

Position paper of the Open Science Initiative at the Department of Statistics (OSIS)

Recently, the foundations of applied research have been called into question, particularly by much-read publications such as “Why most published research findings are false” [6]. As a result, a movement was created that critically re-evaluated the current scientific practices and led to a widespread shift in practices which are now summarized under the umbrella term Open Science: Reproducibility and replicability, Open Data and Open Access as well as general principles of good (empirical) scientific practice.

The Open Science discussion has also found general acceptance in research on methods and methodology. In statistics, for example, sharing code and data is now increasingly widespread. However, the technical implementation and aspects of reproducibility leave room for improvement and need to be systematized. The fact that the empirical parts of methodological research (for example comparison of methods, simulation and benchmark studies) are subject to the same issues that affect the substantive disciplines received little attention until recently. New publications [1-7] suggest many of the problems listed above apply here, too, declaring “[a] replication crisis in methodological research”, and adding to the increasing number of calls for better reproducibility, replicability, and more systematic evaluations in this area [5, 7].

This development indicates that research on methods and methodology in artificial intelligence, machine learning, and statistics is subject to the replication crisis. Therefore, there is a need for action to advance and promote Open Science in this domain. Statistics has a central role to play in the solution of these issues: By providing statistical methods, and critically evaluating their application in empirical research, statistics can contribute by raising awareness of Open Science issues in substantive research. In addition, the field is itself subject to the same problems. The central issues can be summarized, for example, in the “seven sins of methodological statistical research” [1]:

- Fishing expeditions/selective reporting
- Publication bias
- Lack of neutral comparison studies
- Lack of replication studies
- Poor design of comparison studies
- Lack of meta-analyses
- Lack of reporting guidelines

The Department of Statistics at the LMU, therefore, adopts the goal of actively contributing to the resolution of these issues, not least through its important role in teaching, communicating, and thereby promoting good scientific practice. We, therefore, aim to contribute to the following four central domains:

- Research on methods and methodology (e.g. concerning meta- and benchmark studies, development of Open Science guidelines, neutral evaluations)
- Teaching, consulting, and communication (e.g. integration of Open Science concepts in teaching, application in consulting, improvement of science communication)
- Organizational (e.g. recognition of meta- and neutral studies as parts of a dissertation, criteria for job descriptions)
- Technical (e.g. development of standards for sharing data and code, and support in their application)

- [1] Boulesteix, A.-L., Hoffmann, S., Charlton, A., & Seibold, H. (2020). A replication crisis in methodological research?. *Significance*, 17(5), 18-21.
- [2] Hutson, M. (2018). Artificial intelligence faces reproducibility crisis. *Science*, 359(6377), 725-726. DOI: [10.1126/science.359.6377.725](https://doi.org/10.1126/science.359.6377.725)
- [3] Morris, T. P., White, I. R., & Crowther, M. J. (2019). Using simulation studies to evaluate statistical methods. *Statistics in Medicine*, 38(11), 2074-2102.
- [4] Lucic, M., Kurach, K., Michalski, M., Gelly, S., & Bousquet, O. (2017). Are GANs created equal? a large-scale study. *arXiv preprint arXiv:1711.10337*.
- [5] Raff, E. (2019). A step toward quantifying independently reproducible machine learning research. *arXiv preprint arXiv:1909.06674*.
- [6] Ioannidis, JPA. (2005). Why Most Published Research Findings Are False. *PLoS Medicine*, 2(8):e124. DOI :[10.1371/journal.pmed.0020124](https://doi.org/10.1371/journal.pmed.0020124)
- [7] Boulesteix, A.-L., Lauer, S. & Eugster, MJA. (2013). "A plea for neutral comparison studies in computational sciences." *PloS ONE*, 8.4:e61562. DOI: [10.1371/journal.pone.0061562](https://doi.org/10.1371/journal.pone.0061562)