

PARTHENOS – Bridging Research Infrastructures in the Humanities

Poster Session Proposal for DHN 2018

Author: Sheena Bassett, PIN Scrl., Prato, Italy.

Abstract

PARTHENOS supported by two ERICs from the Humanities and Arts sector, DARIAH and CLARIN, along with ARIADNE, EHRI, CENDARI, E-RIHS and IPERION-CH will deliver guidelines, standards, methods, pooled services and tools to be used by its partners and all the research community. Four broad research communities are addressed – History, Linguistic Studies, Archaeology, Heritage and Applied Disciplines and the Social Sciences. By identifying common needs, PARTHENOS will support cross disciplinary research and provide innovative solutions.

By applying the FAIR data principles to structure the work on common policies and standards, the project has produced tools to assist researchers to find and apply the appropriate ones for their areas of interest which are described in this poster session. A virtual research environment will enable the discovery and use of data and tools and further support is provided with a set of online training modules.

Introduction

PARTHENOS aims at strengthening the cohesion of research in the broad sectors of Linguistic Studies, Humanities, Cultural Heritage, History, Archaeology and related fields through a thematic cluster of European Research Infrastructures. Built around two ERICs from the Humanities and Arts sector, DARIAH and CLARIN, along with ARIADNE, EHRI, CENDARI, E-RIHS and IPERION-CH and involving all the relevant Integrating activities projects, PARTHENOS is delivering guidelines, standards, methods, pooled services and tools to be used by its partners and all the research community.

PARTHENOS is achieving this objective through:

- joint policies and solutions, including provisions for cross-discipline data use and re-use;
- the implementation of common AAA (authentication, authorization, access) and data curation policies, including long-term preservation;
- quality criteria and data approval/ certification;
- IPR management, also addressing sensitive data and privacy issues;
- foresight studies about innovative methods for the humanities;
- standardization and interoperability;
- common tools for data-oriented services such as resource discovery, search services, quality assessment of metadata, annotation of sources;
- communication activities, and joint training activities.

The aim of this poster presentation is to familiarise researchers visiting DH Nordic with the PARTHENOS project and its outputs. By doing so, we want to encourage them to make the most out

of our products and services and inspire them to become involved in the activities of the project and Research Infrastructures in general.

How PARTHENOS is bridging research across the Humanities

PARTHENOS started with defining the User Requirements using Use Cases contributed by the many partners and from an extensive literature search based on reports addressing researcher's needs in the Humanities and Arts. The first "User Requirements" report was completed in February 2016 with an updated internal version released in September 2016. All the subsequent work taking place relies on the User Requirements to develop their contributions. PARTHENOS has identified four main user communities:

- **History** (in a broad sense: including Medieval Studies, Recent History, Art History, Epigraphy, etc.)
- **Language-related Studies** (including Literature, Linguistics, Philology, Language Technology, etc.)
- **Archaeology, Heritage & Applied Disciplines** (including Cultural Heritage, Archives, Libraries, Museums, Preservation / Conservation experts, Digital curation / edition / publishing, etc.)
- **Social Sciences** (in a broad sense: Sociology, Political Science, Geography, Anthropology, Cultural Studies etc.).

Researchers across the Humanities need to use standards to ensure that their results are valid, reproducible and reusable. Standards cover practices and processes, protocols and data formats that are necessary to guarantee the quality and interoperability of research outputs. The FAIR Principles provide a structure that supports researchers across all disciplines and which is underpinned by standards. PARTHENOS has addressed those requirements identified as the most important by the four main user communities, focussing on the commonalities between the user communities, to deliver tools and services that enhance findability, accessibility, interoperability and reusability as well as cross-discipline research

The FAIR principles are being used to make a structured overview of existing and needed policies for research data which will be implemented as an interactive "wizard". Researchers from the four identified user communities will be able to select their area of study and the particular aspect of the data life cycle of interest to find the commonly used policies and best practices within their user community. Where policies may not exist, similar policies from other disciplines will be suggested for the researcher to consider. The wizard, can be downloaded and adapted for use as required, e.g. by adding more resources or adding new categories.

An online Standards Survival Kit (SSK) is supported by a printed "Why standards" guide. Cartoon characters from a fictional planet (Digitus) are being used to promote the standards message as these are independent of subject domain or culture. Researchers will be able to look up the standards relating to their disciplines and areas of activity. The SSK is built around scenarios relating to different activities from various Humanities disciplines such as corpus modelling in TEI, multi-spectral imaging and tracking the dissemination of a medieval work. Each scenario is divided into different steps, and involves specific tasks and illustrates the best research practices in a given community, made accessible and reusable for other researchers wishing to carry out a similar project but unfamiliar with the recommended tools, formats, methods to use, etc.

As an example, let us use a sociologist who is a novice in digital methods, but who wants to disseminate a collection of field survey data online, so that they could be used by other researchers

in the long-term. By browsing in the SSK, she or he will find a standardized scenario that meets her or his needs: “Encode and modelize field surveys for their online dissemination”. This

reference scenario consists of nine steps, each with the associated resources. Some of the scenario’s steps are:

- Step 1 - obtain the informed consent of the participants addresses the legal and ethical issues.
- Step 4 - “Anonymize” offers a curated and up-to-date list of resources to help the researcher respect ethical practices and adopt proven techniques for anonymizing the collected data.
- Steps 2 and 6 stress on the importance of using tested standards, such as EAD to “Collect and classify” the data and TEI to “Transcribe the interviews”, highlighting the importance of proper data modelling before disseminating them, and giving- access to appropriate resources on the subject.

Note that more advanced users will also be able to edit the scenarios themselves, by submitting new resources or adding new steps. They can also create new scenarios. The SSK scenarios and steps can be easily extended, reused and customized, thanks to their flexible data model in TEI3.

An example of this is another tool, the PARTHENOS Data Management Plan (DMP) template. This is tailored for the Humanities, which is important for researchers as DMPs are rapidly being integrated into their everyday workflows. The template addresses the domain-specific procedures and practices within the Humanities , paying special attention to the relevant standards and guidelines used in data management. The PARTHENOS DMP template describes the life cycle of the data that is created, collected, archived and preserved by projects and research infrastructures in the Humanities, including information that makes data FAIR: findable, accessible, interoperable and reusable.

In terms of structure, the template will consist of a general section aimed at everyone followed by sections which contain more detailed questions where policies and common practices may vary according to the research area. Likewise, the template is addressed at two main audiences – the researcher and the data repository and some of the questions will only apply to one or the other. It is proposed that template for the researcher may also be domain specific, with archaeology and linguistics as the first examples (capitalising on existing experience with the project). One innovative feature of the template which is designed to make the completion of the template as straightforward as possible is that if the data is to be deposited with one of the listed key data repositories, then all the applicable following questions will be completed automatically according to the repository selected. The guidance will also encourage researchers to follow the common policies and standards used by their community and use best practice for their data management, thus enabling interoperability, preservation and reuse of the data and supporting cross-discipline research.

A generic cloud structure and virtual research environment forms the basis of the PARTHENOS tools and services. The PARTHENOS Entities Model is the underlying framework that researchers will interact with to achieve their tasks in PARTHENOS. The key components consist of a searchable registry (based on a minimal metadata schema, the Back Bone Thesaurus as proposed by DARIAH) of data sources along with the tools to enable researchers to extract, reformat and analyse, annotate

their required information along with services for the deposit and preservation of their research work. Many of these tools and services are pre-existing where access is provided along with user authentication.

PARTHENOS is also providing professional development and training to educate not just researchers but also policy makers, managers, funding bodies etc. A first module which was introduced at the European Summer University in Digital Humanities in Leipzig (July 2016) aimed at informing less experienced researchers about the objectives and benefits of Research Infrastructures and how they can use them. Several new modules are available, covering topic areas such as data management, management of RIs and RI collaboration with more modules under development, along with a series of webinars, which can all be accessed via the project website.

As part of its bridging and dissemination activities, PARTHENOS is establishing international liaisons and relationships with related organisations and other Research Infrastructures. In order to build alliances and foster better understanding across the Humanities of the current challenges and possible solutions to aspects such as digitisation, data management and preservation, several joint RI workshops have already been held on subject areas such as the FAIR Principles with more planned for the next 12 months. PARTHENOS is also building a network of external experts who are reviewing and providing contributing feedback to the project on its reports and outputs.

Please see www.parthenos-project.eu for further information on all the mentioned outputs and access to the training modules.

Bibliography

- PARTHENOS, Report on user requirements, 2016, <https://goo.gl/3lwI5J>
- PARTHENOS, Guidelines for Common Policies Implementation, 2017, <https://goo.gl/GisWAa>
- PARTHENOS, Standards Survival Kit, 2016, <https://goo.gl/VK62qM>
- PARTHENOS, Report on Standardization, 2017, <https://goo.gl/pwCLYS>
- PARTHENOS, The Common Semantic Framework, 2017, <https://goo.gl/dqLWXM>
- PARTHENOS, The Cloud Infrastructure, 2017, <https://goo.gl/OnxQ8g>
- PARTHENOS, Tools and Services, 2017, <https://goo.gl/mL4U0j>