open data and their integration into everyday research practice.

In the field of sports science, however, the potential of open data remains largely untapped (Krüger, Biniossek, Stocker, & Betz, 2023; Lohse, 2025). This is especially true for motor performance research, despite the fact that motor competence is a key factor for promoting physical activity across the lifespan (e.g. Ortega, Ruiz, Castillo, & Sjöström, 2008). While a wide range of test batteries and assessment tools for motor performance exist, no universally accepted gold standard has been established (Lopes et al., 2020). Nevertheless, several internationally distributed test items are commonly used, and the volume of motor performance data collected worldwide is substantial.

Pooling such data could generate sufficiently large samples for answering pressing research questions. Lopes et al. (2020) highlighted the potential of pooled data for examining global prevalence, long-term trends, and cross-cultural comparisons in motor competence, especially in children and adolescents. Moreover, pooled datasets would enable more robust analyses of determinants and correlates of motor performance across diverse populations.

Despite this potential, until recently, no infrastructure existed that fully met the specific needs of motor performance research. This is particularly noteworthy given that sports and exercise researchers have long emphasised the need for collaborative action and expressed strong interest in dedicated solutions—particularly when generating their own data in motor performance studies (Keller, Pleger, Niessner, Ueding, & Krüger, 2025; Kloe, Niessner, Woll, & Bös, 2019). Recently, however, efforts have been made to develop solutions that respond to these needs by aligning with open data principles and offering tailored functionalities for the field.

This discussion paper aims to critically reflect on the potential and limitations of implementing open data practices, such as data pooling, in motor performance research. Using the MO|RE data repository as an example, we explore current developments, reflect on key challenges and practical implications, and consider future directions in the field.

Idea and functions of MO|RE data

Funded by the German Research Foundation (DFG), the open-access MO|RE data repository was developed at the Institute of Sports and Sport Science at the Karlsruhe Institute of Technology (www.motor-research-data.de). The aim of this platform is to facilitate the sharing and pooling of motor performance data, enabling large-scale analyses and global monitoring (Lang et al., 2023). The

Published online: 25 November 2025

Fig. 1 ◀ Main functions of MO|RE data for data users and data providers

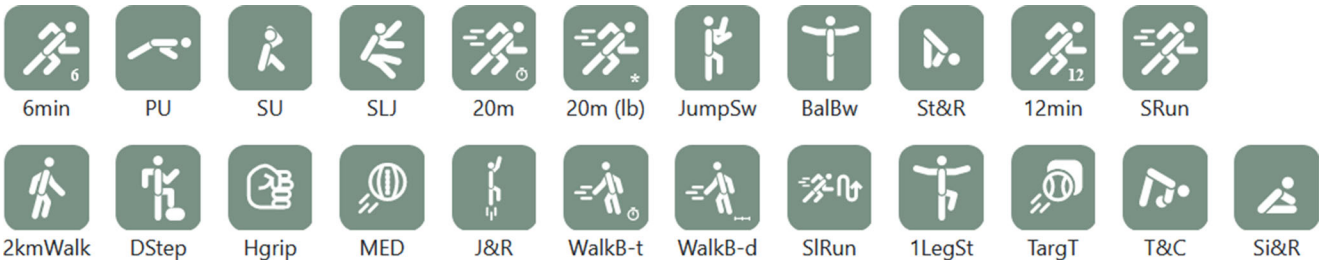

Fig. 2 ▲ Test items in MO|RE data. *6min* 6-minute run, *PU* push-ups, *SU* sit-ups, *SLJ* standing long jump, *20m* 20-m dash, *JumpSW* jumping sideways, *BalBw* balancing backwards, *St&R* stand and reach, *12min* Cooper test, *Srun* shuttle run, *2kmWalk* 2-km walking test, *DStep* Danish step test, *Hgrip* handgrip, *MED* medicine ball push, *J&R* jump and reach, *WalkB-t* walk back (time), *WalkB-d* walk back (distance), *SlRun* slalom run, *1LegSt* single leg stand, *TargT* target throw, *T&C* throw and catch, *Si&R* sit and reach

MO|RE data repository was designed to address the specific needs and requirements of motor performance data in order to overcome existing challenges that inhibit open data practices in sports science.

There are two main functions for users of MO|RE data: (1) *searching and exporting* and (2) *uploading and archiving* research data (**Fig. 1**). These functions are closely tied to the user's role: data users utilise existing datasets from MO|RE data for further research, while data providers contribute their own datasets to the repository. Researchers can, of course, have both roles—uploading their own data and using other people's data for future research. These two functions will be explained in more detail below, beginning with the perspective of data users.

The primary function of MO|RE data is to enable *searching and exporting of data*. This capability lays the foundation for locating and leveraging open datasets, opening the door to a wealth of possibilities. Data users can search for data on specific target groups or other characteristics relevant to motor performance research (e.g. age or socioeconomic status). Searches can be refined by combining keywords and selecting particular motor performance test items. MO|RE data displays matching search results and provides an overview with brief information about each dataset, including links to the respective metadata. In the next step, data users can export datasets and their associated metadata. This feature addresses the need to combine data from various sources into one large dataset, i.e. data pooling.

The second core function—*uploading and archiving data*—includes the following key components: 1) mapping and harmonisation, 2) entering metadata, 3) quality control, and 4) assignment of a persistent identifier. Each of these steps is crucial to ensuring that uploaded data can be reused meaningfully and effectively. The following sections describe these components in more detail.

The *mapping and harmonisation process* in MO|RE data addresses the variety of test items commonly used to assess motor performance in different testing settings. A shared vocabulary is therefore essential, and uploaded data must be standardised to ensure comparability and reusability. To this end, we identified the test items most commonly used in international motor performance research (**Fig. 2**). These test items are integrated into MO|RE data and can be mapped to users' own datasets during the upload process. Detailed manuals describe the standardisation procedures, specifying the required formats and measurement units. They also explain the meaning of individual data fields and attributes, promoting consistency across published datasets and enhancing their comparability and reusability.

Second, *metadata* are required for submitting datasets to MO|RE data. They provide essential information about each dataset and serve as the foundation for organising, understanding, and processing the data. A detailed metadata schema, based on the DataCite guidelines (DataCite Metadata Working Group, 2021), defines required standards. Metadata fields are categorised as mandatory (e.g. file or project title), recommended (e.g. keywords), or optional (e.g. research sponsor). This structure ensures consistent and interoperable documentation of data origin, structure, and usage. For motor performance data in particular, variations in test implementation and data collection methods are highly relevant for future interpretation and reuse.

Additionally, *quality control mechanisms* ensure that published data are of high quality and consistent. A two-stage review model is implemented, based

on the multidimensional framework by Wang and Strong (1996). The first stage involves an automated check using predefined filters to assess the internal quality of the data in terms of accuracy, consistency, and plausibility (Albrecht et al., 2016). For motor performance data, this applies especially to raw data and values within expected or possible values. Filters identify impossible and implausible values for gender-specific age ranges as well as duplicates. In the second stage, the editorial board conducts a manual peer review and acts as an independent quality control instance. As experts in the field, board members evaluate the contextual integrity of the dataset, including the completeness of metadata and the adequacy of anonymisation. They also assess whether the dataset can be accessed and understood based on a clear and detailed description of data collection, processing, and analysis.

Finally, once a dataset is approved by the editorial board members, it gets a *persistent identifier*. In MO|RE data, this is a Digital Object Identifier (DOI), issued via the internal infrastructure RADAR4KIT and registered with DataCite. The assignment of a globally unique and persistent DOI enables accurate citation of datasets, in line with best practices for scientific transparency. Moreover, DOIs can be included in researchers' publication records to increase the visibility of their scientific work.

Challenges

Despite the functionalities of MO|RE data described above, the adoption of open data infrastructure such as the MO|RE data repository remains limited within the wider research community. Several factors may contribute to this reluctance, including technical, legal, institutional, and methodological challenges.

First, a lack of familiarity with the infrastructure and uncertainty regarding technical procedures for preparing and submitting data may discourage researchers from using open data platforms. Navigating metadata standards, data formats, and harmonisation requirements can be perceived as time consuming and complex, particularly for those without prior experience in research data management.

Second, the legal and ethical requirements associated with open data in motor performance research are complex and multifaceted (Kreutzer & Lahmann, 2021). Assessing motor performance typically involves collection of sensitive personal and health data, often from children and adolescents. Ensuring compliance with data protection regulations—such as the General Data Protection Regulation in Europe and Germany—requires rigorous anonymisation and careful access management. Researchers must balance the need to protect participant privacy with the goal of enabling meaningful data reuse. In addition, copyright frameworks and data ownership rights can vary depending on how data are shared and are often difficult to interpret. In response, the Sensitive Data in Sports Science working group, part of the ad hoc Research Data Management committee of the German Society for Sports Science, is actively developing practical guidelines to support researchers in navigating these issues.

A third challenge relates to academic incentive structures. Concerns about data ownership, citation practices, and competitive disadvantages in publishing may discourage data sharing. Under increasing pressure to produce peer-reviewed articles, researchers may deprioritise the preparation and publication of open datasets—especially when such outputs are still perceived as less valuable for academic advancement than peer-reviewed articles (Hahnel, Smith, Schoenenberger, Scaplehorn, & Day, 2023). This issue is particularly relevant in motor performance research, where data collection often requires significant financial, human, and logistical resources. Preparing data for sharing in accordance with FAIR principles can be resource intensive and may exceed the available capacities of individual researchers or research teams. Nonetheless, the scientific community is beginning to recognise data publications as legitimate research outputs. Assigning Digital Object Identifiers (DOIs) and using standardised Creative Commons licences (e.g. CC BY-SA 4.0 or CC BY 4.0) helps to ensure proper attribution, facilitate reuse, and increase visibility—thus contributing to more transparent and rewarding data-sharing practices (McKiernan et al., 2016; Pampel & Elger, 2021).

Finally, beyond legal and institutional barriers, methodological considerations may also limit the implementation of open data practices. One such challenge is the need for strict standardisation of test items in order to enable data pooling

Abstract

Ger J Exerc Sport Res
https://doi.org/10.1007/s12662-025-01076-3

T. Eberhardt · K. Keller · H. Zimmermann · L. Schlenker · K. Bös · C. Niessner

The potential of data pooling through open data infrastructure in motor performance research. A discussion using the example of MO|RE data

Abstract

Open data is a core principle of open science that promotes transparency, collaboration, and innovation through data sharing and reuse. In motor performance research, it offers significant potential to overcome the limitations of small, isolated studies by enabling large-scale interdisciplinary analyses. Despite their benefits, open data practices remain underutilised in this field. This discussion paper examines the opportunities and challenges of implementing open data infrastructure in motor performance research, using the MO|RE data repository as an example. MO|RE supports FAIR principles through tailored functionalities like standardised metadata, harmonisation, and quality control. However, technical, legal, and cultural barriers still limit wider adoption. Recent developments in German sports science show progress toward fostering an open data culture. The paper argues that combining technical solutions with institutional support and cultural change is crucial to unlocking the full potential of open data for advancing motor performance research.

Keywords

Research data management · FAIR principles · Repository · Infrastructure development · Data sharing barriers

and ensure comparability. While standardisation is essential for data quality and interoperability, it can also create constraints that hinder methodological innovation. Researchers may feel limited by predefined protocols that do not align with their specific research objectives, target groups, or local conditions. Striking a balance between harmonisation and scientific flexibility remains an ongoing concern.

Current efforts and future directions

Overcoming the challenges outlined above requires not only technical solutions but also the development of community-based support structures and formal incentives to encourage broader participation in open data practices. Recent efforts in German sports science reflect growing momentum in this direction.

On the technical side, we continuously expand the functionalities of MO|RE data to meet evolving needs. Recent additions, such as the ability to distinguish between public and scientific use files, enhance data governance and access control. User-friendly interfaces, guided upload procedures, and integrated quality checks help to lower the threshold for researchers to share their data and comply with FAIR principles.

In addition, institutional and disciplinary support structures are becoming more established. In the context of motor performance, recent efforts within the German sports science community have made significant progress toward establishing open data practices and adequate infrastructure (Krüger et al., 2023). The integration of MO|RE data into the KonsortSWD consortium of the German National Research Data Infrastructure (NFDI) marks an important step toward embedding motor performance research within broader open science frameworks. In parallel, the German Society for Sports Science has developed practical guidance for managing sensitive data through its ad hoc Research Data Management committee (Krüger et al., 2025). These developments provide a strong foundation for embedding open data practices more firmly within the research culture.

However, future progress depends not only on infrastructure development but also on building trust, clarifying legal responsibilities, and creating academic incentives for data sharing. Establishing clear and broadly accepted guidelines for data citation, ownership, and privacy protection remains a key priority. At the same time, it is essential to foster a culture of recognition in which data publications are valued as legitimate scholarly contributions.

Beyond structural and cultural change, the analytical potential of pooled motor performance data is particularly promising. Pooling datasets from diverse sources makes it possible to overcome the limitations of small, isolated studies and supports more robust, large-scale analyses. Lopes et al. (2020) highlight a range of promising directions: pooled data can be used to examine population-level trends and global prevalence, explore cross-cultural differences, and investigate correlates and health-related associations of motor competence in children and adolescents. They also point to the potential for developing international assessment standards, policy audit tools, and targeted interventions.

Standardising data collection protocols and adopting interoperable formats are critical prerequisites for such large-scale analyses. Once in place, repositories like MO|RE data can facilitate longitudinal studies, meta-analyses, and the development of global normative benchmarks. For instance, initiatives like the European Fitness Landscape (Ortega et al., 2023) demonstrate how pooled datasets can be used to establish international reference values. Similarly, nationally coordinated initiatives, such as the Report Card on Physical Activity for Children and Adolescents (Demetriou et al., 2024) or the region-specific Fitness Barometer in Baden-Württemberg (Eberhardt, Bös, & Niessner, 2021), further illustrate how integrated infrastructure can support both scientific and public health goals.

Ultimately, motor performance research often reaches its limits when based on isolated datasets restricted to specific age groups or short timeframes. A comprehensive open data infrastructure—designed to meet the methodological and ethical demands of the field—can overcome these limitations. It enables researchers to conduct broader and deeper analyses, strengthens cross-study comparability, and enhances the evidence base needed to support physical activity and health across the lifespan.

Corresponding address

Tanja Eberhardt
Institute of Sports and Sports Science, Karlsruhe Institute of Technology
Engler-Bunte-Ring 15,
76131 Karlsruhe, Germany
tanja.eberhardt@kit.edu

Funding. This work was funded by the *Deutsche Forschungsgemeinschaft* (DFG; German Research Foundation)—BO 1551/3-1|; WO 724/3-1 | Project no. 432649291|

Funding. Open Access funding enabled and organized by Projekt DEAL.

Declarations

Conflict of interest. T. Eberhardt, K. Keller, H. Zimmermann, L. Schlenker, K. Bös and C. Niessner declare that they have no competing interests.

For this article no studies with human participants or animals were performed by any of the authors. All studies mentioned were in accordance with the ethical standards indicated in each case.

References

Adhikari, K., Patten, S. B., Patel, A. B., Premji, S., Tough, S., Letourneau, N., Giesbrecht, G., & Metcalfe, A. (2021). Data harmonization and data pooling from cohort studies: A practical approach for

data management. *International Journal of Population Data Science*. https://doi.org/10.23889/ijpds.v6i1.1680.

Albrecht, C., Bös, K., Schlenker, L., Tobias, R., van Wasen, M., Weiß, N., & Woll, A. (2016). *Handreichung Forschungsdatenmanagement in der Sportwissenschaft [Handout on research data management in sports science]*. Karlsruher Institute of Technology. https://doi.org/10.5445/IR/1000061538.

Aschenbrenner, A., & Neuroth, H. (2011). Forschungsdaten-Repositorien [Research data repositories]. In S. Büttner, H.-C. Hobohm & L. Müller (Eds.), *Handbuch Forschungsdatenmanagement* (pp. 101–114). BOCK+HERCHEN. https://doi.org/10.34678/OPUS4-208.

Betancort Cabrera, N., Bongartz, E. C., Dörrenbächer, N., Goebel, J., Kaluza, H., & Siegers, P. (2020). *White paper on implementing the FAIR principles for data in the social, behavioural, and economic sciences*. RatSWD Working Paper Series. https://doi.org/10.17620/02671.60.

DataCite Metadata Working Group (2021). *DataCite Metadata schema documentation for the publication and citation of research data and other research outputs v4.4*. DataCite. https://doi.org/10.14454/3W3Z-SA82.

Demetriou, Y., Beck, F., Sturm, D., Abu-Omar, K., Forberger, S., Hebestreit, A., Hohmann, A., Hülse, H., Kläber, M., Kobel, S., Köhler, K., König, S., Krug, S., Manz, K., Messing, S., Mutz, M., Niermann, C., Niessner, C., Schienkiewitz, A., & Reimers, A. K. (2024). Germany's 2022 Report Card on physical activity for children and adolescents. *German Journal of Exercise and Sport Research*, *54*, 260–275. https://doi.org/10.1007/s12662-024-00946-6.

Eberhardt, T., Bös, K., & Niessner, C. (2021). The fitness barometer: a best practice example for monitoring motor performance with pooled data collected from practitioners. *Frontiers in Public Health*, *9*, 720589. https://doi.org/10.3389/fpubh.2021.720589.

Hahnel, M., Smith, G., Schoenenberger, H., Scaplehorn, N., & Day, L. (2023). The state of open data 2023. *Digital Science*. https://doi.org/10.6084/M9.FIGSHARE.24428194.

Huschka, D., Oeller, C., Notburga, O., & Wagner, G. G. (2011). Datenmanagement und Data Sharing: Erfahrungen in den Sozial- und Geisteswissenschaften [Data management and data sharing: experiences in the social sciences and humanities]. In S. Büttner, H.-C. Hobohm & L. Müller (Eds.), *Handbuch Forschungsdatenmanagement* (pp. 36–48). BOCK+HERCHEN. https://doi.org/10.34678/OPUS4-208.

Keller, K., Pleger, A., Niessner, C., Ueding, E., & Krüger, M. (2025). The status quo of research data management in the German-speaking sports sciences—Results of an online pilot survey. *German Journal of Exercise and Sport Research*. https://doi.org/10.1007/s12662-025-01033-0.

Kloe, M., Niessner, C., Woll, A., & Bös, K. (2019). Open Data im sportwissenschaftlichen Anwendungsfeld motorischer Tests [Open data in the sports science application field of motor tests]. *German Journal of Exercise and Sport Research*, *49*(4), 503–513. https://doi.org/10.1007/s12662-019-00620-2.

Kreutzer, T., & Lahmann, H. C. (2021). *Rechtsfragen bei Open Science: Ein Leitfaden [Legal issues in Open Science: A guide]* (2nd edn.). Hamburg: University Press. https://doi.org/10.15460/HUP.211.

Krüger, M., Biniossek, C., Stocker, M., & Betz, D. (2023). Perspectives and potentials of open data for the sports sciences: The "what", the "why", and the "how.". *Zeitschrift für Sportpsychologie*, *30*(4), 167–176. https://doi.org/10.1026/1612-5010/a000405.

Krüger, M., Keller, K., Niessner, C., Ahrens, B., Bös, K., Eberhardt, T., Eckardt, B., Englert, C., Fleps, M., Holzbach, T., Hovemann, G., Koopmann, T., Memmert, D., Schlenker, L., Wallrodt, S., & Zimmermann, H. (2025). *Stellungnahme zum Umgang mit Forschungsdaten in der sportwissenschaftlichen Forschung [Statement on the handling of research data in sports science research]*. Deutsche Vereinigung für Sportwissenschaft. https://www.sportwissenschaft.de/fileadmin/pdf/download/2025_dvs-Stellungnahme_Forschungsdatenmanagement.pdf

Lang, J. J., Zhang, K., Agostinis-Sobrinho, C., Andersen, L. B., Basterfield, L., Berglind, D., Blain, D. O., Cadenas-Sanchez, C., Cameron, C., Carson, V., Colley, R. C., Csányi, T., Faigenbaum, A. D., García-Hermoso, A., Gomes, T. N. Q. F., Gribbon, A., Janssen, I., Jurak, G., Kaj, M., & Fraser, B. J. (2023). Top 10 international priorities for physical fitness research and surveillance among children and adolescents: a twin-panel delphi study. *Sports Medicine*, *53*(2), 549–564. https://doi.org/10.1007/s40279-022-01752-6.

Lohse, K. R. (2025). Taking steps toward open data in motor control, learning, and development. *Journal of Motor Learning and Development*, *13*(1), 1–8. https://doi.org/10.1123/jmld.2024-0081.

McKiernan, E. C., Bourne, P. E., Brown, C. T., Buck, S., Kenall, A., Lin, J., McDougall, D., Nosek, B. A., Ram, K., Soderberg, C. K., Spies, J. R., Thaney, K., Updegrove, A., Woo, K. H., & Yarkoni, T. (2016). How open science helps researchers succeed. *eLife*, *5*, e16800. https://doi.org/10.7554/eLife.16800.

opes, L., Santos, R., Coelho-E-Silva, M., Draper, C., Mota, J., Jidovtseff, B., Clark, C., Schmidt, M., Morgan, P., Duncan, M., O'Brien, W., Bentsen, P., D'Hondt, E., Houwen, S., Stratton, G., Martelaer, K., Scheuer, C., Herrmann, C., García-Hermoso, A., & Agostinis-Sobrinho, C. (2020). A narrative review of motor competence in children and adolescents: What we know and what we need to find out. *International Journal of Environmental Research and Public Health*, *18*(1), 18. https://doi.org/10.3390/ijerph18010018.

Ortega, F. B., Ruiz, J. R., Castillo, M. J., & Sjöström, M. (2008). Physical fitness in childhood and adolescence: a powerful marker of health. *International Journal of Obesity*, *32*(1), 1–11. https://doi.org/10.1038/sj.ijo.0803774.

Ortega, F. B., Leskošek, B., Blagus, R., Gil-Cosano, J. J., Mäestu, J., Tomkinson, G. R., Ruiz, J. R., Mäestu, E., Starc, G., Milanovic, I., Tammelin, T. H., Sorić, M., Scheuer, C., Carraro, A., Kaj, M., Csányi, T., Sardinha, L. B., Lenoir, M., Emeljanovas, A., & Jurak, G. (2023). European Fitness Landscape for children and adolescents: Updated reference values, fitness maps and country rankings based on nearly 8 million test results from 34 countries gathered by the FitBack network. *British Journal of Sports Medicine*, *57*(5), 299–310. https://doi.org/10.1136/bjsports-2022-106176.

Pampel, H., & Elger, K. (2021). Publikation und Zitierung von digitalen Forschungsdaten [Publication and citation of digital research data]. In M. Putnings, H. Neuroth & J. Neumann (Eds.), *Praxishandbuch Forschungsdatenmanagement* (pp. 521–536). De Gruyter.

Rümpel, S. (2011). Der Lebenszyklus von Forschungsdaten [The life cycle of research data]. In S. Büttner, H.-C. Hobohm & L. Müller (Eds.), *Handbuch Forschungsdatenmanagement* (pp. 25–34). BOCK+HERCHEN. https://doi.org/10.34678/OPUS4-208.

Wang, R. Y., & Strong, D. M. (1996). Beyond accuracy: what data quality means to data consumers. *Journal of Management Information Systems*, *12*(4), 5–33. https://doi.org/10.1080/07421222.1996.11518099.

Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J. J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J.-W., Da Silva, S. L. B., Bourne, P. E., Bouwman, J., Brookes, A. J., Clark, T., Crosas, M., Dillo, I., Dumon, O., Edmunds, S., Evelo, C. T., Finkers, R., & Mons, B. (2016). The FAIR guiding principles for scientific data management and stewardship. *Scientific Data*, *3*, 160018. https://doi.org/10.1038/sdata.2016.18.