The Standardization Survival Kit (SSK)
Marie Puren, Charles Riondet, Laurent Romary, Dorian Seillier, Lionel Tadjou

To cite this version:

HAL Id: hal-01850075
https://hal.archives-ouvertes.fr/hal-01850075
Submitted on 26 Jul 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
The Standardization Survival Kit (SSK)

Bringing best practices to research communities in the Humanities

Marie Puren, Charles Riondet, Laurent Romary, Dorian Seillier, Lionel Tadjou
Inria, PARTHENOS
Why using **standards**?

Manage your data, don’t let your data manage you...

... and produce **interoperable**, **sustainable** & **reusable** results for research.
Why using standards?

“Standards are a key to great digital research, which helps to discover and understand our cultural and societal life.”

Laurent Romary
Chairman of the Technical Committee "Terminology and other language and content resources" of the International Organization for Standardization (ISO)
What are standards?

To be called "standards", they must fulfill three requirements:

- They express a **consensus**
- They are published and easily **accessible**
- They are **maintained**
The SSK: a toolkit for Humanities scholars

https://ssk-application.parthenos.d4science.org/ssk/##/
The SSK: a toolkit for Humanities scholars

- Documenting
- Supporting
- Training
- Communicating
The scenarios

Providing contextual information and relevant examples on how standards can be applied in a given research project.
Hey Mork! With some of my colleagues, I’d like to put musical scores online. But I don’t know what I can do and how to proceed! Can you help me?

Well Tork, use the SSK! I think the scenario “Collaborative Digital Edition of a Musical Corpus” can be very useful for you. It will guide you through all the necessary steps to put these scores online, and present the best practices associated to each of these steps.
Collaborative Digital Edition of a Musical Corpus

A project aims to do a digital edition of a musical corpus. The researchers need to be able to encode a broad range of musical documents in a machine-readable structure. The data to be encoded may include the musical content as provided by the composer (notes, pitches, durations, dynamics, etc.), information on the score (incipit, lyrics writer, etc.), information added by a performer when interpreting the content (timing, phrasing, various annotations, etc.), information on the visual appearance of the score (page layout, musical font, etc.) and analyses of the content in any of the other domains. The edition will be structured around a database in order to allow the users to explore it more easily. Furthermore, the project intends to be collaborative, which means it will offer anyone interested the possibility to contribute.

1. Create a digital corpus of musical compositions.

Select resources to be included in the corpus. After collecting original musical sources, transcribe them adding critical editorial signs and normalizing, where applicable, ancient poetic texts to modern usage. To get directly MEI files, use MEISE (MEI Score Editor).
# Resources and best practices

## GENERAL RESOURCES

**Video**
- Avid Technology
  - Stock.s

**Documentation**
- Barry Kiley, Jai V. Kapur
  - Music Encoding Initiative Guidelines
    - "2018"

**Schema**
- MS Schemata

**Documentation**
- MS Score Editor – MUSE

**Documentation**
- MS's Native XML editor for the MS encoding format

**Website**
- MUSE Wiki

## PROJECT-SPECIFIC RESOURCES

**Paper**
- Laurent Homary, Charles Rondot
  - EAD-OOG: A solution for project-specific EAD schemas
    - "2018/03/21"

**Paper**
- Laurent Homary, Charles Rondot
  - Towards multiscale archival digital data
    - "2017/09/12"

**Tutorial**
- XPath and Schematron for TEB Customization

**Report**
- hal

**Code**
- 

---

This page provides a list of resources and best practices, including general resources and project-specific resources, with a focus on video, documentation, and papers related to audio encoding and archival digital data.
Three layers

- **Scenario**: A complete and generic research use case composed of several steps to be followed.
- **Step**: A unique task to be performed inside a scenario with the help and recommendation of one or several resources.
- **Resource**: A standardized tool / service / document guiding the researcher in her/his tasks completion.
Managing resources

- Standards documentation (ISO, TEI)
- Official publications and reports (D4Science, HAL, Zenodo)

Documentation

- Reference libraries organised by domains and standards
- Maintenance with Zotero

Bibliography

- Code snippets (GitHub)
- Tools & services (D4Science)

Technical resources

- Wikis
  - Blog posts (Hypotheses.org)
  - Discussion lists

User communities
The scenarios are described using the **TEI** format (Text Encoding Initiative). All the information displayed within the SSK proceed from TEI files.
What is next?

- **Browsing** vocabularies and bibliography
- Creating an **account**:
  - to manage **bookmarks**
  - to customize **scenarios** (by combining existing steps from SSK’s research scenarios)
- **Contributing** (in TEI or directly on the interface):
  - creation
  - edition
  - customization
- Accessing a **multilingual** interface
Test the SSK!

https://ssk-application.parthenos.d4science.org/ssk/

ssk@inria.fr

Thank you for your attention!

http://www.parthenos-project.eu/
@Parthenos_EU
@PARTHENOSproject
contact@parthenos-project.eu