Contribution ID: **35**   Type: **Poster**

Metadata Extraction Tool and Schema Mapper for Scanning Electron Microscopy (SEM) images

Tuesday 10 October 2023 14:30 (15 minutes)

Standardized metadata and its proper storage are essential for effective management of scientific research data. The challenge lies in manually compiling such metadata, a process which can be both tedious and prone to human error. To address this problem, we introduce the Mapping Service, developed within the framework of HMC.

The Mapping Service helps to streamline the process of metadata extraction and mapping according to existing community-agreed schemas. This tool has been designed as an adaptable and extensible service suitable across various research disciplines. It allows for add-ons which facilitate the extraction of metadata from otherwise proprietary and non-standard file formats and the mapping to schemas, which strengthens the metadata's interoperability and reusability.

The Mapping Service functions as a platform hosting a diverse suite of plugins. These plugins, each equipped with two primary components—a reader and a mapper—are instrumental in achieving metadata extraction and mapping for various experimental techniques. In many research environments, accessing metadata is a bottleneck due to proprietary formats that demand specific software, often leading researchers to manually transcribe unstructured and poorly-documented metadata embedded within, for example, research images. This manual approach, especially for large datasets comprising hundreds of files, is not only time-consuming but also introduces the potential for human errors. The Mapping Service elegantly addresses these challenges: the reader retrieves metadata from a set of diverse research data, such as images or metadata files, while the mapper discerningly selects key variables prescribed by the user-selected schema from the extracted metadata. These variables are then mapped to their respective schema names, resulting in a systematically formatted JSON metadata document.

Through this poster, we showcase one use case via the mapping of Scanning Electron Microscopy (SEM)/Focused Ion Beam (FIB) tomography metadata to our published schema. This functionality is available as a plugin or "Mapping Component"on the Mapping Service and works to emphasize how a large and complex dataset containing a large amount of research images may be easily and efficiently transformed into a single metadata document using an intuitive user interface.Though the poster showcases the use case of SEM/FIB tomography, the Mapping Service has been designed to be a general-purpose tool. Additional plugins tailored to map from one arbitrary schema to another can easily be integrated, and a suite of such plugins is currently in development. Additionally, the service's standalone web service architecture and user interface ensures ease of adoption without any local dependencies or installations required by end users.

In summary, researchers across various fields seeking a streamlined approach to consistent metadata processing will find the Mapping Service to be a useful tool. With this poster, we aim to illuminate the advantages of the Mapping Service's automated extraction and mapping capabilities as well as the necessary considerations and prerequisites for its implementation in a research environment.

Please assign your contribution to one of the following topics

Metadata annotation and management close to the research process

Please specify "other" (stakeholder)

In addition please add keywords.

mapping, metadata, service, extraction

Please assign yourself (presenting author) to one of the stakeholders.

Data professionals who provide and maintain data infrastructure

Primary authors: KIRAR, Ajay (Karlsruhe Institute of Technology); VITALI, Elias Giulio Georg (Karlsruhe Institute of Technology); INCKMANN, Maximilian (Karlsruhe Institute of Technology); BLUMENROEHR, Nicolas (Karlsruhe Institute of Technology, Steinbuch Centre for Computing); STOTZKA, Rainer; JOSEPH, Reetu Elza (Karlsruhe Institute of Technology); AVERSA, Rossella (Karlsruhe Institute of Technology); JEJKAL, Thomas; HART-MANN, Volker (KIT)

Presenters: VITALI, Elias Giulio Georg (Karlsruhe Institute of Technology); JEJKAL, Thomas

Session Classification: Poster session

Track Classification: Poster session