Improving the Visibility, the Access and the Scientific Use of Heritage Content

Services
Persée brings together digital skills and a pool of software and hardware resources (data modeling, operational process line from digitization to long term preservation, SEO, web design, data display).

Users centered
Persée develops partnerships with researchers and heritage institutions (libraries and museums). They define the descriptive markup of the digitized corpus and collections; they determine the retrieval and display features.

A complete and integrated process line

#1 Digitization
High quality digitization integration of digitized documents OCR

#2 Documentation
Accurate description of the documents XML encoding Author index Permanent ID

#3 Semantization
LOD (ontologies, thesauruses) Specialized index (taxonomy, place names, characters...)

#4 Dissemination
Websites RDF triplestore Web services OAI-PMH Usage statistics

#5 Archiving
Automated and manual quality control checks Long term preservation

Persée Portal for the scientific publications
- Entire collections of academic journals, serials, books and proceedings
- 700,000 full text documents available via www.persee.fr
- 22 disciplines
- Open Access

The Perseids for the research projects
- Corpus of primary sources, archives, gray literature, and published documents
- Dedicated websites
- Specialized index and tools for text analysis
- Open Access

Data Persée for re-using metadata
- Graph representation of the metadata (RDF)
- SPARQL End Point and Sparkis (natural language query)
- Thematic datasets to download
- Tutorials and examples of queries
- Open Data

Open standards to create, curate and preserve the data (TEI, EAD, RDF, METS, MODS, marcXML, MADS, DC, KBART, OAIS)

Business plan with long-term funding

Users' community involvement

Sustainability

Interactions with French and European research infrastructures (Humanum, CoReX, DARIAH)

Open source softwares, distributed IS and interoperability procedures

Nathalie Fargier, June 2017