eContentplus

Project title	**Bernstein – The Memory of Papers** Collaborative systems for paper expertise and history
Action	Cultural and scientific/scholarly content 5.1.1 Promoting the enabling infrastructure – Targeted Projects

1. Proposal summary

The objective of project Bernstein is to create an integrated European digital environment for the expertise and history of paper. The project will interlink all existing European databases of paper reproductions, make their content accessible to specialized image processing tools for the measurement of paper features, and provide an interface to the digital resources of domains related to paper studies or by which the knowledge about papers can be enriched and contextualized. Additionally, a strong dissemination plan including ready to deploy paper expertise software packs will ensure the sustainability of growth and interest in paper studies beyond the project's lifetime.

The importance of paper expertise comes from paper being a "memory bearer": a source of information about the history of societies and an aid in identification of documents. Through examination of its characteristics the date, location and mill of its production can be estimated and information obtained about the technological, economical, social and cultural environment of its origin; while tracing the movement, usage and accidents in time and space of a sheet of paper provides a thread of captured events that is the basis for historical research and exploration of cultural heritage. The enormous amount of papers constituting a sizable part of what can be known of Europe's past – millions of samples stretching back 800 years – is both a richness waiting to be tapped into, and a technological challenge.

The Bernstein project has the ambition to generate the conceptual and technical infrastructure to enable the access to paper expertise to the broadest range of users, accommodating multiple usage scenarios. The principal targeted user communities are the historians and the cultural heritage conservators, although other areas where paper identification capabilities are required are expected to benefit from the project, such as the art market, forensic science, security research and the papermaking industry. Therefore the focus is on historical paper, while applications to modern papers are not excluded. The consortium brings together all the major European actors in the field of digital historical paper expertise (hence the partner choices) coming from both humanities and computer sciences. We strive to become world-leaders in our research and development area. The project will further stimulate collaboration between the academia, government institutions and the industry across European nations. For further details we invite you to take a look at our project's website: http://www.bernstein.oeaw.ac.at.
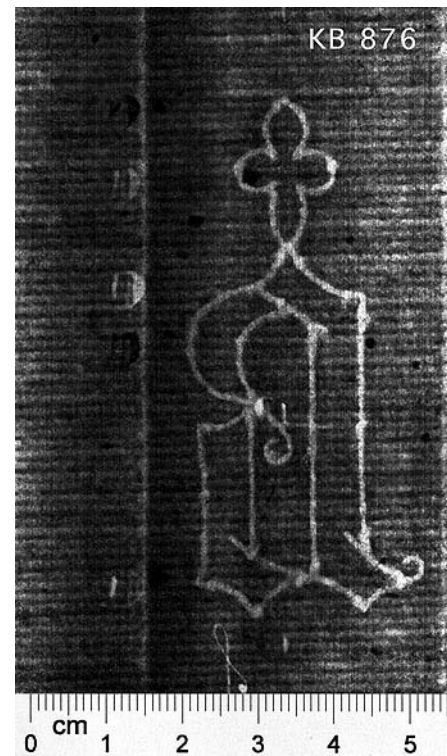

Figure 1 – The primary data used in the Bernstein project are images such as this reproduction obtained by electronradiography, from a paper used in the Low Countries during the 1470s. It shows different features characterizing a unique sheet of paper. In the middle we see the watermark, the horizontal lines are called laid lines, the two vertical ones chain lines, and the variation in greyscale reproduces the variation in the density of the paper pulp. (*KB, http://watermark.kb.nl*)

2. Participant names and profile

1. Austrian Academy of Sciences, Vienna, Austria [*coordinator*] — *Main expertise*: project management, digital image processing, database systems, manuscript studies, watermark standards. — *Products*: paper and watermarks database: http://www.oeaw.ac.at/ksbm/wz/wzma2.htm; paper features measuring software: http://mywebpage.netscape.com/atanasiuvlad/ad751/; visual enhancement: http://mywebpage.netscape.com/atanasiuvlad/bluenile/.

2. Archives of the State of Baden-Württemberg, Stuttgart, Germany — *Main expertise*: database systems, watermark standards. — *Products*: paper and watermarks database: http://www.landesarchiv-bw.de/piccard/.

3. Laboratory for Occidental Medieval Studies in Paris, Paris, France — *Main expertise*: quantitative history of paper, Renaissance printed books. — *Products*: Quantitative history of paper: http://lamop.univ-paris1.fr/W3/.

4. Deutsche National Bibliothek, Leipzig, GERMANY — *Main expertise*: bibliography. — *Products*: International bibliography of paper history: http://www.ddb.de/sammlungen/pdf/ibp.pdf.

5. Dutch University Institute for Art History Florence, Florence, ITALY — *Main expertise*: paper in art history. — *Products*: paper and watermarks database: http://www.iuoart.org/wmdb.htm.

6. Delft University of Technology, Delft, NETHERLANDS — *Main expertise*: artificial intelligence, digital image processing. — *Products*: image processing tools: http://ict.ewi.tudelft.nl/index.php?option=com_sections&id=37&Itemid=97.

7. National Library of the Netherlands, The Hague, NETHERLANDS — *Main expertise*: database systems, incunabula of the Low Countries, watermarks. — *Products*: paper and watermarks database: http://watermark.kb.nl.

8. Institute for Information Systems and Computer Media, Graz, AUSTRIA — *Main expertise*: internet services integration, knowledge management systems, digital libraries. — *Products*: knowledge management system: http://www.hyperwave.com.

9. Liverpool University, Liverpool, UNITED KINGDOM — *Main expertise*: distributed search, name disambiguation, cross-domain resource discovery, text mining. — *Products*: Cheshire3 database integration system: http://www.cheshire3.org; ISTC database development: http://www.bl.uk/catalogues/istc/.

3. Problems addressed and solutions proposed

Integration — The project is a response to the low usability of existent resources despite their potentiality. Particularly paper identification – which is the single most important application – cannot be achieved without Europe-wide collaboration between databases, measurement tools and contextualization resources. The reason is a mix of absence of interoperability, non-standard data formats, lack of coordination between IT projects, poor interdisciplinary approaches and a deprivation of vision and bold innovation spirit. The difficulty of gaining visibility on what is available negatively affects the users' experience of these digital resources, further aggravated during exploitation by the constant switching between various on-line and off-line databases and tools, provided the user speaks the language for which the software was designed. Finally, the would-be developers of new databases and tools can hardly benefit in such a chaotic environment from past efforts.

The solution proposed by the consortium rests on the agreement that collaboration between resources would substantially increase their usage and impact, as compared to their independent existence. In order to avoid major modifications of existent resources and ensure ease of scalability, integration will not happen by forcing each database and each tool to fit to a unique standard, but will be provided by a versatile interface able to address the individuality of all resources. The user will be able to interact with a multitude of contents and content manipulators through a single integrated workspace.

Databases — Integration is not only about software interoperability, but also about data formats. For the description of watermarks current databases use derivatives of the standard of the International Association of Paper Historians, a systems that has its roots in the XIXth century, generates many misclassifications and cannot be used by machines (it describes watermarks semantically, by what they are – "a bull's head with a quatrefoil on top of it" –, not by how they look as geometrical shapes). Regarding the description of other paper features, there is no standard to follow and few quantitative practices are known. Based on the long standing experience of the partners, the project will propose new standards, adapted to the modern technological conditions and to the state of the art in historical paper research.

Tools — Present on-line paper databases work independently from image processing tools conceived for the manipulation of paper reproductions: users have to download images one by one – when technically possible and legally admissible – in order to enhance or measure some features, and have no mean other than textual to search for watermarks even when the shape retrieval software exists. The project aims to bring the existing specialized software on-line and give them access to the databases. The data–tools link will implicitly create the necessary infrastructure to support automated identification of papers.

Contexts — A further problem is that current services in paper expertise are cut off from auxiliary resources such as bibliographies and external resources like catalogues of libraries, archives or museums. The Bernstein project will address the first problem by linking the paper databases to the most complete bibliography of paper history ever produced, supplied by one of the partners. The second issue – of much wider reach – is tackled through the interoperability of a specific paper database – on papers of the books printed in the Low Countries before 1500 – and the library catalogue describing the content of those books, so as to offer to a book historian simultaneous information about the content and the physical consistency.

Users — The last line of action of the project concerns the weak relation between users and content developers. Comprehensive dissemination actions will ensure that users get informed about the projects products, the validation process of the products will involve external users, a particular care will be give to design and ergonomy aspects, and a package of guidelines and paper expertise software will prepare for a harmonic and informed development of paper studies.

4. Target users and benefits of proposed solution for target users

The project addresses users from the following sectors, historians being the main beneficiaries:

Science — historians, particularly historians of the paper, book and manuscripts, art historians; computer scientists specialized in image processing, shape analysis, database integration;

Culture — conservators (librarians, archivists, museographers), private art collectors;

Education — higher education in humanities (history, art history) and arts (conservation technologies);

Market — auction houses, antiquarians;

Industry — art objects restorers; papermaking industry and artisans;

Government — forensic experts, security personnel dealing with paper documents;

Private — private individuals.

The principal benefit brought by the project is paper identification, together with other gains:

Identification — paper samples supplied by users or stored in the databases for which the time and place of production are unknown, can be identified by analysis of their features and comparison with existing benchmarks (all users have an interest in this application);

Authentication — it can be estimated if two papers where produced in similar or identical circumstances (mostly interesting for the governmental and art market sectors);

Contextualization — contextual information about papers can be obtained (benefiting humanities scholars);

Measuring — features measurement is a primary goal in paper studies for art restorers and paper makers given the need of producing quality replicas of historical documents;

Deployment — development packages are particularly interesting for conservators and art collectors who whish to publish databases of paper reproductions from their holdings;

Other — other benefits are the ability to enhance images for visual inspection, to perform searches in vast paper databases and find bibliographical references on paper history.