The Standardization Survival Kit (SSK)

Bringing best practices to research communities in the Humanities
What are standards?

3 keywords
- They express a **consensus**
- They are published and easily **accessible**
- They are **maintained**

Data formats
- **XML**: [https://www.w3.org/TR/xml/](https://www.w3.org/TR/xml/)
- **TIFF**: ISO 12639:1998

Protocols
- **ISO 11554:2017**: Test methods for laser beam power, energy and temporal characteristics
What are standards?

Present at each step of the research process (examples)

- **Production**: ALTO XML, standard for recording layout and logical structure of OCRed text;
- **Processing**: Linguistic annotation encoded in XML TEI (Text Encoding Initiative);
- **Archiving**: OAIS (Open Archival Information System, ISO 14721:2012) => conceptual model dedicated to the management, archiving and long-term preservation of digital documents.
Parthenos & the SSK

Initial goal of the *Standardization Survival Kit* in *PARTHENOS*:

Support and provide expertise to researchers in their use of standards

Which communities?

- Humanities
- Social Sciences
- Heritage Science
Parthenos & the SSK

Positive context of an European project:

- **Diversity**: many experts from different disciplines
- **Synergy**: willingness to cooperate

⇒ Opportunity to build use cases, inspired from real life.
Parthenos & the SSK

Development team

Inria ALMAnaCH

- Laurent Romary (DR, WP leader, supervisor)
- Marie Puren, Charles Riondet (project management, data model)
- Dorian Seillier (UI/UX Design)
- Lionel Tadjou, Damien Biabiany (development web)

Iterations, beta tests

PARTHENOS WP4 on Standardization:

- N. Larrousse & A. Joffres (Huma-Num)
- Klaus Illmayer (OEAW),
- Karolien Verbrugge (NIOD),
- Roberta Giacomini (SISMELE),
- Panos Siozos (FORTH),
- Nikolaos Papadakis (AA),
- and many more.
Concept Evolution

1. Support and provide expertise to researchers in their use of standards;
2. Give context to standards;
3. Link them to a concrete research activity (TaDirah - Taxonomy of Digital Research Activities in the Humanities);
4. Link activities between them --> Describe a research process built on the use of standards.
Not every research step is related to a standard:

- Ethics and legal issues;
- Evaluation and comparison of results;
- ...

But the notion of Best practices is always relevant!
Concept Evolution

Standards & protocols:
A experimental protocol = A suite of tasks using tools and techniques.

- Importance of stabilizing experimental protocols for Heritage Science to facilitate results exchange.
- (Pre)standardization process of protocols
- provide a user "plain" version of protocols normative texts
Concept Evolution

1. Support and provide expertise to researchers in their use of standards;
2. Give context to standards;
3. Link them to a concrete research activity (TaDirah - Taxonomy of Digital Research Activities in the Humanities);
4. Link activities between them -> Describe a research process built on the use of standards.
5. A platform for:
   a. Documenting research best practices in digital environment
   b. "Human readable" expression of experimental protocols
In PARTHENOS and beyond

Propose a user-friendly, sustainable and adaptable service:

- UX design methods
- Being "agile": brainstorming, user/usability tests
- "soft" modeling of data
- Meet the potential user communities.
The SSK & DARIAH

DARIAH working groups: take over intellectual maintenance => **WG Guidelines & Standards**

- Scenarios by other WGs
- **OpenMethods**: Submitted material can be linked (or added) to relevant scenario
- Dariah Marketplace: Context around the presented tools, services, etc.
A short video...

..is worth a thousand words!

https://www.youtube.com/watch?v=JVJ1dUDpX5I
The scenarios

Providing **contextual information** and relevant examples on how standards can be applied in a given research project.
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.
3-level Structure

Scenario description  => title, presentation, image, keywords (techniques, disciplines, objects)

1. Step description  => title, presentation keywords (activities, standards)
   a. resource (bibliographical record)
   b. resource

2. Step description (activities, standards)
   a. etc.
Resources about standards

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

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

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

User communities
- Wikis
- Blog posts (Hypotheses.org)
- Discussion lists
A high-level research guide

Collaborative Digital Edition of a Musical Corpus
Vincent Besançon, Adeline Joffres, Hélène Gautier

Last updated: 4 months ago

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 (music, lyrics, aural aspect, 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 text to modern usage. To get directly MEI files, use MEISE (MEI-Script Editor).

   RESOURCES (specifications, papers, tutorials, etc.)

2. Change format into MEI files if necessary.
   Convert the Sibelius files into MEI files through the plugin SibelMEI.

   RESOURCES (specifications, papers, tutorials, etc.)
Resources and best practices

General Resources

Enrichment of MEI files

MEIMassaging
Django/python powered MEI massaging app
Author(s): DuChemin
URL: SafeValue must use [property]=binding; SafeValue must use [property]=binding; SafeValue must use [property]=binding; https://github.com/DuChemin/MEIMassaging (see http://co.co/nu/secretKey) (see http://co.co/nu/secretKey) (see http://co.co/nu/secretKey)

meiView
JavaScript demo application displaying variant readings of historical musical pieces - zolaemili/meiView
Author(s): zolaemili
URL: SafeValue must use [property]=binding; SafeValue must use [property]=binding; SafeValue must use [property]=binding; https://github.com/zolaemili/meiView (see http://co.co/nu/secretKey) (see http://co.co/nu/secretKey) (see http://co.co/nu/secretKey)

MEItoVexFlow
JavaScript library to render MEI notation using VexFlow - TEI-Music-SIG/MEItoVexFlow
Author(s): TEI-Music-SIG
URL: SafeValue must use [property]=binding; SafeValue must use [property]=binding; SafeValue must use [property]=binding; https://github.com/TEI-Music-SIG/MEItoVexFlow (see http://co.co/nu/secretKey) (see http://co.co/nu/secretKey) (see http://co.co/nu/secretKey)
Easy-to-use & collaborative platform

- Consult and follow the guidelines expressed in the scenarios
- Propose new scenarios
Why contributing?

- Make your research project align with the best practices in your community
- Get peer review and visibility
- Share a project in another form than the usual blog/article (a new way to disseminate your work)
Coherence with our principles

The SSK, a 100% standards web app
Scenarios in TEI

The scenarios are described using the TEI format (Text Encoding Initiative). All the information displayed within the SSK proceed from TEI files.

http://github.com/ParthenosWP4/SSK/tree/master
Managing resources with bibliographic standards

All the references are managed by the open source management software Zotero, and can be found in the a dedicated library.

[https://www.zotero.org/groups/427927/ssk-parthenos/]
RESTful architecture

http://github.com/ParthenosWP4/SSK/tree/dev

- Flexible, easy to deploy and maintain architecture
- Independent entities communicating via REST services.
Contribute to the SSK

A SSK tutorial
Choose explicit titles

The title is the main entry point for the readers. To ensure that the message will driven home, we advice to build titles:

- starting with a **verb** or a **gerund**, describing the process (for a scenario) or the action (for a step) that the user will read about;
- between 10 – 100 characters
**Scenario**: explain the scientific problematic and describe the solution put in place.

**Step**: describe the purpose of the action, how it relates with the previous ones and give an overall presentation of the different kind of methods and tools the resources would point to.

- Extend the acronyms cited and to briefly present the projects mentioned.
- Shouldn’t exceed 1500 characters (but not too short).
- Possible to point to external links
- Lists of items
Associate keywords to the scenario or the step

SSK content is described with a set of controlled vocabularies, particularly created for describing Humanities research. They are:

- Research activities, taken from Tadirah;
- Research techniques, taken from Tadirah;
- Research objects, taken from Tadirah;
- Standards, taken from the SSK Standard Knowledge base;
- Disciplines, taken from aureHAL
Associate keywords to the scenario or the step

When editing the description of a scenario, the available keywords are:

- Research Techniques
- Research objects
- Disciplines

For the steps, the most important keyword is the Activity, that should be unique for each step. It is also possible to pick some techniques, objects and standards.

For each keyword type, we recommend to choose between 1 and 4 terms.
Choose an illustration for the scenario

- The illustration must closely relate with the purpose of the scenario, i.e. not only with the discipline or the period studied.
- Screenshots are accepted.
- Landscape orientation image are recommended
- Maximum size: 2 Mo
- Accepted formats: png or jpg
- It must be published under the licence CC-BY or CC-0. (If CC-BY, provide the author's name).
Identifying state of the art references is a prerequisite before actually add the resources to the steps. When we are talking about resources, we mean a standardized tool, service or document helpful for the task completion.

They have to be of one the following types:

- Specification (normative document) of a standard;
- Technical reports;
- Scholarly papers;
- A single scholarly monograph;
- Lists of bibliographic references;
- Scripts and code samples;
- Computing libraries;
- Collection of structured data;
- Computing tool, software;
- Curating or hosting service;
- Tutorials or guidelines;
- Blog posts;
- Specialised mailing list, forum, etc.
Identify relevant resources

The resources should be separated into several groups:

- General resources like standard specifications
- Project-related resources, i.e. how the standards are used in a real research project.
Link the resources to the step

The resources metadata are recorded in the dedicated SSK Zotero Library (see [here](#)).

The fields required by the SSK are:

- **Title**
- **Author**
- **Date**
- **Url**: The url of the resource is most of the time mandatory (except when the resource is a printed document unavailable online).
- **abstract**: An abstract is very important to give an overview of what the user will find under the linked resource. Zotero is very often able to get this abstract from the webpage, but it may be necessary to check it, and sometimes create it.
Customize a step (advanced mode)

The SSK is adaptable by nature and contributors don’t have to start from scratch their scenario.

It is possible to create a scenario with existing steps as basis.

But if the content of the step doesn’t exactly fit, it is also possible to modify it directly in the new scenario.

Two update mechanisms exist:

- Update the keywords of a step (in particular the standards)
- Update the resources of a step
Link scenarios (advanced mode)

Link scenarios together, or in other words, include a scenario (entirely or partially) into another is an interesting possibility when a scenario is a pre-condition or the continuation of another one.

For instance, a scenario related to the preservation of 3D models can be preceded by a scenario explaining how to create such models.

The most common use cases are the following:

- Add a prerequisite scenario (as a first step)
- Associate a scenario that can be the follow-up of the current (as a last step)
Contribution interface
Contribution Interface
Contribution Interface

1. Terms of Use  
2. Scenario  
3. Steps

For each scenario, you must have at least 2 steps.

You can either create a new step from scratch, or reuse existing ones from other scenarios.

STEP 1
Title of the step

Description

Tags

Resources

You can add this step in several languages.
Create resources

http://ssk.huma-num.fr/#/add-resource

Resource's link

https://hal.inria.fr/PARTHENOS/hal-01850080

Enter the resource's link and submit to Zotero with the button below.

GET RESOURCE METADATA

{
  "key": "MUNDNYZ2",
  "version": 0,
  "itemType": "journalArticle",
  "creators": [
    {
      "firstName": "Marie",
      "lastName": "Puren",
      "creatorType": "author"
    },
    {
      "firstName": "Alix",
      "lastName": "Chagué",
      "creatorType": "author"
    },
    {
      "firstName": "Manuela"
    }
  ]
}
Create resources

http://ssk.huma-num.fr/#/add-resource

Resource has been successfully added to SSK's Zotero library with itemKey: MUNDNYZ2
Roadmap

- IMPORTANT: Write scenarios and steps in a text document in parallel => template available [here](https://www.zotero.org/groups/427927/ssk-parthenos/)
- Add resources:
  - Browse Zotero Library [https://www.zotero.org/groups/427927/ssk-parthenos/](https://www.zotero.org/groups/427927/ssk-parthenos/) and copy the Item Key (last for of URL)
  - Add resources to Zotero with [http://ssk.huma-num.fr/#/add-resource](http://ssk.huma-num.fr/#/add-resource)
- When finishing the scenario: copy the file name you see in the green box:

New Scenario SSK_sc_1557476611209.xml and its steps have been successful added on github !!!