WP2 in Parthenos:

Community involvement and requirements

Prato – 2016-12-14

Sebastian Drude

CLARIN ERIC
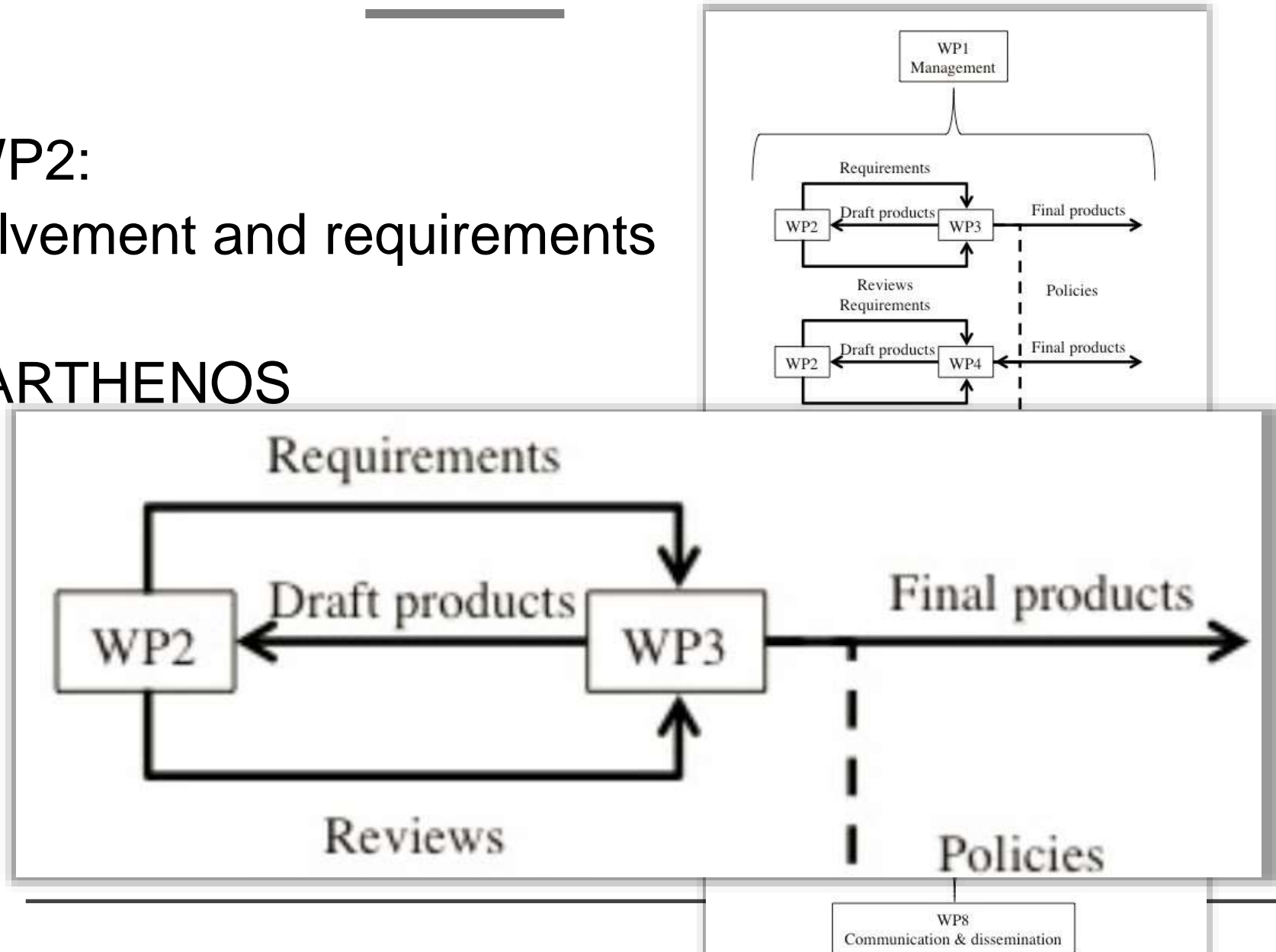
WP2: Community involvement and requirements

1. The role of WP2 in Parthenos

2. Work done in WP2

3. Major findings from the Requirements Report D2.1

Role of WP2 in Parthenos

PARTHENOS WP2:
Community involvement and requirements

At the core of PARTHENOS

Role of WP2 in Parthenos

T2.1 : Definition of users' requirements about data policies

T2.2 : Definition of standardization requirements

T2.3 : Definition of interoperability & related services requirement

T2.4 : Definition of education & training requirements

T2.5 : Definition of communication requirements

WP3: Data policies

WP4: Standardization

WP5: Interoperability and semantics & WP6: Services and tools

WP7: Skills, Professional Development and Advancement

WP8: Communication, dissemination and outreach

Major work in WP2

- Regular biweekly/monthly meetings of the council (30 total since end of June 2015)

- D2.1 **"Report on User Requirements"**: first version Jan. 2016, finalized Oct. 2016

- D2.2 **"Report on the assessment of the education and training plans and activities"** has been submitted (Oct. 2016)

- All "Task Forces" are working in close contact with 'their' respective Work Package, providing early feedback on activities, way before the next deliverables

Parthenos' user communities

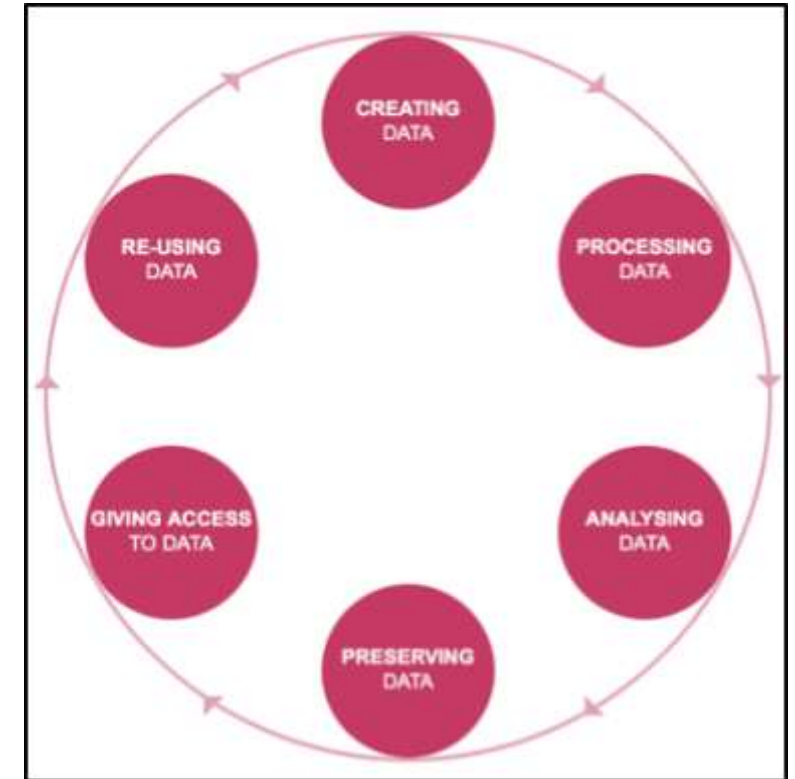| PRIMARY FIELDS:			
1	Studies of the Past	History (in particular recent Hist.), Archaeology, Art History, Epigraphy, …	17
2	Language-related Studies	Literature, Linguistics, Philology, Language Technology, Epigraphy, …	17
3	Heritage & Applied Discipl.	Cultural Hertg., Archives, Libraries, Museums, Digital curation / edition / publishing, …	15
SECONDARY FIELD:			
4	Broader Social Sciences	Sociology, Anthropology, Cultural Studies, Political Science, …	8
OTHER DISCIPLINES:			
		Computer Science	3
		Philosophy	1
		Musicology	1
		Jewish Studies	1
		…	

D2.1: Report on User Requirements

- Covers all major areas relevant to Parthenos and support for (digital) research in general:

  1) Data policies:
     a) Research Data Lifecycle;
     b) Quality Assessment & Assurance;
     c) IPR, Open Data & Open Access
  2) Standardization
  3) Interoperability, services and tools
  4) Education & training
  5) Communication

- Usually not obtained by surveys or similar, but based on earlier reports and documents

D2.1: main findings on (1) data policies

A: Data lifecycle

- Major concerns regards **data preservation** (40 %)

- Improvements to **data accessibility**

- **Data** and **metadata quality**

- Tools and services **in tune** with **researchers' workflows** (often discipline-specific)

- Researchers wish to **control their data**

- **Wider access** to **datasets** would facilitate **collaboration** between researchers and **enhance funding opportunities**

- Understand the **entire scholarly research process** offer applications and services supporting the **corresponding workflow**

The UKDA research data lifecycle model

Basis: **68 use cases** on data lifecycle (**20 docs**)

D2.1: main findings on (1) data policies

B1: Quality Assessment of Digital Repositories (DR)

- Establish **discipline-specific** services of **data management** (different requirements)
- Selection / verifiability of data (**standardized**)
- **Appropriate description** of the variety of resource types
- selection / promotion of **high-quality deposit services**
- clear **guidelines** / **procedures** for management, archiving, and sharing of data
- Use of **PID system** and **Federated Identity**

B2: Quality Assurance of Data and Metadata Items

- Metadata: enable cross-domain **data re-use**
- Datasets with **metadata** (discovery / access)
- **Data**: relevant, understandable, trustworthy
- Quality of **data documentation** (contextual information: evaluate if the data is relevant)
- **Accuracy**: description / data representation
- **Consistency of data values** (conceptual level) of data format (structural level)
- Metadata **functional requirements** (discovery, authentication, administration, …)

Basis: **8 use cases** on quality assessment of DR and quality assurance of data / metadata items (**17 docs**)

D2.1: main findings on (1) data policies

C1: Intellectual Property Rights

- Definition of a **framework of licenses** for data in Portal of Infrastructure's community
- Creation of a **tool to identify the copyright status** for data or collections
- Creation of a **tool to associate the identified copyright status with data** within Portal of Infrastructure's community
- Creation of an AAI (**Authentication and Authorization Infrastructure**)
- Updated **information** on the **IPR legal frameworks** across Europe

C2: Open Access & Open Data

- **Controlled access** to protected data
- Common methods of **data citation**
- Definition of best practice for **reviewing academic articles** before publication
- Identification of **sustainability models**
- If there is a traditional publication process, define the strategy to ensure optimal dissemination of **open access** publication
- Definition of a **minimum set of data to share** under a CC0 or CC-BY license
- Guarantee the **searchability** of data

Basis: **21 use cases** on IPR, Open Data and Open Access (**12 docs**)

D2.1: main findings on (2) standards

- 20 use cases and 20 user requirements on Standards
- Standardization is key to data sharing and reuse

Standards needs for:

- Tools (creation, publication, mapping, visualization etc of data)
- Formats for describing data (metadata)
- Guidelines for describing procedures
- Shareable file formats
- Sustainable repositories

Preconditions for using standards

- Easy to use
- Adaptable to others standards
- Supports work process
- Data is easily findable and accessible
- Budget, time and expertise

D2.1: main findings on (3) interoperability, services and tools

- More than 50 functional requirements from 90 use cases

- General level requirements, shared across several use cases and disciplines, expressing the same needs e.g.: data quality, availability, accessibility and enrichment

- Specific needs driven by specific disciplinary concerns

- Shared set of priorities at the general level (i.e. search and information display tools)

- Detailed, domain driven requirements regarding tools

- Technical and non-technical requirements (i.e. sustainability of tools and datasets) to be considered as action points for other WPs, and inserted in the agenda for the development of mid and long-term actions

D2.1: main findings on (4) education & training

- Information from desk research / analysis of project reports and deliverables + questionnaire

- There is a motivation to systematically develop, organize and set up training and education modules in a comprehensive manner among the analysed infrastructure projects

- Often not realized due to the typical project characteristics (fixed project duration etc.)

- The need for a kind of "umbrella" perspective and developments as in PARTHENOS has been stated clearly

D2.1: main findings on (4) education & training

- Regarding delivery of training:
  - preference for face -to-face meetings
  - combination of workshops, summer schools or virtual meetings with moderated distance-learning modules is the most common and promising way of delivery of training
- Regarding topics for training:
  - raise the awareness about DH in general and research infrastructures in particular
  - insight into Humanities researchers epistemological practices are needed
  - Digital Methods and Tools in Research Infrastructures

D2.1: main findings on (5) communication

Basis:

- A sample of relevant scientific communication platforms such as e-journals and repositories where collected and analyzed to obtain the criteria for their evaluation.

- Goal was to gather information on relevant journals and repositories in the field and to enable their evaluation with regard to their attractiveness for Digital Humanities researchers for publishing their findings there.

- Regarding dissemination strategies, a number of dissemination reports from PARTHENOS-related projects were gathered and scrutinized to get insight into the dissemination strategies.

D2.1: main findings on (5) communication

- Evaluation criteria derived from the analyzed journals and repositories range from the domain and covered topics of the journals to regional and international coverage, languages, formats and outputs accepted to the ability to be quantitatively analyzed.

- The dissemination reports revealed a group of five most evident activities:

  - 1. dissemination activities via project's website

  - 2. partners' institutional websites are used for the dissemination of information

  - 3. newsletters

  - 4. press releases are common means when it comes to dissemination strategies

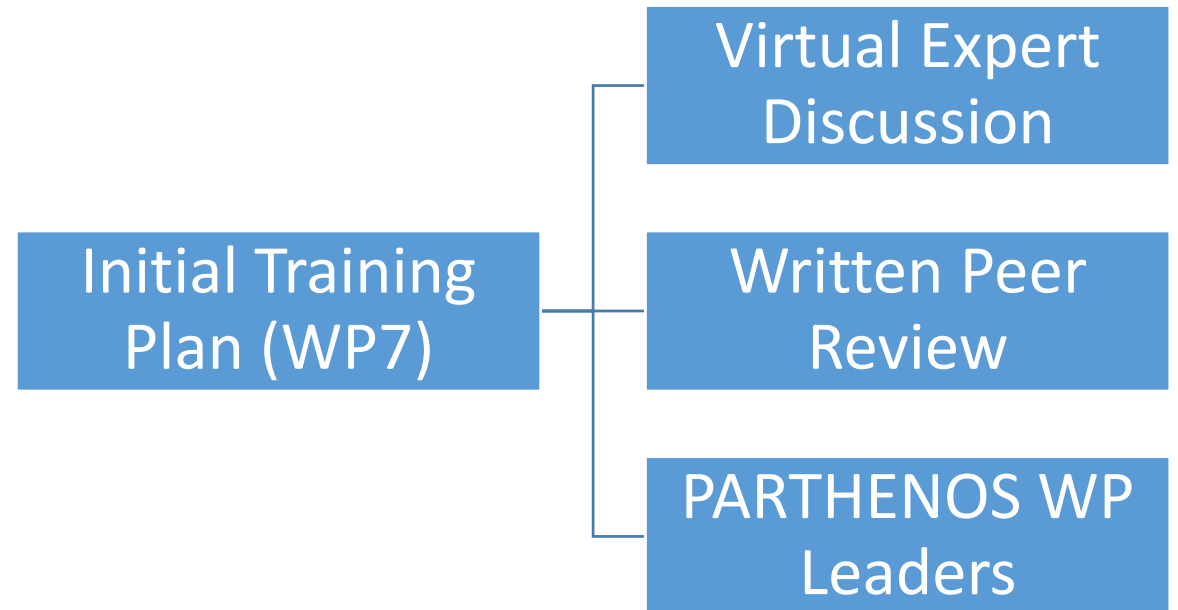 - 5. networking and consulting at conferences in various phases of the projects

Further results

D2.2 "Report on the assessment of the education and training plans and activities" delivered in October 2016

Task 2.4's aim = to support the development of the first version of the training plan by providing structured feedback

Selection of experts / peers reflected the target groups of the training plan (researcher; content specialists in CHIs; technical developers; managers of institutions / projects)

Initial Training Plan (WP7)

Virtual Expert Discussion

Written Peer Review

PARTHENOS WP Leaders

Further results

Results:

- the underlying approach of the training plan is valid
- teaching contents and modes of delivery rated solid, coherent and well thought out by all the experts
- breakdown of target groups / audiences has been evaluated as being clear and sufficient
- asynchronous mode of content delivery proposed in the training plan critically discussed – flexibility vs. barrier of accessibility
- need of awareness raising critically discussed – "Is it actually still necessary to distinguish communities that use RIs from those that do not ?"
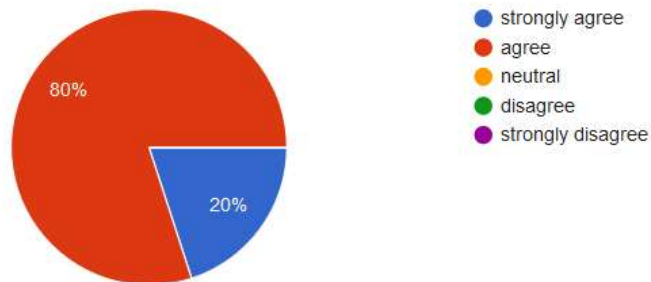
Further results

Results from Leipzig Summer School:

 European Summer University in Digital Humanities in July 2016

 workshop on "Digital Research Infrastructures in the Humanities: How to use, build and maintain them" led by Jennifer Edmond.

 teaching contents reflected the "Phase 1" modules of the initial training plan

The workshop content met my expectations (5 responses)

- strongly agree
- agree
- neutral
- disagree
- strongly disagree

80%

20%

WP2 in Parthenos:

Community involvement and requirements

Prato – 2016-12-14

Sebastian Drude

CLARIN ERIC