Guidelines to Research Data Management at
Leibniz-Zentrum Moderner Orient (ZMO)[1]

"As open as possible, as closed as needed."[2]

Academic research, particularly when it involves working with humans, must follow strict ethical considerations. First of all, there is the researcher's responsibility towards respondents and the people at the centre of the research. Researchers are obliged to safeguard their confidentiality, copyright and personality rights (including data protection) during the research process as well as afterwards. Second, academics must ensure transparency of the research process and account for their findings towards the scholarly communities and the wider public. While these two basic ethical considerations guide academic practice in the humanities and the social sciences, the practical consequences they seem to demand may not always be in alignment. The following guidelines on research data management at ZMO suggest a practical approach to good academic practice that navigates the sometimes conflicting demands arising from these two basic ethical responsibilities.

Introduction

Academic research has always been under public scrutiny and subject to occasional attacks, justified or unjustified, by fellow researchers, politicians, activists, and the broader public.[3] In recent years, though, the debate has intensified. Anti-intellectualism, "fake news" and distrust in experts are on the rise, partly stirred by populist political forces. Prominent cases of plagiarism and other violations of good scientific practice have contributed to intensified demands that research and research data be publicly accountable and accessible. Academics are pushed to share the raw data that led them to

---

their conclusions, the understandable rationale being that such transparency will enhance the verifiability and ultimately the authority of the results and the academics who produced them.

While European policymakers have decided to make transparency and accountability the pillars of the upcoming EU research programme "Horizon Europe" (2021-2027), research bodies (notably the German Research Foundation/DFG), broad scientific societies and associations such as Leibniz, and disciplinary associations have also been drawn into the debate. This process has led to new guidelines on how to collect, store, and share research data. The Guidelines on Safeguarding Good Scientific Practice, adopted by the Leibniz Association on 29 November 2018, now suggest to all Leibniz institutes and researchers to 'fully [document] all steps and results of an experiment or research study and [keep] protocols and research data securely' (§2.1b).[4] At ZMO, we have been doing this for a long time; still, for institutes such as ours, this call has important implications and raises a number of questions.

Disciplinary differences and responses

Crucially, there is the matter of disciplinary differences. Those working in the humanities and social sciences have faced vigorous public scrutiny and criticism, not least because their results seem to be more open to interpretation when compared with the natural sciences. Their practical applicability and, thus, usefulness are also more easily called into doubt. What is more, the very notion of what counts as "data" is specific to different disciplines. In the humanities and social sciences, research data cannot be narrowly understood as survey data, measured data, or intermediary data that can easily be distinguished from published results; rather, our data include annotated forms of representation and critical editions of historical sources; they also include digital representations and adaptations of cultural objects, which, in an ideal world, could and should be made available for much longer than data in other disciplines, as these research data form part of a society's cultural memory.[5]

In the case of quantitative data that can be aggregated and anonymized, the researchers' establishment of, or contribution to, public repositories can indeed be encouraged. At the same time, many of the professional associations related to the disciplines of history and social and cultural anthropology have been cautious about calls to make research data broadly available.[6] While

---

[4] See 'Leitlinie zur Sicherung guter wissenschaftlicher Praxis und zum Umgang mit Vorwürfen wissenschaftlichen Fehlverhaltens in der Leibniz-Gemeinschaft', p. 1. Passed on 27 November 2015. Available at: https://www.leibniz-gemeinschaft.de/fileadmin/user_upload/Bilder_und_Downloads/%C3%9Cber_uns/Gute_wissenschaftliche_Praxis/Leitlinie_gute_wissenschaftliche_Praxis__2018.pdf

[5] Verband der Historiker und Historikerinnen Deutschlands e.V., *Positionspapier des Verbandes der Historiker und Historikerinnen Deutschlands (VHD) zur Schaffung nationaler Forschungsdateninfrastrukturen*, p. 2.

[6] See *ibid.*; Martijn de Koning, Birgit Meyer, Annelies Moors and Peter Pels, 'Guidelines for anthropological research: Data management, ethics and integrity', *Ethnography* (2019), available at https://doi.org/10.1177/1466138119843312; Hansjörg Dilger 'Ethics, epistemology, and ethnography: the need for an anthropological debate on ethical review processes in Germany', *Sociologicus* 67, 2 (2017): 191-208; and Peter Pels, (ed.) 'Data management for anthropologists: the next phase in ethics governance?', *Social Anthropology* 26/3 (2018), 391-413.

verifiability is an important concern to both researchers and the public, it is not the only concern. These associations have rightly warned that there are large amounts of data that cannot or should not be made available at all: data whose ownership rights are unclear or whose owners do not want them shared further; data that have been offered on the condition of personal use only; data that infringe on the privacy of individuals and institutions; and data that are politically sensitive and could compromise the safety of interlocutors. <u>In these cases, the responsibility towards partners in the field weighs more heavily than calls for transparency.</u>

Anthropologists, in particular, view broad access to research data critically because their data are mostly of a context-bound character that is unsuitable for quantitative comparison and aggregation. In addition, knowledge is usually co-produced with interlocutors, which means that any transfer of access or ownership rights to research data to others is possible only with their consent. Based on relations of trust, interlocutors often share personal and sensitive material with the researchers, who are responsible for keeping personal and potentially sensitive materials protected and confidential. Anonymity as the default option and the non-disclosure of fieldwork data are preconditions for anthropological knowledge production. That said, the materials produced in ethnographic fieldwork are typically structured in such a way that anonymising them would not only be extremely time-consuming but also remove so much detail that the material loses much of its meaning. Using such material without the necessary personal and contextual knowledge runs the serious risk of misinterpretation and is therefore not in the interest of scientific knowledge. The sharing of research materials requires a careful process of contextualisation, anonymisation and negotiation with the research participants.

Historians, for their part, gather their material from state, institutional or private archives, or through interviews with contemporaries, whose experiences and memories are made public/ publicly accessible through oral history. While in the latter case, historians are expected to follow the same rules as anthropologists and respect the interests and privacy of their interviewees, archives have their own rules and regulations to which historians must strictly adhere. Researchers receive permission to search and collect evidence for a specific thematic purpose and are expected to inform the archives of the results of the research. Archives often make historians sign non-disclosure agreements. The contents of archival documents, let alone images of documents, cannot be made available without an archive's consent; and while such consent may be acquired for individual academic publications, it is not usually given for large, open-access databases.

Furthermore, the advancement of digital photography has enabled researchers to take large numbers of digital images of archival documents and store these for later analysis. Historians can now spend relatively short periods in archives, and identify and duplicate relevant documents without an immediate need for taking notes. Often, in fact, they no longer even produce field notes that could be made available; instead, the only data they store are images of the original documents, which they are not free to share because these images are the property of the archives. Historians of Asia, Africa and the Middle East, moreover, make wide use of personal archives owned by individuals, families or associations that do not keep official archives. The conditions that apply to these are similar to the ones that frame anthropological research.

Finally, archival research has a subjective dimension. Historians collect data in accordance with the pertinence principle, that is, based on their research topic and question (thematically). Archives are mostly organised according to the principle of provenance (based on the origins of the files). Hence, the data collected for any given project depends on the intentions and interests of the historian and may not be easily transferable to another historian's interests. 'Seeing data' or accessing it through varied platforms also means accessing it through the lens of the collecting historian.

Some forms of data sharing have been found useful by historians and anthropologists: in addition to more traditional forms of publication, there is a growing number of intermediary platforms that allow researchers to make data publicly accessible. Some examples include websites such as www.archive.org, which offers published and unpublished historical sources and research data. Archive.org is a 'non-profit digital library offering free universal access to books, movies & music, as well as 351 billion archived web pages'.[7] There is an increasing trend among researchers to share data or preliminary results in varied and innovative formats including blogs, databases, source collections etc.

Effects on research at ZMO

These questions have been intensively debated at ZMO, where the bulk of research is qualitative rather than quantitative. Most data used at the institute are interviews, archival records, field notes, photographs, and audio-visual media, which are full of personal information and heavily dependent on context. Such data must be protected and cannot be made freely accessible. Transparency and responsibility need to be weighed sensitively against each other to guarantee the quality of future research.

While researchers at ZMO are committed to complying with the conventions of their disciplines concerning research ethics and data management, many of them work in regions characterised by conflict or post-conflict and difficult socio-economic conditions. The personal safety of interlocutors can be at risk in some these regions, especially where they share sensitive information; and even where interviews are anonymized, detailed primary data can often be traced back to individuals. Making such data openly available would pose unnecessary threats to interlocutors; it cannot be considered responsible academic behavior and would run counter to disciplinary research ethics

ZMO encourages the sharing of suitable edited resources such as photographs, translations and documents, insofar as their sharing is in line with the aforementioned research ethics, privacy and copyright concerns. It also encourages the pre-publication of preliminary insights on blogs, websites and other public repositories. Examples of new open-access formats developed at ZMO include the HERA Sourcebook, which is the result of the cooperation project *Cultural Exchange in a Time of Global Conflict: Colonials, Neutrals, and Belligerents during the First World War*. It contains approximately 200 individual entries, each of them introducing one rare historical source (a text, image, sound recording, or object), along with interpretative essays.[8] Examples also include a blog

---

[7] See: https://archive.org/

developed by the project *Domestic Servants in Early Colonial India*, funded by a Starting Grant of the European Research Council (ERC, 2015-2018).[9] And they cover the DFG long-term project MIDA (*Modernes Indien in deutschen Archiven*), which started in 2014, with project teams working at the Centre for Modern Indian Studies/Göttingen, the Institute for Asian and African Studies/HU Berlin and ZMO. This latter project provides the public with a research portal (*Rechercheportal*) that consists of a database, an 'archival reflexicon' (that is, a collection of reflective essays) and thematic resources.[10] All parts of the research portal offer users access to data generated with the special purpose of making it available and visible for a broader international academic community. The database contains systematic, open access information on collections/holdings related to modern India in German archives.

The archival collections in the ZMO Library show how the institute deals with historical research data. A large part of the archival collection consists of personal papers of scholars whose areas of research were related to the regions and topics studied at ZMO. Among the files housed in these holdings are research data related to specific projects, systematically collected from German, European and non-European archives, as well as publications and lectures. These collections have been digitalised and catalogued, with online access to the catalogue lists. ZMO grants access to them on request.

Quantitative research plays a secondary role at ZMO. Over the last few years only one project has worked systematically worked with surveys, opinion polls and quantitative analyses, and the aggregated data of this study were put online in 2018. Local researchers across the Middle East and North Africa (the regional focus of this study), along with researchers further afield, now have access to the datasets through digital online repositories. Due to the design and nature of this large-scale quantitative research, individuals are not identifiable in this study and have not been put in peril.

In general, making such raw data available globally requires a highly developed IT infrastructure. It also demands a high standard of IT security at all nodes of such networks, particularly where sensitive information or personal data are concerned. Where possible and desirable, ZMO will facilitate access to comparable research data and encourage research partners to make their datasets available. That said, the advantages and disadvantages of putting raw and aggregated data online must be weighed carefully. ZMO acknowledges that these datasets constitute important academic accomplishments in their own right. Given its own limited use of quantitative digital data, however, ZMO will in the near future not be in a position to provide access to such data unless it has a potent partner.

---

[8] Project teams worked in London, Utrecht, Poznan and at ZMO. See: http://sourcebook.cegcproject.eu/. The sourcebook is hosted permanently by the In Flanders Fields Museum Ypres and is accessible via their website.

[9] https://servantspasts.wordpress.com/blog/

[10] https://www.projekt-mida.de/rechercheportal/

Guidelines

All of these deliberations have led the centre to draw up its own guidelines for research data management. <u>All researchers based at ZMO are contractually bound to abide by these guidelines</u>. At the same time, this set of rules and conventions has a dynamic character and may be changed or developed further if required or considered appropriate.

1. Researchers at the centre are asked to consider and address the question of research data management and possible future repositories at all project stages, that is, during preparation, implementation, and the dissemination of results.

2. Where quantitative data are concerned that can be aggregated and anonymized, the establishment of, or contribution to, public repositories is encouraged. Where the ownership, personal safety, and other aforementioned reasons for not sharing research data do <u>not</u> apply, researchers are also encouraged to share their data.

3. Open access formats support transparency, offer an additional step of quality control, and allow partners and interlocutors in the field to access ZMO research. As part of its publication strategy, the centre encourages the use of open-access publications, pre-publications, and databases (if there are no legal or ethical reasons that prevent the sharing of such data).

4. <u>Most data collected at ZMO cannot be publicly shared.</u> To allow for the event of queries regarding the scientific validity of a researcher's work, research data must be preserved (see below for details). Preservation, however, is not the same as public access and proactive sharing.

5. Researchers are obliged to preserve and store their research data after publication for a period of ten years. The obligation to preserve concerns raw and aggregated research data (in electronic or hard copy format), not research literature.

6. Researchers may take their research data with them upon leaving ZMO, provided they can always be contacted during the ten-year storage period. If they cannot (or do not want to) guarantee this, they must leave their data at the centre, either in electronic form or as hard copies.

7. Researchers are obliged to weigh carefully which of their research data is fit to be shared or stored in any of these formats. They may take their raw data with them upon leaving ZMO if they supply the centre with a written statement that confirms the data's highly personal and sensitive nature.

8. Our researchers generally work independently, starting from their PhD research. Thus, they own or co-own their research materials and output, and ZMO does not require them to waive ownership of their research materials and output to the centre or to supervisors.

9. All research materials can be made confidentially accessible for an audit in case concerns about the integrity of the research arise. If such concerns arise, the Ombudsperson and the advisory board will appoint three external auditors for the task.