

Research Data Ethics

Andreas Bruns, PhD

University Hospital Heidelberg
National Center for Tumor Diseases (NCT)

The “Data-Driven Dilemma”

Humankind has been gathering and recording information for around 5,000 years. But with the fast development of digital technologies, we are now recording more information in one year than in all our preceding history.



Potential benefits: Big Data comes with great potential in virtually all areas of personal and social life (knowledge, medicine and healthcare, public transport, ...)

Ethical Risks: Big Data comes with diverse and hard-to-track ethical risks for individuals, groups, and societies.



Part 1: Ethical Risks in Research Data

Ethical Risks in Research Data

Case 1:

Marie consents to the use of her genetic data for medical research purposes. Analysis of her data reveals that she has a genetic predisposition for Huntington's disease.

Marie applies for a permanent teaching job at a university. Somehow information about her genetic predisposition reaches the potential employer. Behind closed doors, a decision is made against employing Marie and for employing a “more reliable” candidate instead.

Sensitive personal data can be associated with harms to the data subject.

Data protection is an ethical concern!

Ethical Risks in Research Data

Case 2:

Frank leads the development of his company's facial recognition software. The algorithm is trained with a large database of facial images and passes final tests with higher-than-expected success rates.

However, after release, users come to find that the software often fails to recognise black faces and misidentifies black faces at a high rate.

Research and training data can be associated with **group harms** and **social justice** concerns.

Ethical Risks in Research Data

Case 3:

Karl and Keeley work at the institute of post-colonial studies. Their institute has a large amount of anonymous textual data with discriminatory content about indigenous communities.

Karl thinks that, in the spirit of open science, they should make this data publicly available. Keeley is concerned that this might be wrong in virtue of being a distribution of hate speech and argues that access to the data should be restricted.

Also **anonymous** and **anonymised** data can be associated with ethical risks.

Part 2: Data Ethics Principles

Data Ethics Principles

‘As open as possible, as closed as necessary’

- 1 Protect data subject **autonomy** (privacy, personal choices)
- 2 Prevent **harms** to individuals and groups
- 3 Promote **social justice** (benefit sharing, account for biases)
- 4 Increase **social benefits** of research (data quality)

Part 3: Resources

Ethics Committees

Ethics vote: required, in principle, for *any* research involving human subjects or identifiable human biomaterial and data.

First point of contact: local ethics committee of the relevant faculty, e.g.,
<https://www.medizinische-fakultaet-hd.uni-heidelberg.de/fakultaet/kommissionen/ethikkommission>

Association of Medical Ethics Committees in Germany (AKEK)

https://www.akek.de/en/?option=com_content&view=featured&Itemid=101

Other:

- German Psychological Society: <https://www.dgps.de/die-dgps/kommissionen>
- German Educational Research Association:
<https://www.dgfe.de/dgfe-wir-ueber-uns/ethik-rat-ethikkodex>
- Ethics Committee of the Max Planck Society: <https://www.mpg.de/> (Intranet)
- Support through Data Access Committees (DACs)

Ethics Guidelines

National level:

- DFG Guidelines for Safeguarding Good Research Practice: <https://zenodo.org/records/6472827>
- DFG Checklist Handling of Research Data:
https://www.dfg.de/download/pdf/foerderung/grundlagen_dfg_foerderung/forschungsdaten/forschungsdaten_checkliste_en.pdf
- German Ethics Council Opinions:
<https://www.ethikrat.org/en/publications/kategorie/opinions/?cookieLevel=not-set&cHash=ff524cda9adeef368d7e26e3051ae854>

International level:

- WMA Declaration of Helsinki:
<https://www.wma.net/policies-post/wma-declaration-of-helsinki-ethical-principles-for-medical-research-involving-human-subjects/>
- GA4GH Framework for Responsible Data Sharing:
<https://www.ga4gh.org/product/framework-for-responsible-sharing-of-genomic-and-health-related-data/>

Other Helpful Resources

Open Data Institute: Data Ethics Canvas
Aims to help you identify and manage ethical issues in your data project:

<https://theodi.org/insights/tools/the-data-ethics-canvas-2021/>

Global Indigenous Data Alliance
CARE Principles:

<https://www.gida-global.org/care>

New NFDI Task Force Ethics
Will aim to identify and address ethical issues across NFDI consortia



Further Readings

- Ballantyne, A. (2018). Where is the human in the data? A guide to ethical data use. *Gigascience* 7(7):giy076. <https://pubmed.ncbi.nlm.nih.gov/29961892/>.
- Ferretti, A., Ienca, M., Hurst, S., Vayena, E. (2020). Big Data, Biomedical Research, and Ethics Review: New Challenges for IRBs. *Ethics Hum Res* 42(5). <https://pubmed.ncbi.nlm.nih.gov/32937036/>.
- Floridi, L., Taddeo, M. (2016). What is data ethics? *Philos Trans A Math Phys Eng Sci.* 374(2083):20160360. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5124072/>.
- Gans Combe, C. (2020). Research Ethics in Data. In: Iphofen (ed). *Handbook of Research Ethics and Scientific Integrity*. Springer. https://doi.org/10.1007/978-3-319-76040-7_13-1.
- O’Keefe, K., O’Brien, D. (2018). *Ethical Data and Information Management. Concepts, Tools and Methods*. Kogan Page.
- Rösch, H. (2021): “Forschungsethik und Forschungsdaten”. Putnings et al (eds). *Praxishandbuch Forschungsdatenmanagement*. De Gruyter Saur. <https://www.degruyter.com/document/doi/10.1515/9783110657807-006/html>.
- Wendelborn C, Anger M, Schickhardt C. (2023): What is data stewardship? Towards a comprehensive understanding. *J Biomed Inform.* 140:104337. <https://pubmed.ncbi.nlm.nih.gov/36935012/>.

Discussion Questions

1. What is the point of ethics where the law already provides us with more tangible normative guidance?
2. In your experience, (how) can legal and ethical requirements come apart in the planning and execution of research projects?
3. How to do ethical research where studies require the induction of (supposedly minor) harms, e.g. anxiety induction studies in psychology?

GHGA
THE
GERMAN
HUMAN
GENOME -
PHENOME
ARCHIVE

nfdi Nationale
Forschungsdaten
Infrastruktur

Thank You!