Cross-Community Metadata Sharing: Process Automatization for User Acceptance

In recent years the medical research sharing landscape has evolved, pushing FAIRification by introducing sophisticated platforms. However, the upload of (meta)data to them requires the data to be harmonized with the destinations structure, demanding a high time investment. This increases the threshold for technology adoption among the research community. Here, we introduce a concept for an automatic pipeline for metadata export, aiming to facilitate the process. We mapped odMLtables, a solution for electrophysiology data handling, to the structure of Local Data Hubs (LDHs), a data sharing platform rooted in FAIRDOM-SEEK, and developed by NFDI4Health. Our use case is microneurography data recorded for pain and itch research, with the metadata stored as domain-tailored odML files.

Mayra ELWES[a,1], Barbara NAMER[b], Alina TROGLIO[b], Toralf KIRSTEN[c], Oya BEYAN[a], and Ekaterina KUTAFINA[a]

[a] *Institute for Biomedical Informatics, Faculty of Medicine, University Hospital Cologne, University of Cologne, Cologne, Germany,* [b] *Research Group Neuroscience, IZKF, Department of Neurophysiology, RWTH Aachen University, Aachen, Germany,* [c] *Institute for Medical Informatics, Statistics, and Epidemiology, Leipzig University, Leipzig, Germany*

Introduction

In the past years, our team has established a FAIR local data infrastructure for microneurography data [1]. Microneurography is an electrophysiology method to study peripheral nerve activity in awake humans, critical in pain and itch research [2]. The Local Data Hubs (LDH) [3] were founded within the framework of NFDI4Health, an important German initiative for data infrastructure development, to establish a FAIR data sharing structure for healthcare and can be used efficiently for external data sharing. The LDH encapsulates metadata on different levels within a hierarchical structure. However, the process of transferring the data can be cumbersome, particularly under the condition of ongoing research and frequent needs for updates. Hence, there is a need for a solution, which can collaborate with the existing odMLtables [1, 4] GUI on a controlled and anonymized export to the LDH infrastructure.

Methods

Firstly, we did a requirements analysis of the workflow for the data owners and data seekers, via an interview with two experts. Next, we reviewed the data model at the core of the LDH. Based on the results, an initial linear workflow for metadata sharing and searching via the LDH was conceptualized, presented as an activity diagram in Figure 1.

Results

The resulting requirements put a strong emphasis on integrating the workflow for the data owner as in the already established odMLtables GUI and giving maximal control on choosing the parts of the information for sharing. For the seekers, a structured overview and data viewing within the LDH interface were important. Based on this feedback, we proposed a concept for sharing microneurography metadata, where the user can perform all steps comfortably from the GUI. The extension will allow:

a) to choose the appropriate access type,
b) to fill in the structured descriptive information,
c) to select the odML fields to publish for the defined recipient group
d) to create a overview table from the chosen odML subset
e) to upload to the LDH.

We came up with a mapping of the odML and to the LDH concepts, presented in Figure 2. On the side of LDH, the descriptive information will be filled in for the *Project*, as it is the highest level node grouping all *DataFiles* obtained in the corresponding Project. The data owner may upload the selected fields of the odML tables using the existing *DataFile* structure of the LDHs. The upload of an anonymized raw signal data in NIX format [5] can be added as an additional feature, to a heavy data sharing platform. To facilitate the search, text will be generated from the odML *overview table* and published in the freetext fields of the related *Project* and *DataFile*. Also, the odML overview table, it will be published as a *DataFile*, and will be viewable and searchable in the LDHs web interface.

Figure 1
Concept on how the metadata of the source odML file is represented in the LDH concepts and which information should be streamed from the odMLtables GUI to the LDH.

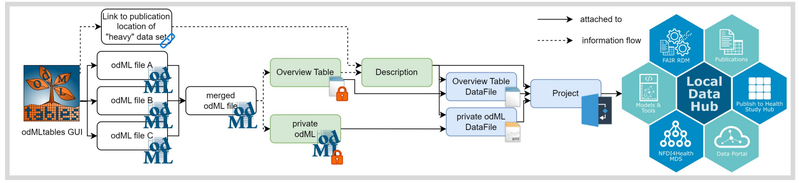
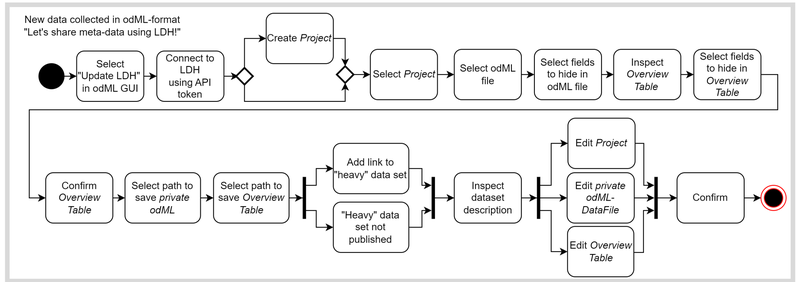


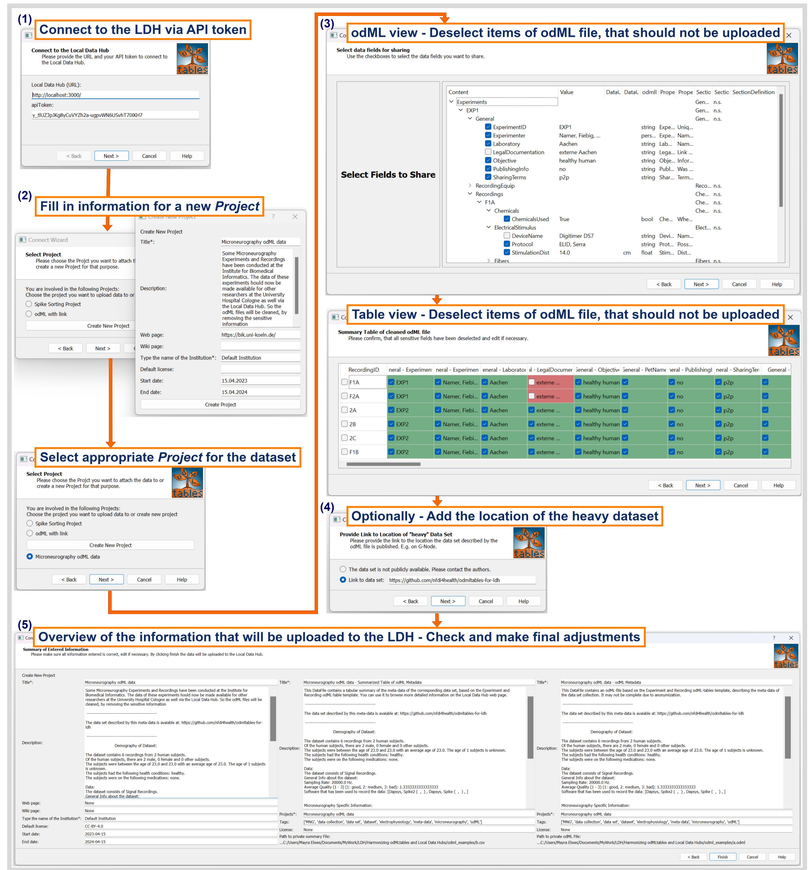
Figure 2
Activity diagram: Workflow of choosing how and what information is published in the LDH.



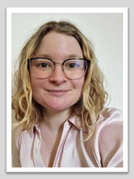
Discussion and Conclusion

The concept is in further development as an open-source Python package and will be extended to further biosignal data use cases. The current status of development can be tracked on Github (https://github.com/nfdi4health/odmltables-for-ldh). Figure 3 gives an impression of the User Interface of of the tool. This work lays a cornerstone for further development of the toolbox, allowing diverse communities of medical researchers to adopt higher level metadata sharing solutions, such as LDHs.

Figure 3
Steps the researcher performs in the extended odMLtables GUI to upload their research data. (1) Logging into the LDH (2) Creating a new Project for dataset (3) Selecting odML fields to hide in the LDH (4) Sharing actual location of dataset (5) Confirming the automatically generated structure and description and making final edits before the upload



Contact



Mayra Elwes
Institute for Biomedical Informatics
University Hospital Cologne
University of Cologne
Kerpener Str. 62, 50937 Cologne
Germany
mayra.elwes@uk-koeln.de

Github

Medical Informatics Europe 2024,
Athens, 25th – 29th of August 2024

References
[1] Troglio A, Nickerson A, Schlebusch F, Röhrig R, Dunham J, et al. odML-Tables as a Metadata Standard in Microneurography. IOS Press; 2023. p. 3–11. doi: 10.3233/SHTI230687.
[2] Kutafina E, Becker S, Namer B. Measuring pain and nociception: Through the glasses of a computational scientist. Transdisciplinary overview of methods. Front Netw Physiol. 2023 Feb 10;3:1099282. doi: 10.3389/fnetp.2023.1099282. PMID: 36926544; PMCID: PMC10013045.
[3] Meineke F, Golebiewski M, Hu X, Kirten T, Löbe M, et. al. NFDI4Health Local Data Hubs for Finding and Accessing Health Data. 1st Conference on Research Data Infrastructure. 2023 Sep. 07;. doi: 10.52825/CoRDI.v1i.375
[4] Sprenger J, Zhel L, Pick J, Sonntag M, Grewe J, et. al. odMLtables: A User-Friendly Approach for Managing Metadata of Neurophysiological Experiments. Front Neuroinform. 2019 Sep 27; 13:62. doi: 10.3389/fninf. 2019.00062
[5] Konradi P, Troglio A, Pérez Garriga A, Pérez Martín A, Röhrig R, Namer B, Kutafina E. PyDapsys: an open-source library for accessing electrophysiology data recorded