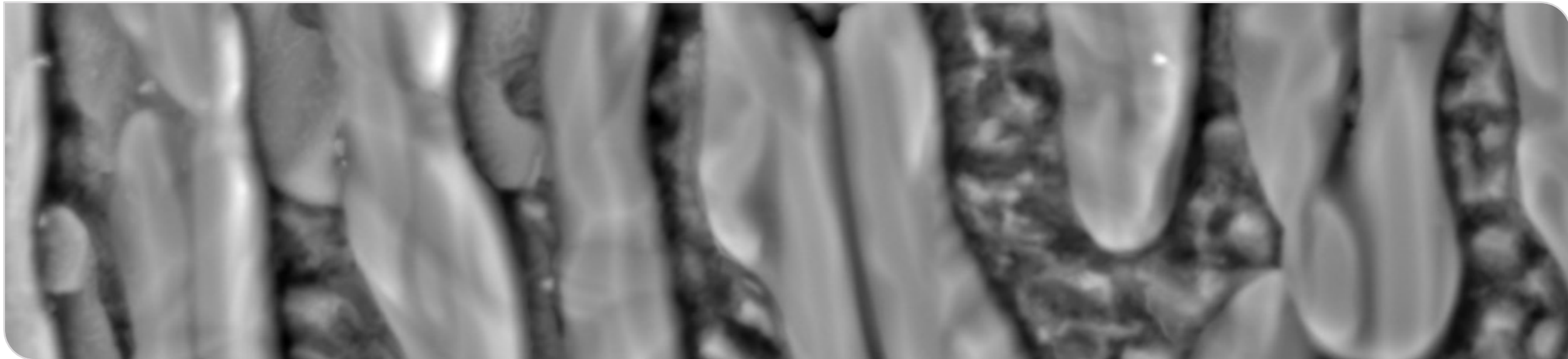# Metadata Extraction and Mapping Service

Reetu Joseph, Elias Vitali, Ajay Kirar, Rossella Aversa

Objective

Develop a 'clickable' tool to extract metadata from TIFF images generated by scanning electron microscopes (SEM) and map the metadata to the published SEM schema* for use with electronic lab notebooks and repositories

* https://ceur-ws.org/Vol-3036/paper21.pdf

Functional Objectives

1. Extract metadata from TIFF images generated by Scanning Electron Microscopes (SEM)

2. Prepare the data for mapping

3. Map the metadata with the terms from the SEM schema

4. Use a browser based GUI, such that users are able to upload the image files and download the corresponding metadata files generated by the mapping tool
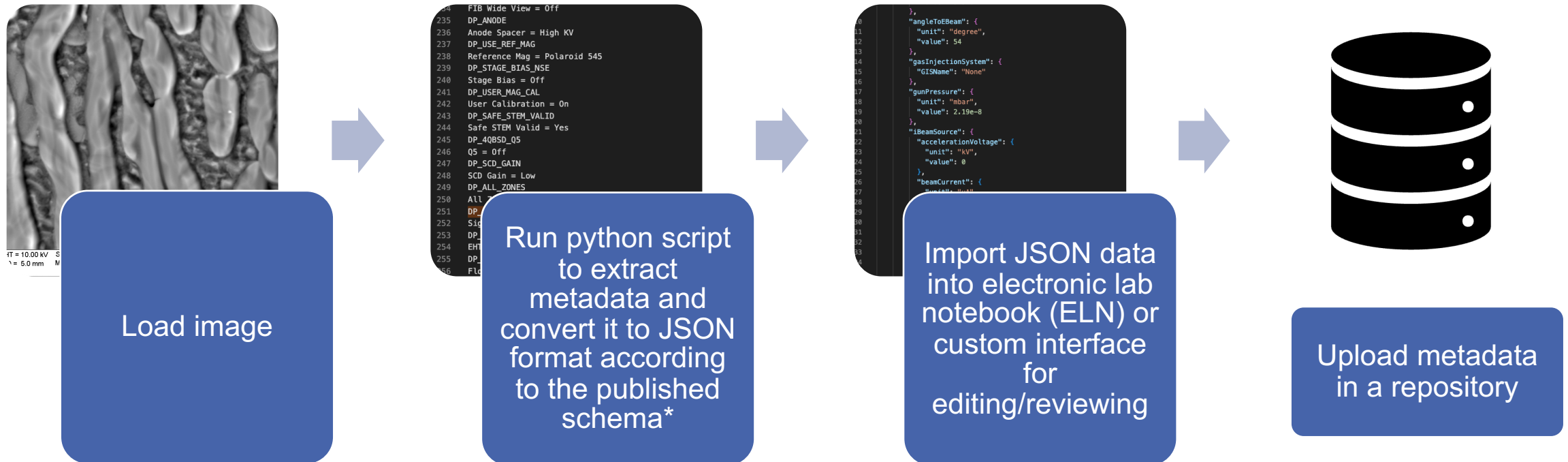
Target Audience

- For materials scientists who perform experiments on SEM

- ELN developers

April 14, 2023   Reetu, Elias, Ajay - SEM Metadata Extraction and Mapping Service   Data Exploitation Methods, Steinbuch Centre for Computing

Methods

- Use Python for data extraction and mapping

- Use the pre-existing MRI Mapping Tool as reference

- Build java Plug-in for the tool to be used with the mapping-service

- Host mapping-service from remote instance

General Workflow

Load image

Run python script to extract metadata and convert it to JSON format according to the published schema*

Import JSON data into electronic lab notebook (ELN) or custom interface for editing/reviewing

Upload metadata in a repository

* https://ceur-ws.org/Vol-3036/paper21.pdf

Graphical User Interface

Mapping-Service GUI

Home   Add mapping scheme   Show all mapping schemes   **Map a document**   REST Documentation

Map a document

Identifier

zeiss_to_json

Enter the ID of the mapping you want to use or select a mapping in the list of mappings.

Document

Choose File   📄 1-as-cast_18_Sch_10k_BSD-Compo.tif

Select the document that should be mapped with an existing mapping.

⬇ Download result   Map document

SEM Mapping Tool GUI:
https://metarepo.nffa.eu/mapping-service/mapDocument.html

Identifier: **zeiss_to_json**

- uses the python extractor located on GitHub
https://github.com/kit-data-manager/SEM-Mapping-Tool
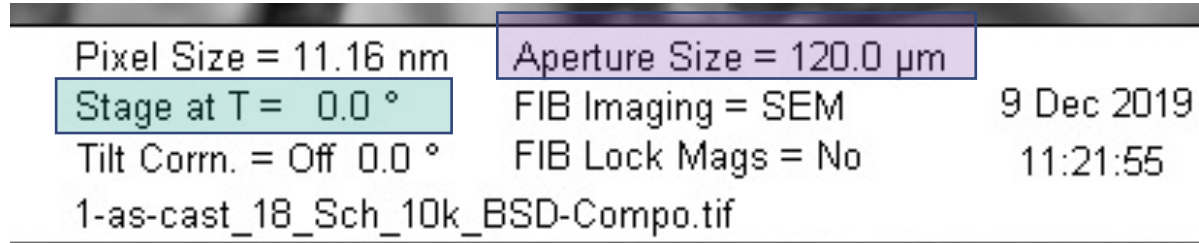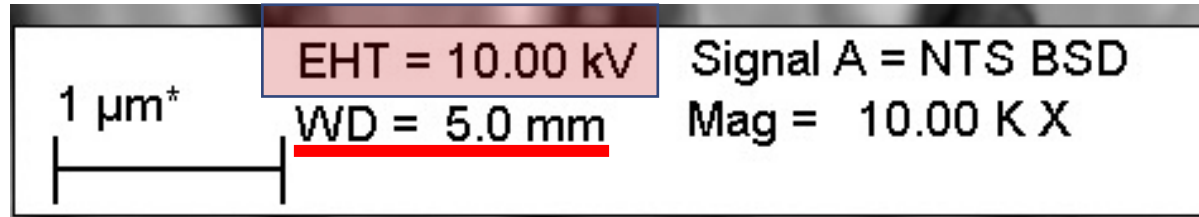
Image Metadata

Extracting Metadata stored in the TIFF file

```
AP_SAMPLE_AT_X
Sample at X = 0.0000
AP_TIME
Time :11:21:55
AP_DATE
Date :9 Dec 2019
SV_VERSION
Version = V05.07.00.00 : 08-Jul-14
SV_USER_TEXT
User Text = 1 as cast
SV_FILE_NAME
File Name = 1-as-cast_18_Sch_10k_BSD-Compo.tif
SV_IMAGE_PATH
Images = d:\images\schlabach\huot_quebec\1_as-cast\2019-12-02_auriga\
SV_USER_NAME
User Name = SABINE
SV_SAMPLE_ID
Sample ID =
SV_SERIAL_NUMBER
Serial No. = Auriga 60-46-18
```

```
1455    AP_FIB_STIGMATOR_X
1456    FIB Stig X =  -3.0 %
1457    AP_STAGE_AT_M
1458    Stage at M =  0.100 mm
1459    AP_STAGE_AT_Z
1460    Stage at Z = 36.853 mm
1461    AP_STAGE_AT_Y
1462    Stage at Y = 93.5093 mm
1463    AP_ACTUALCURRENT
1464    Fil I =  2.290 A
1465    AP_STAGE_AT_X
1466    Stage at X = 74.9182 mm
1467    AP_ACTUALKV
1468    EHT = 10.00 kV
1469    AP_STAGE_AT_T
1470    Stage at T =   0.0 @
1471    AP_SAMPLE_AT_Y
1472    Sample at Y = 0.0000
1473    AP_STAGE_AT_R
1474    Stage at R =  15.0 @
```

~1500 lines

Extracted metadata is mapped to the Schema

April 14, 2023   Reetu, Elias, Ajay - SEM Metadata Extraction and Mapping Service   Data Exploitation Methods, Steinbuch Centre for Computing

Import Metadata Document for Viewing/Editing

- Electronic Lab Notebooks
  - Schema can be imported as a template
  - Enable relevant fields for correlative characterisation
  - Working with ELN Developers
    - Herbie
    - Chemotion
    - Kadi4Mat
    - eLabFTW
    - Metadata Extractors WG of MaRDA
    - OpenBIS
    - …

- Metadata should be uploaded in a repository like the MetaStore, which registers a schema, and then metadata can be validated against the schema fulfilling FAIR principles

Custom Interface to Add, Review and Correct Metadata

SEM Metadata Editor

Click to upload JSON document

Acceleration Voltage ⓘ

Value ⓘ

10

Unit ⓘ

kV

Download
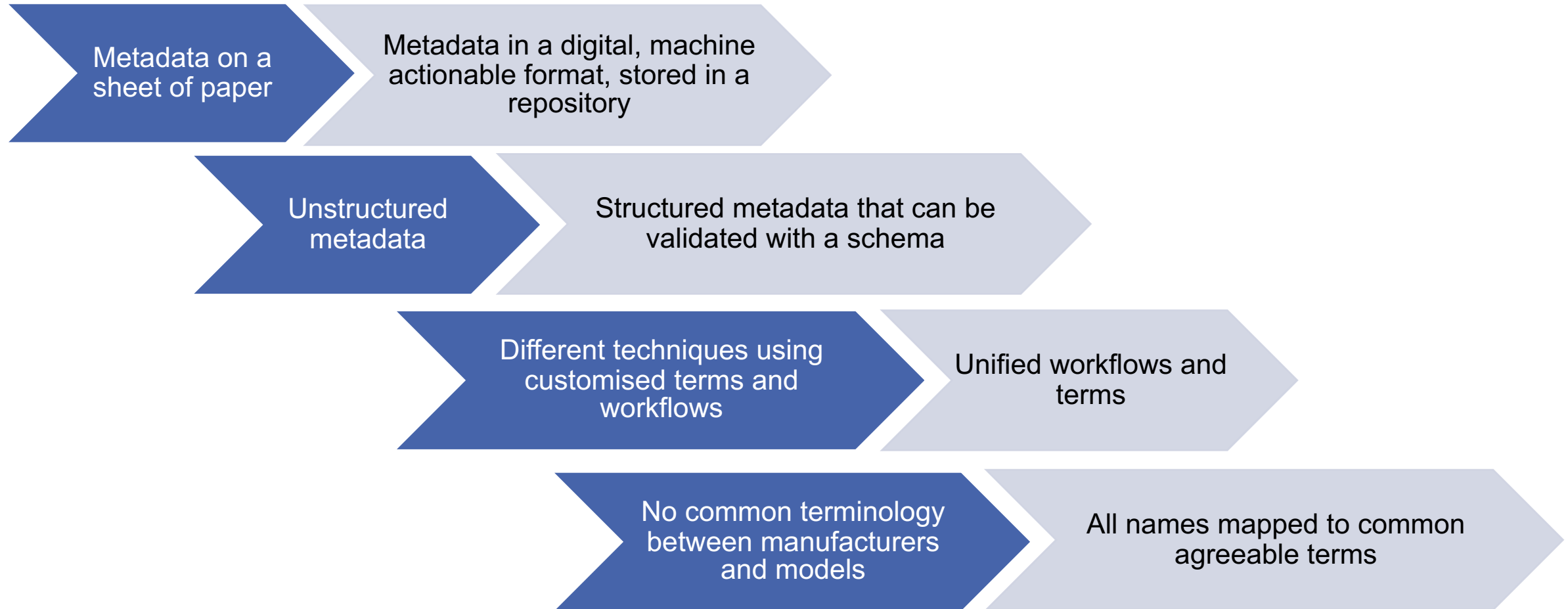
- Based on the published schema
- With drop down menus
- Information on each field
- Integrated schema validation

Available at: https://kit-data-manager.github.io/Metadata-Schemas-for-Materials-Science/

Live Demonstration

Reetu, Elias, Ajay - SEM Metadata Extraction and Mapping Service   Data Exploitation Methods, Steinbuch Centre for Computing

Past vs Future with Schemas

Metadata on a sheet of paper → Metadata in a digital, machine actionable format, stored in a repository

Unstructured metadata → Structured metadata that can be validated with a schema

Different techniques using customised terms and workflows → Unified workflows and terms

No common terminology between manufacturers and models → All names mapped to common agreeable terms

April 14, 2023   Reetu, Elias, Ajay - SEM Metadata Extraction and Mapping Service   Data Exploitation Methods, Steinbuch Centre for Computing

Additional Features in Development

- Add readable/workable filetype output along with JSON metadata document (csv)
- Batch processing

Future Work

- Add support for further instruments/manufacturers
- Implement a proper logging system for backend of new interface
- Merge/unify with existing plugins and mappers (e.g. GEMMA)
- New, simplified, and user-friendly interface for mapping service

Mapping Service Update

Reetu, Elias, Ajay - SEM Metadata Extraction and Mapping Service   Data Exploitation Methods, Steinbuch Centre for Computing

Mock Up

Technologies

- TypeScript

- JavaScript

- HTML

- CSS

April 14, 2023   Reetu, Elias, Ajay - SEM Metadata Extraction and Mapping Service   Data Exploitation Methods, Steinbuch Centre for Computing
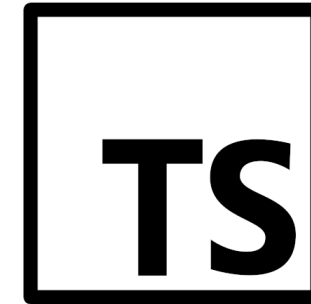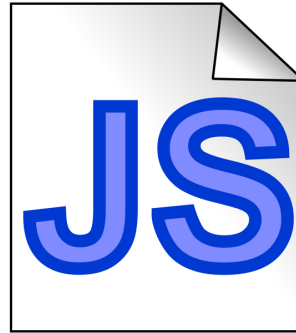
Progress

Choose a Mapping

Select or search your mapping here.

Drag & Drop your files or Browse

Map document

Thank You ☺

Reetu, Elias, Ajay - SEM Metadata Extraction and Mapping Service   Data Exploitation Methods, Steinbuch Centre for Computing