

# EOSC Support Office Austria: Visions, needs and requirements for research data and practices

---

Wolfgang Kraus (University of Vienna), Susanne Blumesberger (University of Vienna)

This interview is also available for download: [<https://doi.org/10.5281/zenodo.7804887>]

In 2015 the vision of a federated system of infrastructures supporting research by providing an open multi-disciplinary environment to publish, find and re-use data, tools and services led to the launch of the [European Open Science Cloud](#) (EOSC). Against this background, bodies such as the [EOSC Association](#) on the European level and the [EOSC Support Office Austria](#) on the national one have been established.

Within this framework and since research has always been at the heart of EOSC, we are eliciting visions, needs and requirements for research data and practices from researchers who are located at public universities in Austria. Let's see what Wolfgang Kraus, cultural and social anthropologist, has to say!

**“Our data should be made available for a wider public, including the research subjects!”**

**Susanne Blumesberger (SB)** Thank you very much for doing this interview with me. What does your work currently focus on?

**Wolfgang Kraus (WK):** In my current research I am mainly working with data from earlier fieldwork I did in Morocco between 1983 and 2005. I'm now busy organizing and archiving my material and republishing some of it in new contexts. Beyond that, my main interest is data management and archiving. Together with colleagues I have established a digital archive for ethnographic data in order to make these materials available for various forms of reuse.

**SB:** What are the things that drive your work and why?

**WK:** I could answer this question in two ways. One is relating to my personal research, the

second to my work in the context of data archiving. The latter is based on the idea that ethnographic data have an ongoing interest over time. Since they are situated in time and space, ethnographic data are historical by nature and may become relevant and interesting in ways that are quite different from their original research context. In addition to that, the richness of ethnographic data often is not fully exploited in the original analysis, so they have a potential surplus value. This value extends beyond the scientific community. In our

*“Ethnographic data are historical by nature and may become relevant and interesting in ways quite different from their original research context.”*

archival work, we aim at making data available not only for other researchers but also for a wider public, and most importantly also for the research subjects, those who were the partners in the research. Obviously, privacy and data protection concerns must be respected when we do so.

**SB:** What kind of data are you working with?

**WK:** As an ethnographer, I am mainly working with primary data, but also with empirical material culled from publications by others. It is an ongoing practice in our field to gather empirical data for contextualization, comparison and other purposes from other researchers' publications. This made me think about the added value of having original data available rather than just trying to reconstruct empirical material from the analyses of researchers. In my own ethnographic research I have been producing various kind of data such as field-notes, recorded conversations, interviews and narrations relating to historical events, transcriptions of recorded material, photos and other visual data.

**SB:** What kind of digital tools and services do you use for your research?

**WK:** My ethnographic research has never been very software-intensive. Software was mainly

*“In ethnographic research the question of data quality, whether data can be relied on, arises the moment we begin collecting data.”*

involved in digital recordings and in organizing the material. I never worked with qualitative

data analysis software but organized my material in Microsoft Word and attached keywords to make the data searchable. For purposes of digital archiving we mostly use word processing, audio and imaging software. In the context of the Ethnographic Data Archive, the archival system has a central role, which is the PHAIDRA repository at the University of Vienna.

**SB:** Are you happy with how things are?

**WK:** For my own data processing needs, I'm fine with the tools at hand. Concerning data management and archiving, I think that PHAIDRA is a great solution and I enjoy working with it. A future functionality one might think of could be to process and work with data without downloading them. It would also be great to have a possibility for users to comment on data. We are reflecting a lot on collaboration with the research subjects. One buzzword in this context, and an important concept, is “indigenous metadata”. An option for users to comment on and contribute to metadata might be very interesting.

**SB:** What are the key steps in defining, collecting, analyzing, interpreting, storing data in order to know they are (and remain) high in quality?

**WK:** In ethnographic research the question of data quality, whether data can be relied on, arises the moment we begin collecting data. In the case of factual information, the obvious question is if the information is correct. However, most of the material we produce and use is not just factual information; it needs to be interpreted in the light of its specific context. So the question of data quality is not a simple one to answer. The critical question to ask is: is this data item really what I take it to be? Does it

justify the expectations I have in it? This includes much more than just the question of factual correctness. I'll give you an extreme example. In a book on music in Morocco published in the 1930s, the author described that musicians performing on a regional variety of the shawm could play this wind instrument for a long time without a break to breathe. He explained that young boys born into a family of musicians underwent an operation where a red-hot needle was used to pierce two holes into their neck through which they then could breathe while blowing their instrument. This is total nonsense. In fact they employed a technique that is now widely known as "circular breathing". How could such a blunder happen? In an ethnographic perspective, all knowledge is produced in a social relationship between researcher and research subject. As we don't know the research context, we can only speculate what happened. Maybe an interlocutor just did not know and made up an explanation in order not to disappoint the researcher. In this case however it is more likely that people were making fun of the researcher. So even if an information may be factually wrong it may serve to understand the social relationship that provided the context for it. Both for researchers and for repositories, this means that documenting the research context is as important as preserving the data. In evaluating ethnographic data, the context must always be taken into account and a critical spirit is always needed. We could do a separate interview just on analyzing and interpreting data! It is easier to answer the question concerning the processing and storing of data. In the Ethnographic Data Archive, we mainly work with analog data. Here the process of digitizing is critical and we're trying to get high quality data by paying a lot of attention to the process. We have done quite some experimenting with hardware and workflows to define best practices. This has also financial

dimensions; one cannot always have the best things money can buy, but one can get higher or lower quality digital copies depending on the processes. So we try to develop best practices for routine processes and define them in ways that can be reproduced by others. We aim at producing digital objects that remain technically relevant for a long period. I made a visit to the Royal Anthropological Institute in London which has an important and very extensive collection of documents and photos. They have an arrangement with Wiley for digitizing their material and making it accessible on a commercial basis these materials. What they told me is that the quality of the digital copies is not very reliable, so when they need a good quality copy for a publication they have to redo it. This is something we are trying to avoid.

SB: Are you reusing other researchers' data?

WK: Up to now I have not been reusing data in the usual sense of picking them from data archives and working with them. I don't know if I have missed a lot of interesting material. I may be mistaken but my impression is that in the French-speaking world some digital research practices, such as data archives, are less developed than they are for instance in the US. Therefore the historical material on Morocco that I am interested in is perhaps less available than data on American indigenous groups.

SB: Would you like to add something?

Perhaps this: something I tried out in my recent work with my old material, and that I found interesting and promising, is the possibility to make the data used in publications available via the Ethnographic Data Archive. I have published a few articles based on texts in the Berber language in the journal "Études et Documents Berbères" in which I interpreted and

contextualized narrations I had recorded in the field. I took the opportunity to publish the excerpts I used from my audio recordings as data objects in the repository and linked them in the text. This is an interesting option. In addition to reading my transcription and analysis people can listen to the original recordings. It is not something revolutionary and has been done before, but I think it is opening up interesting possibilities.

SB: Thank you!



*Wolfgang Kraus is associate professor (ao. Univ.-Prof.) at the Department of Social and Cultural Anthropology, University of Vienna, Austria. As Study Programme Director he is responsible for organising teaching and studies in social and cultural anthropology. He has done extensive field research on tribal organisation and historical knowledge in central Morocco. Among his main research interests are kinship, ethnography, visual anthropology, audio documentation, data management and research ethics. He is leading the Ethnographic Data Archive (eda) at the University Library Vienna.*

*“Both for researchers and for repositories, documenting the research context is as important as preserving the data.”*