

Collaborative work with data in the human rights field

Introduction	1
Advantages of collaborative work with data	2
Challenges of collaborative work with data	2
Recommendations for collaborative work with data	3
Consistent standards	3
Data dictionaries and thesauri	4
Training	5
Security and access control	5
Communication and engagement	6
Learnings	6

Introduction

The protection and promotion of human rights (HR) worldwide requires collaborative work across different professional fields, countries and institutions. Regarding data management, it is clear how working as a network allows us to have more complete datasets and perform comparative analysis. Collaboration, however, entails many challenges that could jeopardize the effectiveness of the projects if they are not tackled at the right time.

This document is intended to provide some tips for HR practitioners immersed in collaborative projects based on data. These recommendations come from HURIDOCs' experience with partners working in networks, and from international forums whose goal

was to develop global databases on topics related to HR.

Advantages of collaborative work with data

Collaborative work allows us to go where we couldn't go by ourselves. If addressed properly, collaboration can have a positive impact on the different stages of the data lifecycle —from data collection to sharing. Here we present some advantages of collaborative work with data:

- We enlarge our data collection range: we can collect data from places where we don't have staff or offices.
- Data collected and processed collaboratively provides us a bigger picture which allows us to get a comprehensive understanding of global and local dynamics.
- If we collect data collaboratively and systematically we increase our chances to perform comparative analysis across locations or contexts.
- Global data maximizes our advocacy impact.
- Multicultural and multidisciplinary teams offer more diverse and culturally sensitive approaches for work.

Challenges of collaborative work with data

All the advantages mentioned above depend greatly on our ability to anticipate risks and take action to ensure all the network members are well coordinated. Next, we present some of the most common challenges we will face when working collaboratively with data:

- Data collection is made at the local level, which can be distant to the decision-making level.
- Hierarchical networks have clear rules for the whole data management process, but less ability to react or adapt to local scenarios. Centralized decisions require top-down communication to reach all nodes. Decentralized networks are more likely to cope with local needs, but they make it more difficult to homogenize criteria.
- Rigid criteria (regarding data structures and definitions) facilitate data processing and analysis, but sometimes are not capable of covering all the scenarios. Flexible criteria can adapt to special circumstances, but heterogeneous data is more difficult

to analyze. To work collaboratively, it is essential to find balance between rigidity and flexibility.

- Combining different interests and purposes: every member of the network needs to understand their role, as well as the general purpose of the network and how it helps members' specific purposes.
- Information accountability: from the external point of view, the data belongs to the network, but inwards each one is responsible for the data they provide, in spite of the filters it passes before being recorded in the database.
- Working collaboratively with structured data requires many previous agreements. Planning tasks increase when networking.
- Culturally sensitive work is mandatory when collecting and managing data from different places. The same questions might not have the same interpretations everywhere. The language is only the first obstacle.
- More people working with the data always entail more concerns about sensitive information.
- Collective work demands commitment from every network member.

Recommendations for collaborative work with data

Based on the challenges previously identified, it is possible to issue some recommendations that can improve our work as a network.

Consistent standards

Having a common dataset involves unifying data collection and processing methodologies and tools. In the data collection process, the survey or interview form should be standardized and the personnel in charge of gathering the data must be trained to minimize personal subjectivity. The questions and possible answers should be presented in such a way that reduces ambiguity and additional interpretations.

We also need to define standards regarding data quality and verify that those are fulfilled before recording the data in the database. On this basis, we can identify some mandatory

fields to consider a record as “complete” or we can verify that there are no contradictions in the data. For further information on data quality and validation, [you can check our guide about database migration](#).

If we want our data to be reliable and comparable we need to be consistent when applying the predefined standards.

Data dictionaries and thesauri

If we are going to work as a network we need a common language to ensure we are referring to the same things with the same words. Each natural language has its own vocabulary and grammar rules, so we need to develop the same to work with our data.

Thesauri are lists of terms that we create to reduce ambiguity and ensure consistency in order to work with data. In the HR field, we need to categorize different aspects related to HR and be sure that everyone involved in the data management understands the terms the same way. For example, the perpetrator of a violent act can have different degrees of involvement ranging from carrying out the act to failing to prevent an unlawful act. Using a thesaurus with clear definitions about different degrees of involvement will greatly minimize the subjectivity of the categorization process. If you want to learn more about how to standardize vocabularies and define categories you can consult the [HURIDOCs’ microthesauri collection](#) or [this guide on how to build reliable indicators on HR issues](#).

At the same time, we need certain rules on how to record data, especially if there are several people involved in the process. The data dictionary provides information about every field in the database: naming conventions, constraints, relationships, format, etc. The following chart shows an example of a simple data dictionary which helps to record information about a victim:

Attribute name	Format	Required	Example
Victim’s name	(text) Surname, Name	Yes	<i>Doe, Jane</i>
Birth date	(date) DD/MM/YYYY	Yes	<i>19/02/1970</i>
Occupation	Select from: Occupation categories by HURIDOCs	No	<i>Legislators and senior officials</i>
Civil status	Select from: single, married, widowed, divorced, living with common-law spouse,	No	<i>Single</i>

	with partner, other.		
Description	Free text (max. 1000 characters)	No	...

The data dictionary contains the database’s metadata, so it is also called metadata repository. To learn more about how to build your data dictionary, you can check [this resource of the El Grito Toolkit created by Witness.org](#).

Training

The network is always as weak as the weakest link, so be sure that every member of the network understands the common project. Training is the best way to eliminate the misinterpretation mistakes that could jeopardize the reliability of our data.

Be sure that everyone understands what we need to collect and how to do it, as well as how every specific action will contribute to the general goal of the network. The thesauri and the data dictionary are very sensitive elements, so we should dedicate specific training sessions for them.

We should expect different levels of data and tech savviness among the network members, so training is essential to put everyone on the same step. Investing time and resources in training will save us a lot of headaches in the future.

Security and access control

As we have seen, security threats are always bigger if more people are involved in sensitive information management. If we are going to grant access to every network member, we need to define roles and translate them into different access levels and permissions to view, edit, add or delete records. These kinds of permissions are normally conditioned by the user access level and the information sensitivity degree.

As a network, we should also reflect on data ownership and we need to put crystal clear who owns each piece of data and what is the data use policy, that is to say, what members can do with data or what can be done on behalf of the network.

If we address these aspects consciously and collectively, we will avoid potential conflicts between members.

Communication and engagement

Communicating important decisions to the whole network is essential to keep everyone updated and engaged. Just as there are various roles and access levels within the network, there are different communication levels, since not everyone is interested in everything.

Having said that, it is clear that important agreements that affect the whole network must be communicated timely and widely in the network. This would be an example of a top-down approach in communication, but you should provide additional down-top and horizontal channels of communication within the network. In general, engagement is stronger in networks where each member feels equally important.

In this sense, key decisions should be duly documented and all the relevant documents for the network operations should remain accessible for everyone.

Learnings

In addition to the recommendations for collaborative work with data that we have already provided, we present some general learnings that we have extracted from our experience with partners and as network participants ourselves:

- Being part of a network entails big responsibilities. The quality of our data depends on the work of every node in the network.
- Communication should be as horizontal as possible to ensure that no one feels disconnected from the network.
- Each member needs to understand how their contribution will support the network and, on the other hand, how the network results will serve its particular goals.
- The network goal must be of interest for every partner.
- Credibility is a long-term collective goal, but it only takes one wrong step to lose it.

- Good coordination is the result of effective communication and capacity strengthening.