

Positionspapier zur Open Science Initiative am Institut für Statistik (OSIS)

Zu Beginn der 2000er Jahre wurden die Fundamente der angewandten Forschung in Frage gestellt, u.a. durch viel gelesene Publikationen wie "Why most published research findings are false" [6]. Insgesamt entstand hier eine Bewegung, die die gängige Wissenschaftspraxis in Frage stellte und zur Auseinandersetzung mit Themen führte, die unter dem Begriff *Open Science* zusammengefasst werden: Reproduzierbarkeit und Replizierbarkeit, Open Data und Open Access, sowie allgemein Prinzipien guter, empirischer wissenschaftlicher Praxis.

Auch in den method(olog)ischen Wissenschaften fand der Open Science Gedanke grundsätzlich Anklang. So ist in der Statistik heute das Teilen von Code und Daten teilweise verbreitet. Allerdings sind die technische Umsetzung und Aspekte der Reproduzierbarkeit oft verbesserungsfähig und bedürfen einer Systematisierung. Dass der empirische Teil method(olog)ischer Arbeit (z.B. Methodenvergleich, Simulationsstudien und Benchmark Experimente) aber denselben Problemen unterliegt wie in angewandten Forschungsdisziplinen, wurde bis vor kurzem hingegen eher wenig beachtet. Neuere Publikationen [1-7] legen jedoch nahe, dass auch hier viele der oben angesprochenen Probleme Bestand haben. So wurde unlängst "[a] replication crisis in methodological research" [1] ausgerufen und es mehren sich Forderungen nach besserer Reproduzierbarkeit, Replizierbarkeit und systematischer Forschung in diesem Feld [5, 7].

Dies liefert Hinweise, dass auch method(olog)ische Forschung in Künstlicher Intelligenz, Machine Learning und Statistik von der Replikationskrise betroffen ist. Es besteht daher Handlungsbedarf, Open Science in diesem Bereich voranzubringen. Dabei kommt der Statistik bei der Lösung eine zentrale Rolle zu. Durch die Bereitstellung statistischer Verfahren und die kritische Würdigung der Anwendung dieser Verfahren in der empirischen Forschung trägt sie einerseits aktiv dazu bei, auf Open Science Problematiken in den Substanzwissenschaften aufmerksam zu machen. Andererseits ist sie in ihrer methodologischen Ausrichtung selbst von dem Problem betroffen. Zentrale Problemfelder können z.B. in den "seven sins of methodological statistical research" [1] zusammengefasst werden:

- 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

Das Institut für Statistik der LMU setzt sich deshalb zum Ziel, mit der Open Science Initiative in Statistics (OSIS) einen aktiven Beitrag zur Verbesserung zu leisten. Nicht zuletzt umfasst dies auch ihre wichtige Rolle in der Lehre, Kommunikation und in der Verbreitung guter wissenschaftlicher Praxis. Es werden deshalb Aktivitäten in vier zentralen Bereichen angestrebt:

- Method(olog)ische Forschung (z.B. Forschung zu Meta- und Benchmarkstudien, Entwicklung von Open Science Guidelines, Neutrale Studien)
- Lehre, Consulting und Wissenschaftskommunikation (z.B. Einbinden von Open Science Konzepten in der Lehre, Anwendung in der Beratung, Verbesserung der Wissenschaftskommunikation)

- Organisatorisch (z.B. Anerkennung von Meta- und (neutralen) Benchmarkstudien als promotionsrelevante Leistungen, Kriterium bei Stellenausschreibungen)
- Technisch (z.B. Entwicklung von Standards für Data und Code Sharing, konkrete Umsetzungshilfe)

[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*, 2005; 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 (2013):e61562. [DOI: 10.1371/journal.pone.0061562](https://doi.org/10.1371/journal.pone.0061562)

