

# Practices and representations of data sharing and data protection within research communities practicing or accompanying ethnography.

---

Marianne Noel (CNRS, LISIS) Florence Revelin (CNRS, LESC) Alix Levain (CNRS, AMURE) PARDOQ research group (AMURE) Camille Noûs (Cogitamus)

## Little is known about how the HSS are empirically dealing with the open data movement

- The STS literature has shown the diversity of opening-up practices; but little is known about how the HSS research communities are empirically dealing with the contemporary open data movement (ODM) in the "age of open science".
- In the context of the increasing digitization of practices and results, this movement, now transposed in European political agenda, has an impact on the very core of research activities and its support.
- Faced with new requirements of funding agencies, a growing number of studies from French academic institutions focus on data management at the laboratory and university levels, most often to support internal policies aiming at implementing data management tools.

## Questions

- How are qualitative research practitioners affected by the Open Data Movement?
- How do they perceive the injunctions and norms in favor of open data, and integrate opening-up in their research practices?
- What shifts does it bring about?

## PARDOQ, a qualitative research survey carried out in France on data sharing & data protection

- PARDOQ project, the first qualitative survey conducted in France on the practices and representations of data sharing and data protection within research communities practicing or accompanying the collection of ethnographic materials.
- 19 semi-directed interviews with members of research support networks and with practitioners addressed in the fall 2020, completed with a critical discourse analysis and participant observation.

## Profiles of the interviewees (12 practitioners)

Types of collectives	Statutes
Researchers in anthropology, having a strong interdisciplinary culture with biophysical sciences in a context of conservation and constitution of museum collections	<ul style="list-style-type: none"> <li>- Non-permanent researcher (PhD&lt;5 yrs), M</li> <li>- Assistant professeur, mid-career, F</li> <li>- Permanent researcher, mid-career, M</li> <li>- Professor, end of career, M</li> <li>- Permanent researcher, emeritus, F</li> </ul>
Researchers in sociology with a critical culture (STS)	<ul style="list-style-type: none"> <li>- Non-permanent researcher (PhD&gt;5 yrs), M</li> <li>- Permanent researcher, ECR, F</li> <li>- Permanent researcher, mid-career, M</li> </ul>
Researchers with a transdisciplinary culture in HSS	<ul style="list-style-type: none"> <li>- Assistant professor, ECR, F</li> <li>- Non-permanent researcher (PhD&gt;5 yrs), F</li> </ul>
Others	<ul style="list-style-type: none"> <li>- PhD in history, F</li> <li>- Non-permanent researcher in psychology (PhD&lt;5 yrs), F</li> </ul>

## PARDOQ, a qualitative research survey carried out in France on data sharing & data protection

PARDOQ Report (in French) is rich in empirical results, with many long verbatims.

<https://hal.archives-ouvertes.fr/hal-03238067v1>

Focus on:

- How do respondents name/design what they collect?  
Diversity of their relationships to data  
3 ideal types: “author”, “collector” and “collaborator” (p. 91)
- Passing on one's archives at the end of one's career is the most common (and most studied) form of sharing (the main channel through which ethnographers approach the topic).  
Ex: The costly (and salutary!) transition from a personal archive to a collection that can be appropriated by others

The ethnobotanist  
Claudine Friedberg in  
front of her collected  
material (*herbier*), oct.  
2009



## The notion of « data » is not obvious for ethnographers

- This survey highlights the way in which researchers conceive the empirical materials they collect: the notion of "data" is not obvious for them, neither on the epistemological level, nor a fortiori on the legal level. This is reflected, for example, in the fact that not all the practitioners interviewed use it to designate the information and empirical materials on which they base their analyses.
- One of the very concrete consequences of this “specific” relationship to “data” is that researchers practicing ethnography only exceptionally project themselves into the reuse of ethnographic data produced by others, and symmetrically do not think of their materials as appropriable ex post by others, i.e. detachable from their context of production and their authors.



## The notion of « data » is not obvious for ethnographers

The notion/category of “data” is generally not spontaneously mobilized by practitioners whose work is based in whole or in part on an ethnographic approach:

*“I talk about data but I don't like it. Others don't give it to me, they share it with me or they lend it to me, but I don't have a better term for it... My results? I don't use them that much, I'll say "in my interviews" or "in my observations".”* (Amandine, assistant professor, oct. 2020)

## Types of “sensitive” data mentioned by researchers at Univ. Montpellier Paul Valéry (Dillaerts et al., 2020, p. 24)

- “in the field of prisons”, respondent 1
- “privacy of the persons concerned or of third parties, possibly incrimination of third parties”, respondent 2
- “heritage data (archaeological: data are suppressed as they are studied, so the issue of their permanent recording is important)”, respondent 3
- “personal, medical data”, respondent 4
- “anonymity and drug practices, deviance and exclusion”, respondent 5
- “personal and private data (but I'm not sure if this corresponds to “sensitive data”)”, respondent 6
- “archives of the Intelligence Services”, respondent 7
- Etc.

## The notion of « data » is not obvious for researchers

- Ethnography remains a practice that is most often solitary, very artisanal, and increasingly practiced under precarious statutory and material conditions. Not all ethnographers have the same access to resources, the most important of which are working time and the skills to be acquired or to be used within a team.
- The history of ethnographic practice is marked by the slow construction of collective conventions, most often informal, which aim to regulate the risk for the interviewees both in terms of protection against the intrusion of third parties and against possible negligence by the ethnographers themselves.

The notion of « data » is not obvious for researchers

The Pardoq survey shows the extent to which practices of metadata documentation and opening-up of ethnographic data are numerous and sometimes quite long-standing (Strasser 2019).

Ex. of researchers who are part of a systematic approach to listing and classifying biological and cultural diversity.

## The notion of « data » is not obvious for researchers

- In this field as elsewhere, digital tools and infrastructures implies a transformation in the practices in which the researchers and the projects studied are fully involved and of which they are sometimes fervent promoters.
- However, these practices are not spontaneously conceived or linked to the legal framework and to the institutional offers of services in which they fit or could fit.

## A general movement towards opening-up?

- What can be observed is not a general movement towards opening-up, in which practitioners would identify a break or an acceleration from 2016 onwards.
- Rather, it is an ongoing exercise (where they compose and arbitrate between different choices) in which the forms of opening-up and sharing are not standardized. These arbitrages occur at different points in the research process, depending on considerations related to the nature of the data and the objectives of its collection and exposure.

## A general movement towards opening-up?


The practices of sharing ethnographic materials are thus, most often, far from the descriptive schemes of the data “research data lifecycle”.

**Une introduction à la gestion et au partage des données de la recherche**


**Le cycle de vie des données**

**Définition**  
Le cycle de vie des données de recherche (Research data lifecycle) décrit le processus d'utilisation des données de leur création à la publication et à leur réutilisation ultérieure.


Il existe plusieurs représentations du cycle de vie des données de la recherche.  
Le centre d'archives de données de la recherche, UK Data Archive, fournit un modèle de cycle de vie des données sur lequel les chercheurs peuvent s'appuyer pour gérer leurs données.

Ce modèle  définit les six étapes suivantes :

- **Création ou collecte** des données (creating data) ;
- **Traitement** des données (processing data) ;
- **Analyse** des données (analysing data) ;
- **Conservation** des données (preserving data) ;
- **Accès** aux données (giving access to data / data discovery) ;
- **Réutilisation** des données (reusing data).



*Cycle de vie des données de la recherche (adapté de UK Data Archive)*



Educational material produced by the INIST for research support communities

Any questions?  
[noel@ifris.org](mailto:noel@ifris.org)