

ECP 2005 CULT 038097/Bernstein

Bernstein

Final Project Report

September 2006 – February 2009

<http://www.memoryofpaper.eu>

<http://www.bernstein.oeaw.ac.at>

Deliverable number/name	<i>D7.8 Final Project Report</i>
Dissemination level	<i>Public</i>
Delivery date	<i>23 March 2009</i>
Status	<i>Final</i>
Author(s)	<i>The Bernstein Consortium</i>



eContentplus

This project is funded under the *eContentplus* programme¹,
a multiannual Community programme to make digital content in Europe more accessible, usable and exploitable.

¹ OJ L 79, 24.3.2005, p. 1.

1 Table of content

1	TABLE OF CONTENT	2
2	PROJECT OBJECTIVES	3
3	CONSORTIUM.....	4
4	PROJECT RESULTS/ACHIEVEMENTS.....	6
4.1	BERNSTEIN WATERMARK TERMS	6
4.2	INTEGRATED WORKSPACE	7
4.2.1	<i>Catalogue</i>	7
4.2.2	<i>Result list</i>	9
4.2.3	<i>Statistical functionality</i>	9
4.2.4	<i>Atlas</i>	10
4.2.5	<i>Paper bibliography</i>	11
4.2.6	<i>Expertise</i>	12
4.2.7	<i>Dissemination kit</i>	12
5	TARGET USERS AND THEIR NEEDS.....	13
6	UNDERLYING CONTENT	14
7	SUMMARY OF ACTIVITIES.....	16
7.1	DATA AND DATA INTEGRATION	16
7.2	CONTEXTUAL DATA	17
7.3	DISSEMINATION	18
7.4	MANAGEMENT	18
8	IMPACT & SUSTAINABILITY	19
8.1	IMPACT	19
8.2	SUSTAINABILITY.....	19
9	FURTHER INFORMATION.....	21

2 Project Objectives

The present project revolves around paper. The significance of paper derives from being the ubiquitous physical carrier for information exchange until the present times. As such, the study of paper is used on the one hand to identify undated paper documents or expertise documents of questioned authenticity. On the other hand, paper studies have also a historical dimension, revealing aspects of technological evolutions, economical infrastructures, state policies, etc. interwoven into human networks across countries. Knowledge of this kind is partly obtained from the physical characteristics of paper, a source of ‘hidden information’ as opposed to the visible inscription on a paper object. Watermarks visible when holding a bank note against light are an example for such hidden information. They constitute also the most prominent characteristics examined by historians as well as reproduced and documented in catalogues. But a great amount of such and similar paper data with a broad geographical and temporal spread is necessary in order to create a reliable information source.

The goal of project Bernstein *was* the creation of a European integrated digital environment for paper history and expertise. Bernstein connects now all European watermark databases accessible through the Internet at the onset of the project. It offers a comprehensive and unrivalled information source about paper. The databases are augmented by specialized image processing tools for measuring, authenticating and dating papers, and by a plenitude of contextual data with bibliographical, historical and geographical (GIS) contents. A substantial further project goal was the dissemination of the achieved results to a broad audience in the form of a series of exhibitions, a book about paper history and watermarks and an easily installable software package for paper cataloguing.

3 Consortium

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. The project consortium consists of nine partners from six countries, among which the largest collections of paper and watermarks.

1. Austrian Academy of Sciences, Vienna, Austria
(OEAW, <http://www.oeaw.ac.at>),
Commission for Palaeography and Codicology of Medieval Manuscripts in Austria
(<http://www.oeaw.ac.at/ksbm>),
Commission for Scientific Visualization (<http://www.viskom.oeaw.ac.at>),
▶ Management, mediaeval watermarks database & repertories, image processing.
2. Archives of the State of Baden-Württemberg, Stuttgart, Germany
(LABW, <http://www.landesarchiv-bw.de>),
▶ world's richest digital collection of watermarks.
3. Graz University of Technology, Institute for Information Systems and Computer Media, Austria
(TUG, <http://www.iicm.edu>),
▶ integration software (implementation), user interface.
4. Laboratory for Occidental Medieval Studies in Paris, France
(LAMOP, <http://lamop.univ-paris1.fr>),
▶ contextual historical datasets, historical GIS.
5. German National Library, Deutsches Buch- und Schriftmuseum, Leipzig, Germany
(DNB, <http://www.d-nb.de>),
▶ bibliography on paper, paper collection.
6. Dutch University Institute for Art History Florence, Italy
(NIKI, <http://www.iuoart.org>),
▶ Renaissance paper database, art historical expertise.
7. Delft University of Technology, Information and Communication Theory Group, Netherlands
(DUT, <http://tudelft.nl>),
▶ image processing, data mining.
8. Koninklijke Bibliotheek, National Library of the Netherlands, The Hague, Netherlands
(KB, <http://www.kb.nl>),
▶ Watermarks in Dutch incunabula database.
9. University of Liverpool, Great Britain
(LU, <http://www.liv.ac.uk>),
▶ integration architecture, bibliography integration.

Further institutions which have contributed prominently to the project are:

Istituto Centrale per il Restauro e la Conservazione del Patrimonio Archivistico e Librario,
Rome, Italy

(ICPAL, <http://www.patologiaibro.beniculturali.it>),

► Italian portal, Italian watermark terms, exhibition, book.

Paper and Watermark Museum of Fabriano, Fabriano, Italy

(<http://www.museodellacarta.com>),

► exhibition, book, Zonghi database.

Institut Valencià de Conservació i Restauració de Béns Culturals, València, Spain

(IVCR, <http://www.ivcr.es>),

► Spanish portal, Spanish watermark terms, exhibition.

State Historical Museum of Russia, Moscow, Russia

(SHM, <http://www.shm.ru>),

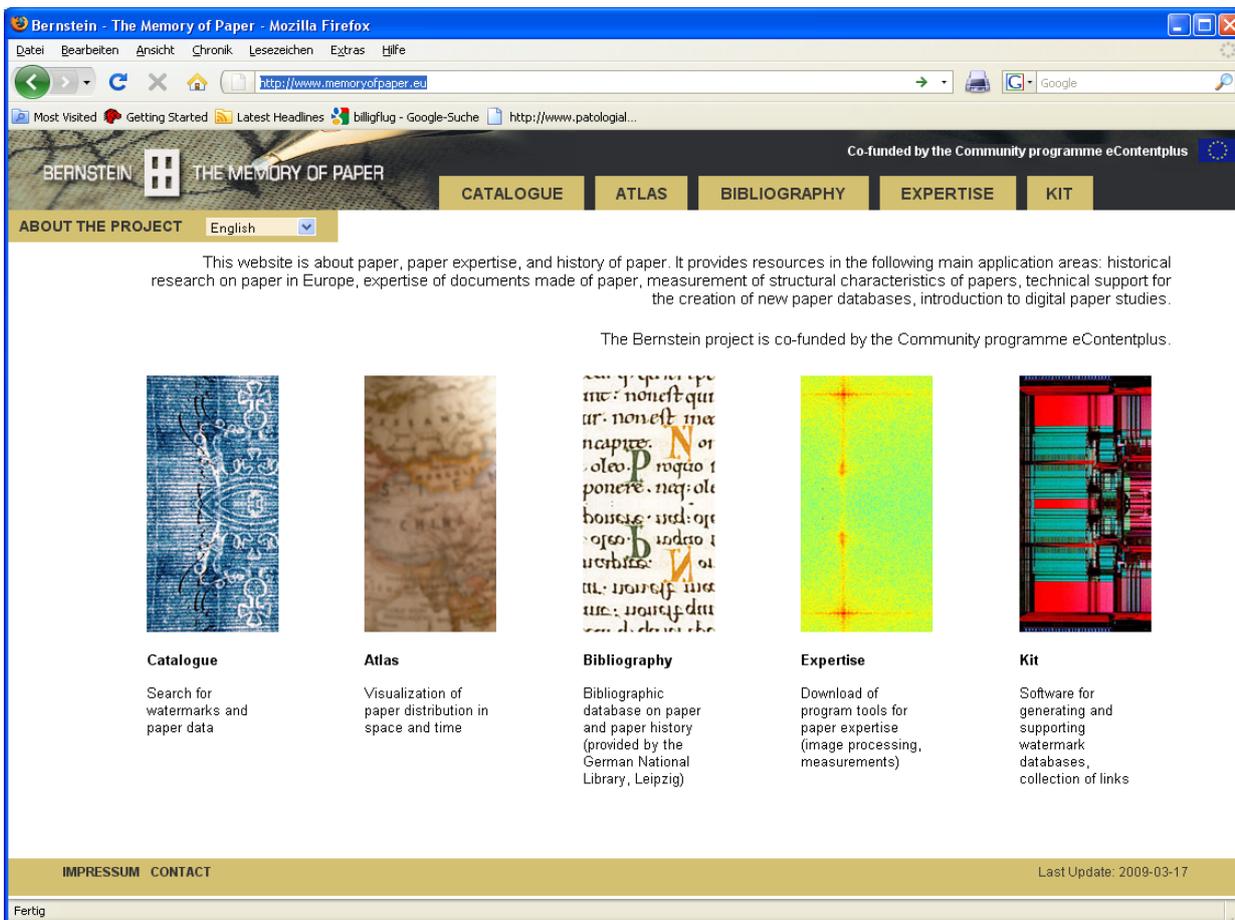
► Russian portal, Russian watermark terms.

4 Project Results / Achievements

4.1 Bernstein Watermark Terms

An essential step for the integration of the four watermark databases (Piccard-Online, WILC, WZMA, NIKI – ordered according to the number of records in each) was the adoption of a textual watermark description standard for the project. This textual watermark description standard (www.bernstein.oew.ac.at/watermark_terms.pdf) should achieve two goals: firstly, creation of a common multilingual nomenclature in the form of a vocabulary for all relevant terms for describing watermarks (Bernstein vocabulary for watermark description). Watermarks of the same type can now be described in six languages (English, French, German, Italian, Russian, and Spanish) by the same names and data interoperability like search across all databases which is now made possible. Secondly, definition of a textual watermark description standard offering a classification scheme for hierarchically organized watermark types which has so far been realized on three hierarchy levels (see www.bernstein.oew.ac.at/Bernstein_systematics.pdf).

Both aims, nomenclature and classification scheme, were necessary in order to realize a unified model for better mutual matching of the contents of the existing digital watermark databases. In the future, the textual watermark description standard will be useful and extensible to other collections of watermarks beyond this project's lifetime. It could become the standard for new collections, which have not yet been described or digitized completely.



Bernstein - The Memory of Paper - Mozilla Firefox

http://www.memoryofpaper.eu

BERNSTEIN THE MEMORY OF PAPER

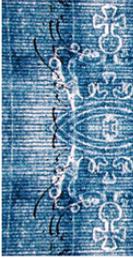
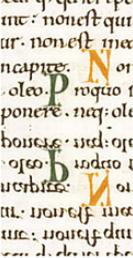
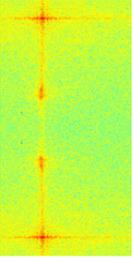
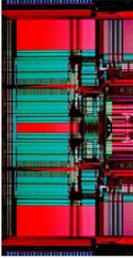
Co-funded by the Community programme eContentplus

CATALOGUE ATLAS BIBLIOGRAPHY EXPERTISE KIT

ABOUT THE PROJECT English

This website is about paper, paper expertise, and history of paper. It provides resources in the following main application areas: historical research on paper in Europe, expertise of documents made of paper, measurement of structural characteristics of papers, technical support for the creation of new paper databases, introduction to digital paper studies.

The Bernstein project is co-funded by the Community programme eContentplus.

				
Catalogue	Atlas	Bibliography	Expertise	Kit
Search for watermarks and paper data	Visualization of paper distribution in space and time	Bibliographic database on paper and paper history (provided by the German National Library, Leipzig)	Download of program tools for paper expertise (image processing, measurements)	Software for generating and supporting watermark databases, collection of links

IMPRESSUM CONTACT

Last Update: 2009-03-17

Fertig

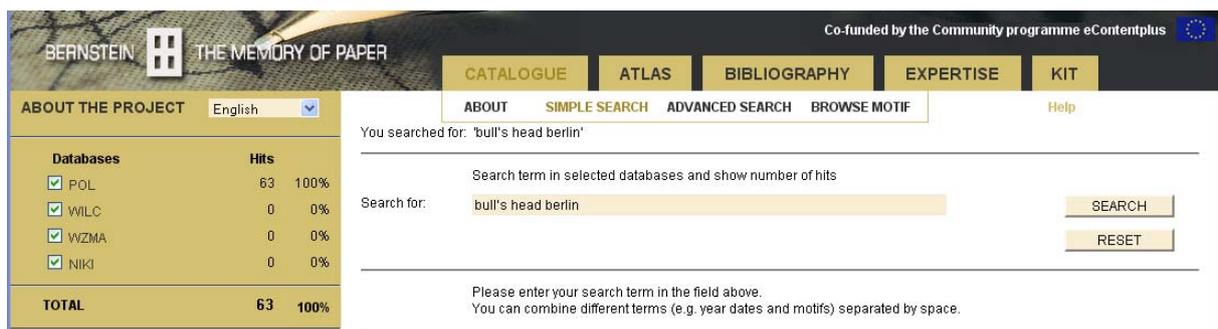
Figure 1: Start page of the integrated workspace (Bernstein web portal)

4.2 Integrated Workspace

The integrated workspace (<http://www.memoryofpaper.eu>) is the backbone of the project that provides the digital environment necessary for the integration of resources. Specifically, this is an Internet application interfaced in six languages (English, French, German, Italian, Russian, and Spanish) that gives access to all the Bernstein resources, of which the main components are the catalogue, the atlas, the bibliography, the expertise, and the toolkit (figure 1).

4.2.1 Catalogue

The main component is the *Catalogue* which allows search in and data retrieval from the various online databases (figure 2). The search can be formulated in six languages. All search terms are translated automatically into the supported languages according to the Bernstein vocabulary for watermark descriptions. A search in Italian for *sirena* is carried out also for *mermaid* (English) and *Meerjungfrau* (German). Catalogue offers three modes for the search: *simple search*, *advanced search*, and *browse motif*. In simple search the user just enters the search terms in a search field. *Simple search* is performed within all record fields available in each of the databases (i.e. motif, place of use, depository, date, height, distance of chain lines, and reference number).



Databases	Hits	
<input checked="" type="checkbox"/> POL	63	100%
<input checked="" type="checkbox"/> WILC	0	0%
<input checked="" type="checkbox"/> WZMA	0	0%
<input checked="" type="checkbox"/> NIKI	0	0%
TOTAL	63	100%

Figure 2: Catalogue, simple search (Search for bull's head and Berlin)

In *advanced search* the user can combine several search fields, specify single dates or intervals and use specific search items which are not available in all databases (figure 3). The *browse motif* search is an implementation of the Bernstein systematics for watermark motifs. It offers the possibility to navigate in the tree structure of the systematics by names or by icons.



Datenbanken	Treffer	
<input checked="" type="checkbox"/> POL	132	90.4%
<input checked="" type="checkbox"/> WILC	0	0%
<input checked="" type="checkbox"/> WZMA	14	9.6%
<input checked="" type="checkbox"/> NIKI	0	0%
GESAMT	146	100%

Figure 3: Catalogue (German), advanced search (motif= bull's head, date=1470-1480, place of use=Austria)

Each search is performed within the ‘original’ databases. The databases were not mirrored but kept at the original locations of the database holders. This provides the advantage that the data is always up-to-date and also avoids copyright problems. The access to the technically different databases is implemented through SRU-gateways. The response times have proven to be satisfactory until now.

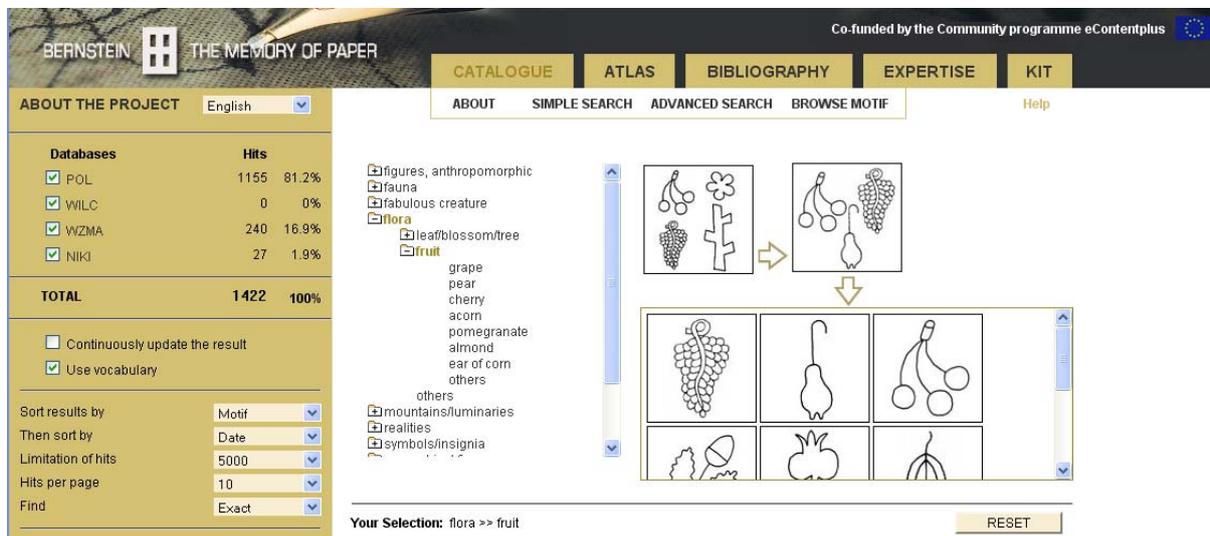


Figure 4: Catalogue, browse motif (flora/fruit)

Any subset of watermarks resulting from a search in the databases can be visualized in three ways: list of the found items, statistical properties of the returned subset, and geographical distribution by cartographic mapping.

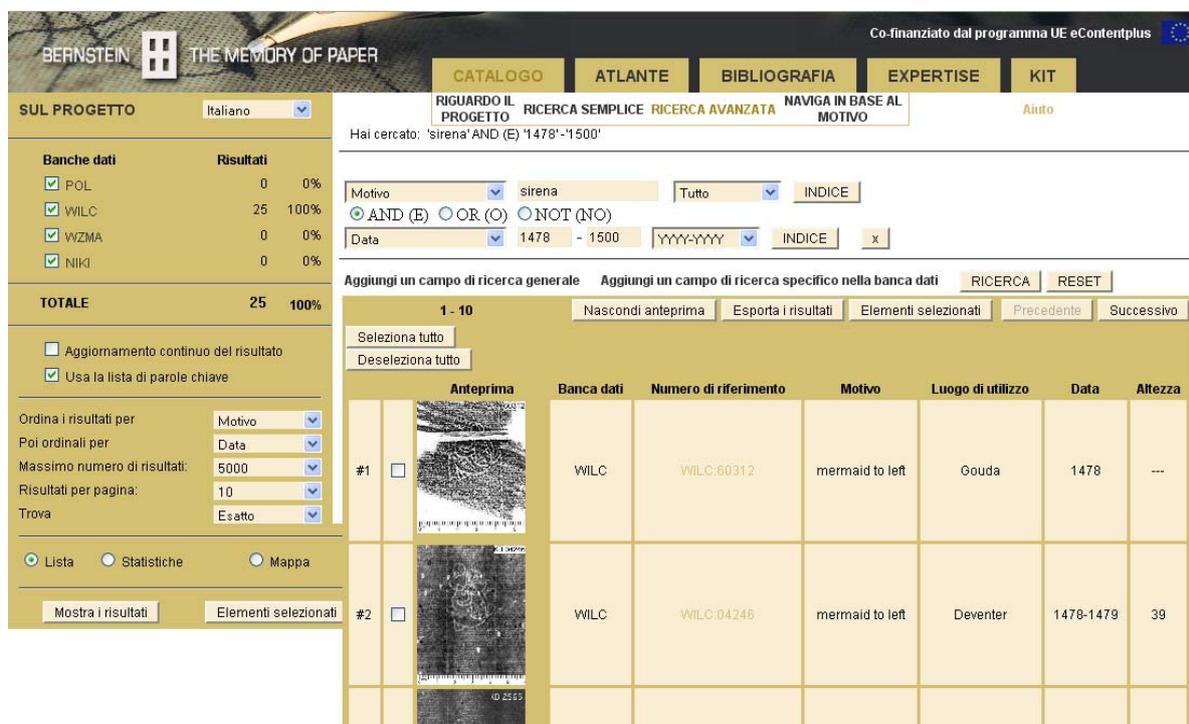


Figure 5: Catalogue (Italian), result list with thumbnails (motif=mermaid, date=1478-1500)

4.2.2 Result List

The major way to show the results of a search is a list which can be ordered by any of the data fields (motif, date, repertory, place of use, ...). This list shows: thumbnail (if show thumbnail is selected), database, reference number, motif, place of use, date, and height (figure 5).

In the result list, the user can bookmark individual entries by clicking the checkboxes at the beginning of each row or he can bookmark all shown results by clicking *Select all*. It is possible to show or hide the thumbnails for the watermarks and to export the results of the search as a file suitable for download.

4.2.3 Statistical Functionality

With well over 120,000 items in the combined databases of the project, it becomes necessary to provide means for visualising the statistical properties of the collection of watermarks. Statistics are nearly as important for historians and experts as the information about individual items. At present, except limited such functionality in the WZMA database, none of the original resources provides statistics of the contents – Bernstein reached a solution satisfying the expectations of users in this regard. A quantitative statistical description of the user's selection gives insights into the structure of the data and facilitates its interpretation.

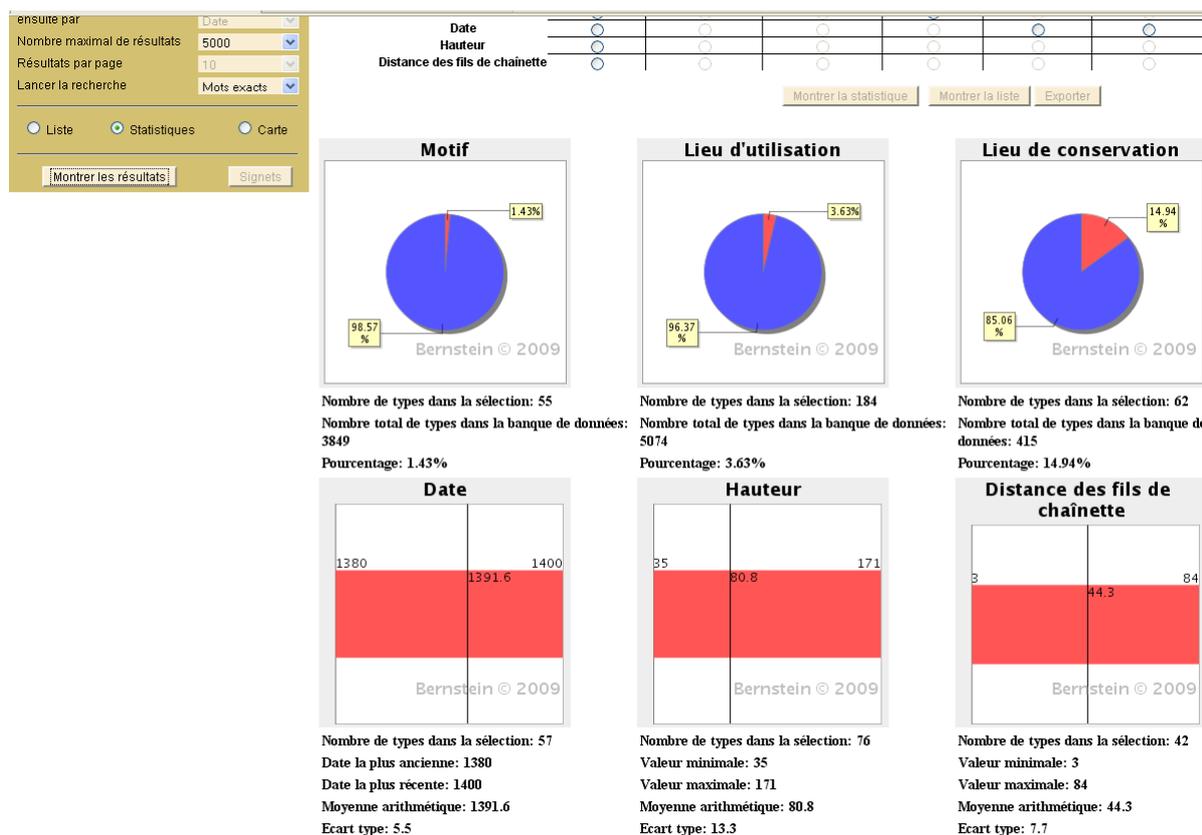


Figure 6: Statistics (French) – basic statistic parameters (motif=bull's head, date=1380-1400)

The statistics module offers a wide range of possibilities. Basic statistics parameters like mean value and standard deviation (figure 6) are displayed numerically and as a graphical representation in the form of a bar, pie or bubble diagram. Users can analyze single parameters (e.g., number of watermarks per year) or tie and inspect together two parameters (e.g., watermarks per year and country, figure 7).

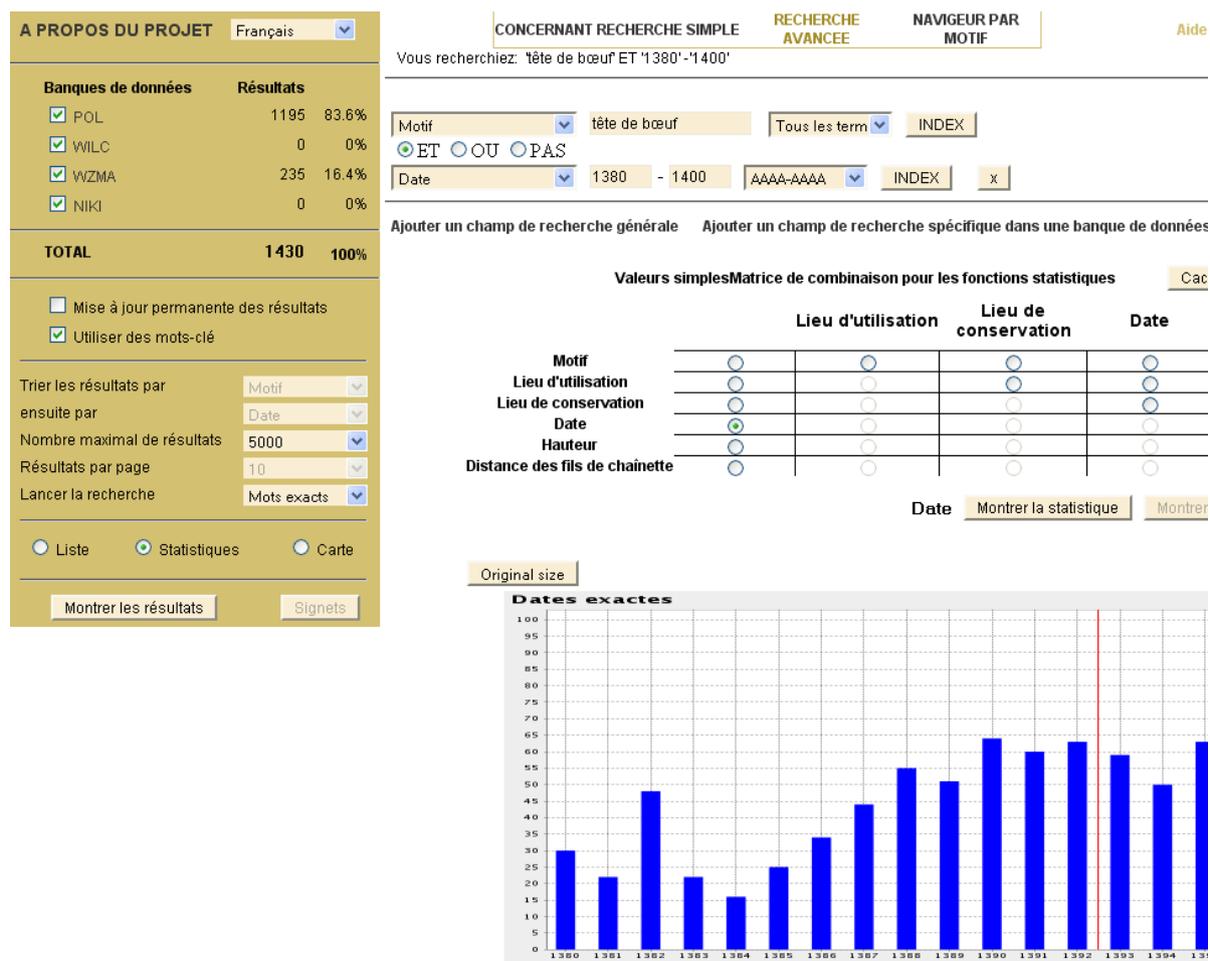


Figure 7: Statistics (French) – distribution of the motif bull’s head in the year 1380-1400

4.2.4 Atlas

The purpose of the *Bernstein Atlas* is to provide a tool for historical research based on representations of the distribution in space and time of watermarks and other paper characteristics stored in the Bernstein databases and supplemental datasets. Such a dataset of 14,000 place names representing approximately 7,500 unique and identified places with their geographical coordinates was generated. It contains all place names extant in the Bernstein databases, the digitized watermarks and papers repertoires, the Bernstein paper bibliography, the incunabula reference works *Incunabula Short-Title Catalogue* (ISTC, <http://www.bl.uk/catalogues/istc>), *Gesamtkatalog der Wiegendrucke* (GW, <http://www.gesamtkatalogderwiegendrucke.de>), and the plague dataset (see further down the “7.2 Contextual data” section). The Bernstein databases were augmented with geo-referenced data. This makes it possible to visualize the geographical distribution of a search result in *Catalogue* in regard of the place of use of the paper documents (figure 8, p. 11). The field *place of use* can be specified as city or region according to the European standard NUTS (Nomenclature of territorial units for statistics – NUTS Statistical Regions of Europe, http://ec.europa.eu/eurostat/ramon/nuts/splash_regions.html)

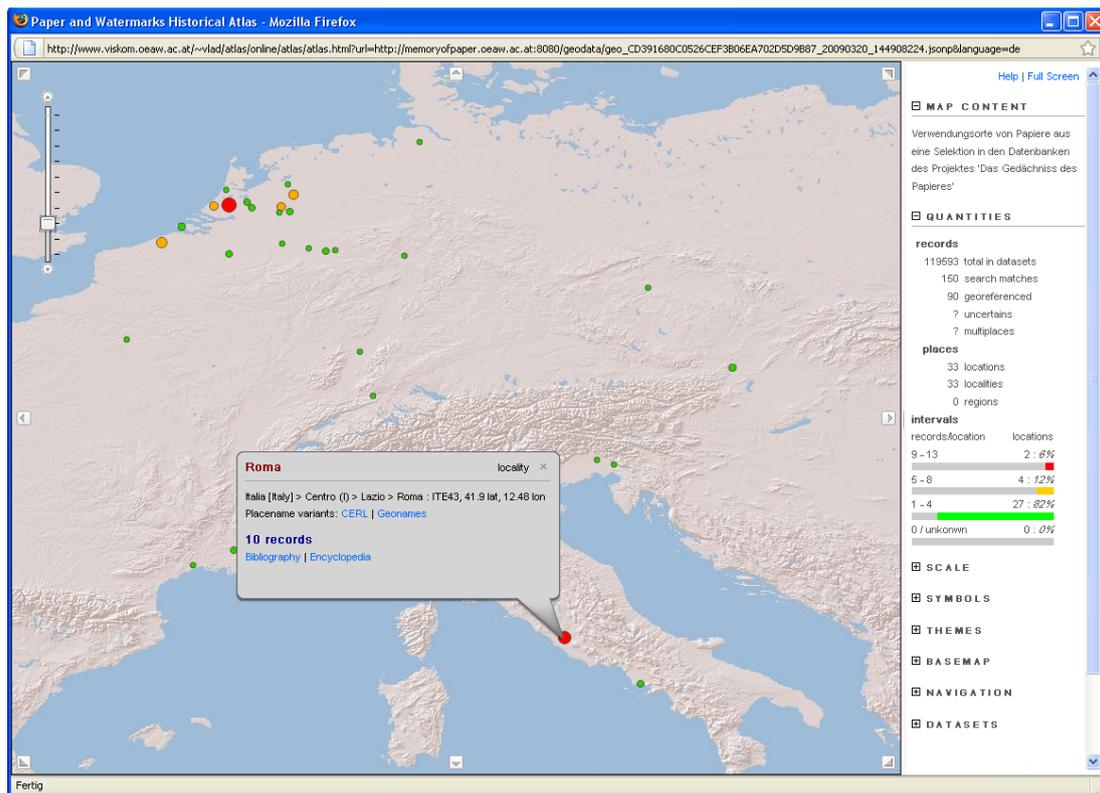


Figure 8: Atlas – place of use of paper with watermark motif mermaid

4.2.5 Paper Bibliography

Paper, its history, characteristics, trade, development, and use are of interest to scholars in the fields of history, codicology, bibliography, musicology, art history, and also for people active in paper conservation and forensic science. All over the world experts contribute to this domain of knowledge: they publish dictionaries, create watermark repositories, watermark handbooks, and publish articles in various periodicals. The publications in this domain as a whole represent the present level of knowledge. Unfortunately, many areas of this domain are ignored by scholars and other users—they fall out of focus due to a lack either of documentation or an easy access to that documentation.

The *Bernstein Bibliographic Database* contains a sizable subset of citations of publications across the above fields, written in many different languages. This material does not claim to be complete, and it has a slight bias towards German language publications. The data comes from the continuous documentation of the German Book and Writing Museum of the German National Library in Leipzig which accounts for above tendency. On the object level, the bibliographic database has been a multilingual project right from its beginning. Originally, classification, subject headings and geographical terms were described according to a German point of view only. As a further result of the Bernstein project, the structure of the bibliographic database was completely changed and it allows now a multilingual handling of above aspects.

The Bernstein interface to the bibliographic database gives the user the ability to retrieve records by both searching for as well as browsing specific terms. The user can search either across all fields, or specifically by title, author, location of print, year, shelf mark, bibliographic references, ISBN, subjects, and subject heading (separate search masks for each motive, location, person and corporation). The browse interface enables the look-up of terms stored in the database without prior knowledge by browsing the indices.

4.2.6 Expertise

The goal of the expertise activities in project Bernstein was to provide methods and software for measuring watermark and paper characteristics and to facilitate the dating of paper documents. Such a tool is the program AD751 performing laid line measurements. The user has the choice between a standalone and an online version. At the moment, three further tools can be downloaded for local use. PAT is the Paper Analysis Tool. It detects the chain and laid lines in a digital image of a piece of paper and calculates their number and average distance. AWDT is the Automatic Watermark Detection Tool. It detects automatically watermarks in X-ray and backlight images. QET is the quality enhancement tool. It removes noise from X-ray and backlight images in such a way that paper features become better visible. Further software tools are in development and will be added in future.

4.2.7 Dissemination kit

The so called '*dissemination kit*' plays an important role for the sustainability of the project. It is the realization of a case study where a user has a collection of watermarks and wishes to create their own watermark database. The dissemination kit is a ready-to-use tool set that gives people the means to set up and operate their own paper study services in a very short time and with little effort. It is provided as a downloadable pack containing data, software, and documentation. A user-friendly installation package was developed and allows users the easy installation and initialization of a watermark database compliant with the Bernstein standards and immediately linkable with all other Bernstein databases. The distribution is as a Microsoft Windows Installer (MSI) package. The responses and feedbacks to the Bernstein project showed a massive demand for the dissemination kit. It ensures that the quantity of data in Bernstein will grow permanently and thus guarantees the sustainability of Bernstein.

5 Target Users and Their Needs

The project is characterized by much diversified sources of demand for the functionalities it offers. Hereafter we describe who the user communities are, what their specific needs are, and what is the critical level needed to satisfy these requirements.

5.1. Cultural demand/Historians – The original initiative for the project originated with historians wishing to study the culture of Europe at the onset of the Renaissance by means of paper studies. Following aspects to all of which our project brings solutions are of utmost importance for the successful work of historians: identification of date and place of origin of paper based on objective measurements, statistical and historical cartography of paper characteristics in order to map paper technology and trade in space and time, and, finally, the capability to do paper history in the broader contexts of European culture and societies.

5.2. Economic demand/Curators and Industry – For art dealers the ability to correctly authenticate, date and locate paper documents is the fundamental basis on which their business relies. The same is true for curators of public collections (libraries, archives, museums), who need to know the authenticity and value of objects they possess or wish to acquire or sell.

5.3. Societal demand/Forensic experts – Paper documents such as passports are the most widely used identification documents for individuals across the world. The ability to quickly and surely identify fakes can prevent crimes and facilitate law enforcement. Paper expertise by police agencies is done however not at the place where the paper documents are examined (borders, police checks, banks, ...), but rather upstream in the chain of criminal expertise, by experts in laboratories. They rely mostly on their experiences and less on machines. A clear need for paper analysis software and networked databases for authenticating documents exists here.

5.4. Industrial demand/Paper makers – Together with curators, paper manufacturers (artisans and big companies) provide specialized papers reproducing ancient paper characteristics for restoration of damaged cultural objects. From the art historian's point of view it is important that their products match closely the old originals and that confusion about where and when a particular type of paper did occur be avoided in order not to produce inadequately restored objects. Therefore knowledge of paper history and cartography is crucial for paper makers.

6 Underlying Content

Our resources consist of content (images, metadata and contextual resources) and content processing software (for measurements, data mining, statistics and cartography).

6.1. *Images* – The images are reproductions of the physical structure of papers through several techniques: radiography, backlight, rubbing and tracing. Several features are made visible: the sieve of the paper mould (watermarks, chain and laid lines), parts of the wooden frame and the paper pulp distribution. The measurements of the variation of each of these features make a unique identifier for each paper sheet and mould and more generally for the paper mill and the know-how of a region or time period. Thus the identification of the origins of paper can be established and by comparison of watermarks (which were specific for individual paper makers and replaced every 2 to 4 years), paper documents before the 17th century (the age of manually produced papers) can be dated with a precision of ± 9 months (by intersecting the watermark date ranges of the several paper batches usually composing a manuscript, book or newspaper).

Quantity: In total we have about 120,000 images distributed among four on-line databases: LABW, Germany (91,750), KB, Netherlands (16,000), OEAW, Austria (9,550) and NIKI, Italy (2,200). This covered quasi all digital primary resources on paper studies available in the world at the onset of the project in 2006.

Quality: Our collections represent the reference material for historical studies on paper and watermark expertise. LABW provides the entire “Piccard” repertory of watermark tracings, a monumental work that in the print version spans 25 volumes and covers the watermarks used in the whole of Western Europe, from the 13th to the 19th century. Although other watermark repertories exist, none is equal in importance to Piccard and no other is digitized or expected to be so in the near future. KB brings its own impressive collection which documents by electron radiography and rubbing paper types of all incunabula printed in the Low Countries (the first books produced after the invention of print with movable types by Gutenberg). OEAW’s (KSBM) collection is yet from another area, recording by X-rays Austrian manuscripts of the late Middle Ages. NIKI’s special contribution to the project consists in outstanding reproductions of art drawings and prints of such key figures of European culture as Leonardo da Vinci and Rembrandt.

6.2. *Metadata* – The images are backed in the databases by metadata providing information first of all about the measured characteristics of the paper and also the classification of the watermarks. Information about the date and place of production of the papers and about the documents for which they were used (e.g. for books: title, author, publisher name, and date).

6.3. *Contextual resources* – Besides content pertaining strictly to paper, we provide contextual data that helps advance the study of paper. This is provided by DNB in form of the most recent and complete bibliography on paper studies ever collected. The paper data and bibliography integration of the last two volumes of this work are part of our project achievements. We mention also another contextual resource, that is not owned by our partners, but for which we provide interoperability: the Incunabula Short-Title Catalogue (ISTC) of the British Library, the complete online catalogue of all incunabula printed in Europe (29,244 editions).

6.4. Content processing software – The processing software for the expertise infrastructure is provided by OEAW and DUT. These are:

a) pre-processing software for improving the image quality for human and machine examination (BlueNile for image filtering in the frequency domain) and for removing ink traces from backlight reproductions;

b) measurement tools for chain and laid lines density evaluation (AD751 and Rembrandt);

c) paper comparison tools based on the sieve structure (Rembrandt);

d) paper dating experimental tools based on watermark similarity,

e) LAMOP's tools for historical cartography and statistical aspects made operable in the project by upgrading the databases undertaken by their owners. LAMOP is pioneering historical cartography for paper studies.

7 Summary of Activities

7.1 Data and Data Integration

Databases – The core of the project is the integration of several databases. This activity turned out a much harder task than expected. All four databases differ technically and logically. The first step was to harmonize the terminology used to describe the paper characteristics across the databases. The result of this harmonization was at first a vocabulary of watermark terms in three languages (English-German-French). This vocabulary for watermark description has been augmented by Italian, Russian and Spanish since then and is now available in six languages.

The second step in the harmonization was the development of a common classification scheme for watermarks. The existing motif groups of the relevant databases were unified into a Bernstein classification scheme. This classification scheme is not only an important part of the multilingual access, but also necessary for other elements of the Bernstein workspace, such as the dissemination kit. A first version of this classification scheme has been agreed upon and it describes 12 main motifs on the first level and further motifs on the second and third levels. Further levels are in preparation.

Integrated workspace – A LINUX server was installed at the Austrian Academy of Sciences in order to host all online resources of Bernstein, especially the integrated workspace (Bernstein web portal). This server will continue to operate for at least five years after the project's end.

The integrated workspace was developed further and includes now also a powerful statistics and a cartography module. The free Java chart library “JFreeChart” was used for creating the complex statistical charts in the workspace. It has a flexible design and is easily extendible for fulfilling special demands. The industry standard software ArcGIS was installed on the server to support the Bernstein Atlas and its integration into the workspace. A web designer was employed in order to improve the ergonomics, readability, and design of the Bernstein workspace.

Numerical paper description standards – Papers without watermarks or with not accurately assignable watermark motifs are difficult to identify within the present databases. Art historians, archivists, and musicologists are however all dealing with papers which predominantly do not contain watermarks at all or contain only parts of watermarks. Therefore, one of the goals of the Bernstein project was to find additional paper parameters, which allow the comparison of papers with and without watermarks if necessary. Direct dating of paper only on the basis of its structure seems impossible. Instead, it appears necessary to find sufficient corresponding paper parameters in order to relate undated paper with dated one. Such investigations were summarized by the term ‘numerical description of paper’ in the Bernstein project. In this context, existing software for numerical paper description was adapted and tested with selected examples.

Numerical parameters in the databases such as height and width of watermarks, or chain line distance were integrated as search criteria into the Bernstein portal. Existing image analysis software calculating numerical paper parameters was adapted for tests in the Bernstein project and is now freely available for download. No single program able to handle all types of recordings in the Bernstein databases exists. In order to allow for future developments of virtual mould reconstructions and data mining in paper and watermarks databases, it is recommended to record the full area of the paper object and not only the area surrounding the watermarks.

7.2 Contextual data

This activity segment aimed at enriching already existing data in the Bernstein databases through web releases, reorganization, metadata and contextualization.

(1) *Web release* – The German National Library in Leipzig (DNB) compiles the most comprehensive bibliography on all aspects of paper. The bibliography contains currently more than 30,000 bibliographic records. The bibliographic database of the DNB cannot be accessed directly by the Bernstein workspace due to technical and organisational reasons. Therefore the DNB exports their existing Allegro database into a standardized XML version for interoperability and transfers the data to the Bernstein server. The bibliography is stored in the centrally located Bernstein server as a Cheshire3 database (<http://www.cheshire3.org>). This is the third generation of the Cheshire system started more than 10 years ago at UC Berkeley and recently further developed in a partnership between Berkeley and the University of Liverpool. Cheshire platforms are used by several national services in Europe as well as by several services and projects in the United States.

(2) *Reorganisation* – a) Concerned was the Piccard-Online database, which represents three quarters of the Bernstein contents. This resource is based on a work done in the mid-20th century for the print media, that didn't require the consistency and explicitness necessary for digital databases. For example person names might have different spellings and social status might be obvious to historians, yet not mentioned. The task consisted in making the content of the database fields consistent across the 90,000+ records (ex.: “Kg. Maxililian” > “Maximilian I., Kaiser des H.R.R.”), splitting fields with multiple information types in several individual fields (ex.: 1. “Function: Emperor”, 2. “Name: Maximilian I.”), and adding explicit data (ex. for Maximilian: “Authority: secular”). – b) “Bull’s head” watermarks. 5.000 out of 25.000 watermarks of the type “bull’s head” in Piccard-Online have been classified in equivalence groups, according to their similarity. It enables historians to compare various historical factors, such as industrial production areas, trade routes, consumption patterns, etc., through the means of watermarks. It provides also a supplementary database navigation method, between watermarks of the same equivalence group. This method was introduced by the developer of the partner database WILC and acclaimed much by the users.

(3) *Metadata* – a) Georeferences: place names from the Bernstein and related datasets were attributed with metadata concerning the geographical coordinates and administrative units to which they belong. See the section on the Atlas above for details. – b) Incunabula identification: a total of 799 incunabula editions were identified in Piccard-Online (564 editions from 4,000+ watermarks) and Briquet (235 editions for 400+ printed books). Each record was supplemented with a field giving the incunabula identification number in the world reference work on the subject, the ISTC, the URL in the ISTC and URLs to images of the book. – c) Repositories identification: the current repositories where the incunabula are kept were provided in digital form based on the printed reference on the matter, the GW.

(4) *Contextualization* – a) Incunabula authors: a dataset was prepared containing biographical information about the 3,500+ authors of the 28,000 incunabula. The information is structured to allow its use in digital databases. It provides historians with quantitative material that connects information on paper with information on persons, social and cultural environments. – b) Plague: along the same line of thought, a digital dataset of locations where the plague occurred in Europe during the Middle Ages and the Renaissance was produced. The dataset is part of the Bernstein geo-references.

7.3 Dissemination

(1) *General public* – Statistics for the project website usage show a constant increase since its launch in August 2006 – an encouraging testimony of the interest in the project and the subject it promotes: paper studies (<http://www.bernstein.oeaw.ac.at/twiki/bin/view/Main/WebStatistics>). The travelling exhibition *Bull's Head and Mermaid – The History of Paper and Watermarks from the Middle Ages to the Modern Period* turned into the most effective dissemination activity of Bernstein. It was held in five cities, in Germany, Austria, and Italy, and a sixth exhibition will be opened in Torino (Italy) at the end of April 2009. Further exhibitions are planned in Austria, Germany, France, the Netherlands, Slovenia, and Spain. The exhibition comes with a catalogue in English, German, and Italian. This catalogue stands for its own as a technical guide and a scientific book about all aspects of watermarks, watermark collections and the project Bernstein.

(2) *Scientific media* – Project participants contributed scientific papers to specialized journals in Humanities and Sciences and participated in congresses. See the list at <http://www.bernstein.oeaw.ac.at/twiki/bin/view/Main/ProductsDissemination>.

(3) *Contacts* – The principal collaboration contacts with partners outside the project during the past years were – a) with the Fabriano Paper and Watermarks Museum for the sieve experiment and creating the Zonghi online database; – b) with VirginiaTech University for inter-linking, harmonizing and developing the Briquet printed and archive databases; – c) with the British Library and the Staatsbibliothek zu Berlin for adding contextual data and georeferences to ISTC and possibly GW; – d) with the Electronic Cultural Atlas Initiative (ECAI, <http://www.ecai.org>) for connecting the Bernstein GIS with ECAI datasets; – e) with the Laboratorio de Restauración of the Universitat de València for the Spanish terminology and thesaurus; – f) with the State Historical Museum of Russia, Moscow for the Russian translation and watermark terms vocabulary; – g) with Istituto Centrale per il Restauro e la Conservazione del Patrimonio Archivistico e Librario in Rome for the Italian terminology and the watermark vocabulary; – h) with the Statens Museum for Kunst, Copenhagen for a new project application; – and i) with the Bates College, Lewiston, Maine, for a joint project – and many more.

7.4 Management

Main aspects of the project management were meetings between partners and with contacts outside the project. This was partly supported by electronic communication means: a public project webpage, a development website open to the partners and a mailing list.

8 Impact & Sustainability

8.1 Impact

The project is expected to have a considerable impact on paper studies, its “target market”, by facilitating and bringing innovation in the field and broadening and synergising the area.

Facilitating role. Obviously the creation of digital resources (the paper databases, image processing software and geographical and bibliographical contextual resources) will facilitate a work until now done “manually”, improving the speed and breath of research (in excess of 120,000 records accessible with one query).

Innovation role. By integrating content (the paper databases) with content processing tools (the image processing, cartography and bibliography software) existing data can be used in ways not possible before (for example the ability for measuring paper features provided by image processing and the statistical information generated by paper databases make the dating of paper documents possible).

Broadening and synergising role. The individual data resources, software and know-how having converged into the project were initially intended for specific user communities (databases for historians, image processing for experts, and integration capacity at partner TUG for applications outside the Humanities). The project enabled each one to broaden its reach and become valuable for new users. Conversely, the polyvalence of the projects products has the potential to create synergies between user communities.

8.2 Sustainability

The lasting and evolving of the projects achievements is based on the credit of the consortium partners and spin-in & spin-off effects.

Credit: The partners who are the real holders of the project’s digital products are in their countries institutions of national importance (national libraries, academies...). This is one of the ways how the consortium is ensuring the continuity of its work.

Spin-in: The project has already created an effect of ‘spin-in’ by which new parties (such as database holders) got interested in joining the effort. While some collaboration took part in the framework of the project, others need a substantial allocation of resources and thus represent objectives of new potential projects (for example integrating 31 non-Bernstein paper databases created or identified during the project). This is only possible if the project’s outcome is maintained active.

Spin-out: In the context of the project ‘spin-out’ activities refer to contacts between the primary users (paper historians and experts) and users from untargeted fields. For example, the linking of a watermark database (WILC of the Koninklijke Bibliotheek, The Hague) and an incunabula database (ISTC of the British Library) is a by-product seen as the first step in integrating Bernstein into other European cultural assets such as *The Europeana* (<http://www.europeana.eu>). Again, these development opportunities are also a guarantee for the lasting of the Bernstein endeavours.

Migration: The time after the end of the project will come when the software used to provide the services developed now will become obsolete. In order to maintain these services, migration of data and software upgrade will be necessary. To ensure that the project partners store the data in non-proprietary, unencrypted, well documented and human readable formats (XML for the KB) or use database software that can export data to such formats (MySQL and similar software for the other partners). As for the software, the solution adopted was to make the source code publicly available under an Open Source licence.

9 Further Information

In addition to the information provided in this report the reader might find of interest to consult the following resources:

- The Bernstein website (<http://www.memoryofpaper.eu>) which is the Internet ‘face’ of the project and gives access to its ‘body’, the integrated workspace.
- The Bernstein project presentation as a downloadable Microsoft PowerPoint file (http://www.bernstein.oeaw.ac.at/Bernstein_project_presentation.ppt)
- The TWiki development platform (<http://www.bernstein.oeaw.ac.at/twiki>) serving as a common blackboard and document repository for the project’s partners.
- The itinerant exhibition, for learning more about watermarks (check locations and order the catalog at <http://www.bernstein.oeaw.ac.at/twiki/bin/view/Main/ProjectExhibitions>).

Abbreviations

BH-GIS	Bernstein Historical Geographical Information System
BL	The British Library, London, United Kingdom
CERL	The Consortium of European Research Libraries
DNB	Deutsche Nationalbibliothek, Leipzig, Germany
DUT	Delft University of Technology, Delft, Netherlands
ECAI	Electronic Cultural Atlas Initiative
GIS	Geographical Information System
GW	Gesamtkatalog der Wiegendrucke (State Library, Berlin)
KB	Koninklijke Bibliotheek, The Hague, Netherlands
KSBM	Commission for Paleography and Codicology of Medieval Manuscripts, OEAW
IPB	International Paper Bibliography
ISTC	Incunabula Short Title Catalogue (by BL)
LABW	Archives of the State of Baden-Württemberg, Stuttgart, Germany
LAMOP	Laboratory for Occidental Medieval Studies in Paris, Paris, France
LU	Liverpool University, Liverpool, United Kingdom
NIKI	Dutch University Institute for Art History Florence, Florence, Italy
OEAW	Austrian Academy of Sciences, Vienna, Austria
TUG	Technical University Graz, Graz, Austria
VISKOM	Commission for Scientific Visualization, OEAW
WILC	Watermarks in Incunabula printed in the Low Countries (by KB)
WZMA	Watermarks of the Middle Ages (by OEAW (KSBM))