

Himanis

Historical
MANuscript
Indexing for
user-controlled
Search



Project Summary

Manuscripts are among the most important witnesses to our European shared cultural heritage. In recent years, large quantities of historical handwritten documents are being scanned and made available through web portals. Yet, the wealth of information conveyed by the text captured in these images remains largely inaccessible. General users and researchers more and more expect to query handwritten resources in plain text like printed books, but current handwritten text recognition technology is still far from offering sufficiently accurate transcripts on historical documents. This situation raises the need of searching approaches specifically designed for text images.

HIMANIS aims at developing cost-effective solutions for querying large sets of handwritten document images. To this end, innovative keyword spotting, indexing and search methods will be developed, tested, adapted and/or scaled up to meet the real world conditions required. HIMANIS will associate Computer Science (A2iA, UP-VLC and RUG), Humanities (IRHT) and Cultural Heritage (The European Library, the Archives Nationales of France and Bibliothèque Nationale de France) institutions and produce technology to generate new, research-based knowledge from historical manuscripts. As a challenging and particularly interesting case study, the large collection of registers produced by the French royal chancery (14th-15th c.) will be used as a test bed. This will allow sustainable use and management of this important cultural resource. Implementing user feedback will also increase our understanding of the meanings that it holds for people and how they

perceive, use and interpret it, thereby meeting societal challenges in promoting its use and re-use.

The expected outcomes of this project are:

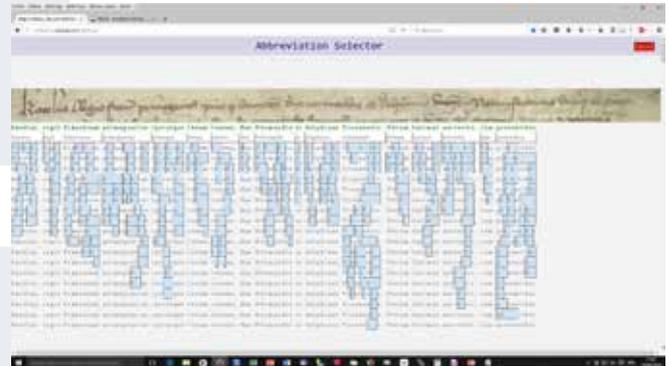
- a) a new indexing/searching technology for historical manuscripts, brought to the market by A2iA;
- b) a new paradigm to study our historical heritage, as conveyed by manuscripts, by using full text search technology.
- c) a new vision of the raise of nation states in Europe via a new study of the corpus under this paradigm.

Application and impact

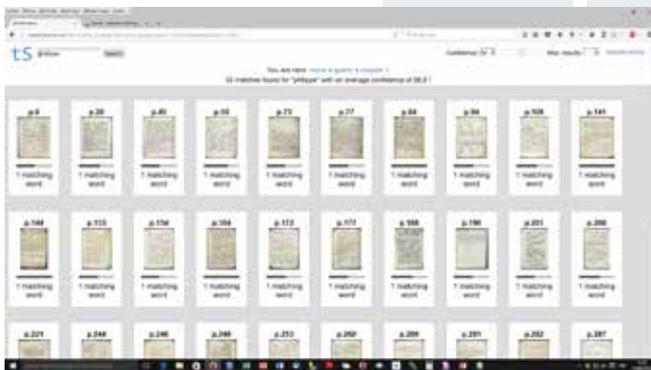
HIMANIS will (produce technology to) generate new, research-based knowledge from historical manuscripts. The vast collection of the CHANCERY corpus possesses in France an iconic value and is part of a common representation of the Past of the Nation: it belongs to the Trésor des chartes which has been held for centuries with the relics of the monarchy and is the starting point of a centralization process which is still very present in nowadays France. This corpus of material artifacts are the neutral tangible cultural heritage on which scientific research as well as national mythologies and intangible ways of thinking are based. They are, sometimes through various distorting mirrors a key stone in the multiple construction of self-narratives or of the European history. An enhanced access to these resource (interactive, plain text search, etc.) will allow sustainable use and management this important cultural resource by understanding the meanings that it holds for people and how they perceive, use and interpret it, and meet

Himanis

Historical
MANuscript
Indexing for user-
controlled Search



Abbreviation Selection



UPV tranScriptorium Demo

societal challenges in promoting its use and re-use. This new path will help creating new perceptions of the Past, which is erroneously seen as dark and brutal. Such a direct access will offer the opportunity to balance the different opinions and identify in the medieval Past some of the roots of Renaissance humanism, and emphasize and promote its effective, positive aspects.

Therefore, to foster the accessibility of these documents is required. In this context, the expected impact of the project objectives achievement will unleash the economic potential of European cultural heritage enabling the emergence of new services and business models, and providing a better way to handle this information for educational and scientific institutions.

The great visibility of the scenarios described in Section A will play a fundamental role for achieving this expected impact.

With the project objectives achievement it is expected also to have an important impact in the area of content providers of cultural heritage for the creation of their digital collections. In this sense, a successful result derived from the project objectives accomplishment will lead to the availability of a new technology that will enable users to access to historical documents. Finally, and no less important, for the case of non-experienced users in paleographical, the developed techniques will lead to innovative solutions the will help them to quickly achieve enough experience in this field.

Coordinator

Institut de Recherche et d'Histoire des Textes (France).

Participants

- Universitat Politècnica de València (SPAIN).
- A2iA (FRANCE).
- Rijksuniversiteit Groningen (NETHERLANDS).

Dates

01/11/2015 - 31/10/2017

Budget

Total project funding: € 409.746,00

Funding awarded: € 390.869,00

Subject area(s)

Digital libraries, Historical manuscripts, Keyword searching and indexing, Paleography and Diplomatics, Medieval Studies.

Project website

www.irht.cnrs.fr/fr/recherche/les-programmes-de-recherche/himanis

<http://himanis.hypotheses.org>