Editorial

GJAE is Preparing for Future Developments in the Scientific System: New Policies Regarding Data and AI

Silke Hüttel[1] iD, Jens Rommel[2] iD, Ramona Teuber[3] iD, Martin Banse[4] iD and Carola Grebitus[5] iD

[1] Editor and University of Göttingen, Germany

[2] Editor and Swedish University of Agricultural Sciences, Uppsala, Sweden

[3] Editor and University of Gießen, Germany

[4] Managing Director of GEWISOLA and Thünen Institute Braunschweig, Germany

[5] Guest editor and Arizona State University, Tempe, USA

*Correspondence: Silke Hüttel, silke.huettel@uni-goettingen.de

As one of the few diamond open access journals in the agricultural and food economics domain, the German Journal of Agricultural Economics (GJAE) and its non-profit owner, the German Society for Economic and Social Sciences of Agriculture (GEWISOLA), are clearly committed to open science. Open science is broadly defined as scientific practices that enhance the openness, transparency, replicability, and reproducibility of research results, data, and methods (Finger et al., 2024). Open science relies on transparency through open data, codes and all materials supporting the research process to ensure that findings can be replicated and reproduced, i.e., consistent findings for the same research question are obtained using new data (replicability), and consistent findings are obtained using the same original data and methods (reproducibility).

As foundational aspects of the scientific system, replicability and reproducibility validate published results, and thus create confidence in reliability, robustness and generalisability of scientific findings (Ankel-Peters et al., 2023). Replicability and reproducibility also promote scientific progress by fostering the discovery of new evidence, expanding our understanding of economic problems and challenging existing theories or findings. These cornerstones ultimately contribute to the efficient functioning of the scientific system, i.e., shifting the boundaries of current knowledge and understanding. Such a functioning system is needed, particularly to address societal challenges for a sustainable future (Gomes, 2025). Replication and replicability are particularly important for evidence-based policy and management advice for a sustainable future, not least because evidence must often be assessed and synthesized across multiple studies. An open science system ensures straightforward assessments, replications and syntheses of scientific findings (Finger et al., 2023; Finger et al., 2024; Fraser et al., 2025; Gomes, 2025).

To ensure an efficient replication and replicability process, and ultimately contribute to an efficient functioning of the scientific system, replication studies should form a constant pillar in

publishing, and FAIR (findable, accessible, interoperable, and reusable) principles should guide research data management and publication (Wilkinson et al., 2016). These principles form the basis of GJAE's data policy and procedures for replication studies, therewith supporting a future-oriented open science system.

In light of this, all future articles published in the GJAE will be handled under our new data policy, which is added to the appendix of this editorial and is available on our website. Under this policy (i) a **data availability statement** is required for articles *published* in the GJAE. This is irrespective of whether the data can be shared and/or published. For example, if the work described in an article is based on data that are under an embargo or cannot be shared due to privacy or ethical issues this needs to be stated in the manuscript at the time of submission. The data availability statement should mention any restrictions regarding data access. If data is restricted, e.g., third-party data, legal or ethical constraints, this shall be explained in the data availability statement together with detailed information describing how other researchers can obtain the restricted data. This step offers utmost transparency for replicability and enables future research to follow up in a time- and resource-efficient manner when expanding validity of research findings. Our selection of data availability statements offers a broad range of options following widely accepted best-practice examples, ranging from open availability in GJAE's Data Archive or other repositories or public domains, to embargoed data. If no data were used or generated, a data availability statement is not applicable.

Pursuing the goal of replicability and reproducibility in an open science system means for GJAE that providing data "upon request" is not acceptable (Tedersoo et al., 2021). This is simply because the process of approval would not be open, and scientific evidence points to a high non-response rate of authors (Krawczyk, Reuben, 2012; Neilson, Premji, 2024). In addition, such practices would bear the risk of subjective or biased approvals of access to data and code (Acciai et al., 2023).

(ii) The GJAE encourages authors to make data available in the **GJAE Data Repository at the ZBW** – Leibniz Information Centre for Economics' Journal Data Archive *whenever* legally and ethically feasible (e.g., Jamali Jaghdani et al., 2024).

(iii) The GJAE also encourages authors to **make program code, software, etc. openly available using replication packages**. Not being able to share the data should not discourage authors from additional steps. For instance, we suggest authors to provide a replication package using synthetic data sets that mimic the original data for testing their code (Wimmer, Finger, 2023). A recent example is published by Isenhardt et al. (2025).

(iv) The GJAE explicitly **welcomes replication studies**. These include research investigating the reproducibility and robustness (using the original data and same methods, using the original data but different methods), the replicability (using the same question with the same or comparable methods but new data), and the generalizability (using new data and new methods) of findings (Ankel-Peters et al., 2025; Christensen et al., 2019; Finger et al., 2023). Replications should fall into GJAE's aims and scope, and should follow the recently published protocol of Ankel-Peters et al. (2025). That is, replications should typically start with a summary of the original study and an assessment of the computational reproducibility from the original code and data. This should be followed by a robustness reproduction, i.e., the main output of the original paper and robustness of the outcome over different assumptions, model specifications, estimation techniques, different sampling and data processing. Testing supportive analyses of the original paper could follow. When new data is collected, authors should critically reflect upon contextual factors that may affect the replication outcome. We expect authors to define the type of replication they are conducting, to motivate their replication appropriately, and to discuss generalizability and statistical power critically. Pre-analysis plans and pre-registration are particularly welcomed for replications. We welcome authors to reach out to the editors with questions on planned or completed replication studies. For additional general guidance on open science practices and replication, we recommend Christensen et al. (2019). In

the GJAE, all replication studies undergo the standard double-blind review process. Once the replication study is accepted for publication, the authors of the original study are offered the possibility to publish a comment on the replication.

Following the Transparency and Openness Promotion Factors[1] we strive for GJAE's data policy to be at level III in data citation, level II in data transparency and replication studies, and level I in analysis code, material and design transparency. Not yet implemented are any pre-registration or pre-analysis plan requirements.

While the data policy denotes an important step for the GJAE towards an open science system in the near future, the GJAE will need to continue investing time and efforts to prepare for a scientific system with Artificial Intelligence (AI). Relevant for the scientific discovery process are all AI systems. For instance, language processing in Large Language Models (LLMs) could be used for text and qualitative data analytics but also hypothesis generation from a large space of theories and support experimental designs, data collection or other forms of synthesizing. Other AI could be used for any data combination, visualization and machine learning-based analytics, prediction and synthesis (Wang et al., 2023; Bail, 2024). AI can accelerate the scientific research discovery process, as well as manuscript writing and publishing by means of considerable time and transaction cost savings. AI has already become an indispensable tool for researchers, for example, in economics (Korinek, 2023), social sciences (Bail, 2024) and in the applied agricultural and food economics domains (Hüttel, Hess, 2024; Fraser et al., 2025). More generally, productivity gains through AI are becoming increasingly important for labour productivity (Bick et al., 2024).

GJAE follows the Committee on Publication Ethics (COPE Council, 2024) position statement on Authorship and AI tools (2024): generative AI tools cannot take responsibility for the submitted work, and hence do not meet the requirements for authorship, such as the ability to declare competing interests or to agree to the license agreement. Any use of AI tools in the writing of a manuscript, image/graphic generation, or in the collection and analysis of data, must be disclosed transparently in the paper. AI used for literature research needs a verification of the sources. GJAE rejects submissions containing non-existent references. The use of AI tools in the review process is not permitted, as the confidentiality of the submissions cannot be guaranteed.

AI-aided research has also the potential to improve research studying human behaviour using surveys, experiments, and the like (Bail, 2024), and is therefore highly relevant for the transition towards more sustainable and resilient agri-food systems. For instance, experimental and survey-based testing of how behavioural factors or emotions affect acceptance of new policies need to rely on information schemes in the experimental setting. Generative AI may support the information schemes but can also support pre-testing or even replace human participants (Bail, 2024). While this may be particularly attractive for small and hard-to-recruit populations, such as, farmers, evidence as to how accurate predictions made by AI would be is mixed (Corrigan et al., 2025).

Many potential biases in the training data – including collected and published data, codes and articles (Bail, 2024; Hüttel, Hess, 2024) may impede productivity gains and exacerbate biases when using AI. Furthermore, data generated with AI based on experimental instructions alone – without any seed or training data – have not yielded evidence comparable to human subjects (Corrigan et al., 2025).

Another example where the use of AI can become challenging is the p-value debate: documented misinterpretation of statistical inference-based results, discrimination of results with

---

[1]   https://topfactor.org/summary

large standard errors (Amrhein et al., 2019) or equating low p-values with economic significance (Heckelei et al., 2023; Rommel, Weltin, 2021; Aurbacher et al., 2024) may then misinform future research.

Other challenges related to AI, such as research ethics, replication and replicability, environmental impacts and the proliferation of low-quality research are already largely documented (Marshall, Naff, 2024). This is why many calls for new AI ethics urge for explaining the use of AI in the research process, providing a critical reflection and disclosure of limitations – as is necessary to do for any other research method (Resnik, Hosseini, 2025). Future policies for journals and publishers in the scientific system will also have to address questions on how to ensure replicability of AI-aided data collection, ranging from synthesizing available data sets, aiding experimental and survey design, and AI-generated data. Key elements of open science-open source code-based AI algorithms gain importance and become, therefore, inevitable in an open data and open code infrastructure (Gundersen et al., 2025).

Against this background we see the new GJAE data and AI policies as a guidance for GJAE authors to conduct research that is in line with the requirements of open science. Ultimately, this will enable the scientific community to replicate and reproduce research published in the GJAE, which will increase trust in science and provide evidence-based policy recommendations that are more sound.

Acknowledgement

References

Acciai, C., Schneider, J.W., Nielsen, M.W. (2023): Estimating social bias in data sharing behaviours: an open science experiment. Scientific data 10 (1): 233. https://doi.org/10.1038/s41597-023-02129-8.

Amrhein, V., Greenland, S., McShane, B. (2019): Scientists rise up against statistical significance. Nature 567 (7748): 305-307. https://doi.org/10.1038/d41586-019-00857-9.

Ankel-Peters, J., Brodeur, A., Dreber, A., Johannesson, M., Neubauer, F., Rose, J. (2025): A protocol for structured robustness reproductions and replicability assessments. Q Open. https://doi.org/10.1093/qopen/qoaf004.

Ankel-Peters, J., Fiala, N., Neubauer, F. (2023): Do economists replicate? In: Journal of Economic Behavior & Organization 212: 219-232. https://doi.org/10.1016/j.jebo.2023.05.009.

Aurbacher, J., Bahrs, E., Banse, M., Hess S., Hirsch, S., Hüttel, S., Latacz-Lohmann, U., Mußhoff, O., Odening, M. Teuber, R. (2024): Comments on the p-Value Debate and Good Statistical Practice. German Journal of Agricultural Economics 73 (1). https://doi.org/10.52825/gjae.v73i1.988.

COPE Council (2024): COPE position - Authorship and AI - English. https://doi.org/10.24318/cCVRZBms.

Bail, C.A. (2024): Can Generative AI improve social science? Proceedings of the National Academy of Sciences of the United States of America 121 (21): e2314021121. https://doi.org/10.1073/pnas.2314021121.

Bick, A., Blandin, A. Deming, D. (2024): The Rapid Adoption of Generative AI. NBER Working Paper 32966. https://doi.org/10.3386/w32966.

Christensen, G., Freese, J., Miguel E. (2019): Transparent and Reproducible Social Science Research. How to Do Open Science. University of California Press, Berkely. https://doi.org/10.2307/j.ctvpb3xkg.

Corrigan, J.R., Grebitus, C., Rousu, M.C. (2025): On the Usefulness of Using Current LLMs for Experimental Auction Valuation. SSRN Electronic Journal. https://doi.org/10.2139/ssrn.5217193.

Finger, R., Grebitus, C., Henningsen, A. (2023): Replications in agricultural economics. Applied Economic Perspectives and Policy 45 (3): 1258-1274. https://doi.org/10.1002/aepp.13386.

Finger, R., Henningsen A., Höhler, J., Huber, R., Rommel, J. Grebitus, C. (2024): Open science in agricultural economics. Q Open. https://doi.org/10.1093/qopen/qoae029.

Fraser, I., Brooks, J., Bazzani, C. (2025): Open access and open science: some implications for the agricultural economics profession. Q Open. https://doi.org/10.1093/qopen/qoaf009.

Gomes, D.G E. (2025): How will we prepare for an uncertain future? The value of open data and code for unborn generations facing climate change. Proceedings. Biological sciences 292 (2040): 20241515. https://doi.org/10.1098/rspb.2024.1515.

Gundersen, O.E., Cappelen, O., Mølnå, M., Nilsen, N.G. (2025): The Unreasonable Effectiveness of Open Science in AI: A Replication Study. Proceedings of the AAAI Conference on Artificial Intelligence 39 (25): 26211-26219. https://doi.org/10.1609/aaai.v39i25.34818.

Heckelei, T., Hüttel, S., Odening, M., Rommel, J. (2023): The p-Value Debate and Statistical (Mal)practice - Implications for the Agricultural and Food Economics Community. German Journal of Agricultural Economics 72 (1): 47-67. https://doi.org/10.30430/gjae.2023.0231.

Hüttel, S., Hess, S. (2024): Are lessons being learnt from the replication crisis or will the revolution devour its children? Open Q science from the editor's perspective. Q Open. https://doi.org/10.1093/qopen/qoae019.

Isenhardt, L., Seifert, S., Wiltfang, T., Hüttel, S. (2025): Replication Data for: The Costs of Farmland Fragmentation: Evidence from Farmland Transactions in Eastern Germany. ZBW Journal Data Archive. https://doi.org/10.15456/gjae.2025168.0821751054.

Jamali Jaghdani, T., Glauben, T., Götz, L., Svanidze, M., Prehn, S. (2024): The Stability of the Global Wheat Trade in the Post-Soviet Space (replication data). ZBW Journal Data Archive. https://doi.org/10.15456/GJAE.2025002.0938236463.

Korinek, A. (2023): Generative AI for Economic Research: Use Cases and Implications for Economists. Journal of Economic Literature 61 (4): 1281-1317. https://doi.org/10.1257/jel.20231736.

Krawczyk, M., Reuben, E. (2012): (Un)available upon request: field experiment on researchers' willingness to share supplementary materials. Accountability in research 19 (3): 175-186. https://doi.org/10.1080/08989621.2012.678688.

Marshall, D.T., Naff, D.B. (2024): The Ethics of Using Artificial Intelligence in Qualitative Research. Journal of empirical research on human research ethics: JERHRE 19 (3): 92-102. https://doi.org/10.1177/15562646241262659.

Neilson, C.J., Premji, Z. (2024): A study of search strategy availability statements and sharing practices for systematic reviews: Ask and you might receive. Research synthesis methods 15 (3): 441-449. https://doi.org/10.1002/jrsm.1696.

Resnik, D.B., Hosseini, M. (2025): The ethics of using artificial intelligence in scientific research: new guidance needed for a new tool. AI and ethics 5 (2): 1499-1521. https://doi.org/10.1007/s43681-024-00493-8.

Rommel, J., Weltin, M. (2021): Is There a Cult of Statistical Significance in Agricultural Economics? Applied Economic Perspectives and Policy 43 (3): 1176-1191. https://doi.org/10.1002/aepp.13050.

Tedersoo, L., Küngas, R., … , Sepp, T. (2021): Data sharing practices and data availability upon request differ across scientific disciplines. Scientific data 8 (1): 192. https://doi.org/10.1038/s41597-021-00981-0.

Wang, H., Fu, T., … , Zitnik, M. (2023): Scientific discovery in the age of artificial intelligence. Nature 620 (7972): 47-60. https://doi.org/10.1038/s41586-023-06221-2.

Wilkinson, M.D., Dumontier, M., … , 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.

Wimmer, S., Finger, R. (2023): A note on synthetic data for replication purposes in agricultural economics. Journal of Agricultural Economics 74 (1): 316-323. https://doi.org/10.1111/1477-9552.12505.

Appendix

Data Policy GJAE

Replicability and reproducibility of research denote a cornerstone in the scientific system. They enable the validation of published results, and therethrough create confidence in the reliability, robustness and generalisability of scientific findings. Replicability and reproducibility also promote scientific progress by fostering the discovery of new evidence, expanding our understanding of economic problems and challenging existing theories or findings.

For these reasons, GJAE is committed to open science, replication research and the FAIR (findable, accessible, interoperable, and reusable) principles for research data (Wilkinson et al., 2016).[2] This means that (i) a data availability statement is a requirement for papers published in the GJAE, (ii) GJAE expects authors to make data available in GJAE Data Repository at the ZBW - Leibniz Information Centre for Economics' Journal Data Archive whenever legally and ethically feasible, and (iii) expects authors to make program code, software, etc. openly available using replication packages that are archived in the GJAE Data Repository at the ZBW - Leibniz Information Centre for Economics Journal Data Archive.

Data Availability Statement

Data availability statements confirm the presence or absence of shared data, are short and standardized to inform readers about the availability of the data used in the research process. For this purpose, the GJAE is offering a template, see at the end of the appendix.

The data availability statement should mention any restrictions on accessing the data. If the data is restricted (third-party data, legal or ethical constraints), this must be explained in the data availability statement. In addition, authors are asked to provide detailed information on how other researchers can obtain the restricted data in their readme-file (see below). In this case, we encourage authors to provide synthetic data sets that mimic the original data for testing the code.

Authors that have employed empirical methods and data will receive an invitation email by the GJAE Data Repository at the Journal Data Archive (send to [journaldata at zbw dot eu]) to deposit their data if legally and ethically feasible, and if these data not stored prior to the submission in a comparable repository. This will happen in the final step of the publication process.

It is generally not acceptable that data be provided "upon request" if the request must be approved by the authors themselves.[3]

Data Citation

In recognition of the relevance of data as an output of research effort, GJAE endorses the FORCE11 Data Citation Principles. Therefore, authors are asked to cite the data in the same way as article, book and web citations and include data citations as part of their reference list.

---

[2] Please see the following table for researcher responsibilities and a further explanation by the go fair initiative: https://www.snf.ch/SiteCollectionDocuments/FAIR_principles_translation_SNSF_logo.pdf and https://www.go-fair.org/fair-principles/ (last accessed July-09, 2024).

[3] Please see for further information Krawczyk, Reuben (2012) and Tedersoo et al. (2021).

When citing or making claims based on data, authors should refer to the data at the relevant place in the manuscript text and in addition provide a formal citation in the reference list. We recommend the format proposed by the Joint Declaration of Data Citation Principles:

[dataset] Authors; Year; Dataset title; Data repository or archive; Version (if any); Persistent identifier (e.g. DOI)

Replication Packages

We expect all our authors of accepted papers that contain empirical work, simulations, or experimental work to provide information about the data, programs, and other details of the computations sufficient to permit computational reproduction. Data, programs and other relevant material should be archived in the GJAE Data Repository at the ZBW - Leibniz Information Centre for Economics' Journal Data Archive, if legally and ethically possible.

Data and Variables

Data should be provided as part of the replication package unless the exceptions for restricted data apply (see above) or unless they can be fully reproduced from other accessible data within a reasonable time frame and with reasonable resources.

Each variable in the provided datasets should have a meaningful name or description (label), or authors may provide separate codebooks or similar metadata that describe the allowed values and their meaning. It is also possible to refer to publicly accessible documents that fulfil this purpose.

Readme-File

All replication packages must include a "Read me" file (clearly labelled) containing a list of all files included and guiding a user on the types of files and how to use them to do replication. If necessary, you will find detailed instructions on how to create a readme file here.

Program Code

Programs that produce computational results such as estimation, simulation, model solution must be included. Ideally, these programs reproduce all the computational exhibits in the paper with minimal human intervention.

A master script is encouraged. When no master script is included, please provide sufficient and precise step-by-step instructions in your readme file, allowing users to exactly reproduce the generated outputs with the least amount of effort.

Experimental Instructions and Surveys

Details regarding experimental procedures and instructions, and questionnaires are necessary to evaluate submissions of experimental research and research based on primary data collection. This information is also important to facilitate replicability and related work by subsequent researchers. Authors of experimental and survey-based papers must include detailed experimental instructions, along with screen shots for computerized experiments or record sheets for non-computerized ones, or the questionnaire. These instructions, questionnaires and related materials should be provided in the replication package. In cases where the instructions and materials were not presented to participants in English, both the original version and an English translation of materials shall be provided.

Further Information

- A step by step guide on how to make your data submission a success can be found here: https://aeadataeditor.github.io/aea-de-guidance/
- Detailed guidance on creating a readme file is available here

Standard Templates for the Data Availability Statement for Author Use

Below you will find a list of standard templates for the "Data Availability Statement". Please choose the correct statement(s) for the availability of your data.

Table 1. GJAE data availability statements

Availability of data	Template for data availability statement
Data openly available in the GJAE Data Archive	The data that support the findings of this study are openly available in the GJAE Data Archive. [DOI is added by the GJAE editorial office].
Data openly available in a public repository that issues datasets with DOIs	The data that support the findings of this study are openly available in [repository name e.g. "zenodo"] at http://doi.org/[doi], reference number [reference number].
Data openly available in a public repository that does not issue DOIs	The data that support the findings of this study are openly available in [repository name] at [URL], reference number [reference number].
Data derived from public domain resources	The data that support the findings of this study are available in [repository name] at [URL/DOI], reference number [reference number]. These data were derived from the following resources available in the public domain: [list resources and URLs]
Embargo on data due to commercial restrictions	The data that support the findings will be available in [repository name] at [URL / DOI link] following an embargo from the date of publication to allow for commercialization of research findings.
Data subject to third party restrictions	The data that support the findings of this study are available from [third party]. Restrictions apply to the availability of these data, which were used under license for this study. Data are available [at URL] with the permission of [third party].
Data sharing not applicable - no new data generated	Data sharing is not applicable to this article as no new data were created or analyzed in this study.
Data sharing not applicable - no new data generated, or the article describes entirely theoretical research	Data sharing not applicable to this article as no datasets were generated or analysed during the current study

Source: own data

When data is available and linked, authors will need to provide a citation of the data in their reference list.

Data Citation

[dataset] Authors (Year): Dataset title; Data repository or archive; Version (if any); Persistent identifier (e.g. DOI)