

## **ECP 2005 CULT 038097/Bernstein**

### **Bernstein**

# **Deliverable no. 2.2, GIS data batch no. 1**

<b>Deliverable number</b>	<i>D2.2</i>
<b>Dissemination level</b>	<i>Public</i>
<b>Delivery date</b>	<i>31 July 2007</i>
<b>Status</b>	<i>Final</i>
<b>Author(s)</b>	<i>The Bernstein Consortium</i>



***eContentplus***

This project is funded under the *eContentplus* programme<sup>1</sup>,  
a multiannual Community programme to make digital content in Europe more accessible, usable and exploitable.

---

<sup>1</sup> OJ L 79, 24.3.2005, p. 1.

## 1. Background

---

The goal of the Historical Geographical Information System of Project Bernstein (BH-GIS) is the representation across space and time of the distribution of watermarks and other paper features stored in the Bernstein databases. Additionally, Bernstein should be able to provide bibliographical data incorporated in the maps. The first step towards the GIS is the generation of a list of places occurring in the Bernstein databases matched to geographical coordinates (geo-list). This is the content of the deliverable D2.2 'GIS data batch no. 1' presented here. The webpage on GIS aspects on the Bernstein development site is <http://www.bernstein.oeaw.ac.at/twiki/pub/Main/PaperCartography/>.

## 2. Work accomplished

---

The work accomplished so far consists in a matched list of place names, coordinates and additional metadata. The list is explained hereafter and the process and issues generating the list described further down.

### 2.1 Data sources

The processed geographical data belong to the following sources:

- A. Piccard-Online watermarks and paper database (POL, <http://www.landesarchiv-bw.de/piccard/>);
- B. Briquet printed catalog of watermarks and paper (BPC);
- C. Incunabula Short-Title Catalog, quasi complete (short) list of existent incunabula (ISTC, <http://www.bl.uk/catalogues/istc/>);
- D. Gesamtkatalog der Wiegendrucke, quasi complete (extended) list of existent incunabula (GW, <http://www.gesamtkatalogderwiegendrucke.de/>);
- E. International Paper Bibliography (IPB, <http://www.ddb.de/sammlungen/pdf/ibp.pdf>).

This represents 8.000 records, equivalent to approximate 3.000 individual localities.

### 2.2 Description of the geo-list

The core information provided by the geo-list are name of places as they will appear on the Bernstein maps and their respective geographical coordinates. Additional geographical information is given about variant names of the places and various levels of administrative units to which the places presently belong. The metadata identifies the source where the place names was found and records wherever an ambiguity surrounding the exact location which is meant by the source was found. Also given is information about wherever the place name given by the source refers to the place where the paper was produced (paper mill), where it was used (printing house, scriptorium), or where it is currently conserved (library, archive...).

The next two pages show the beginning of the geo-list. A digital version can be downloaded from [http://www.bernstein.oeaw.ac.at/twiki/pub/Main/PaperCartography/BH-GIS\\_data\\_sample.xls](http://www.bernstein.oeaw.ac.at/twiki/pub/Main/PaperCartography/BH-GIS_data_sample.xls). The fields read as follows:

- A. *Source id nr.* – Number of the record in the source from were the place name was found.
- B. *Source* – Short name of the source from were the place name was found.
- C. *Information type* – Records if the place name is that of the where the paper was produced, where it was used, or where it is currently conserved.
- D. *Accepted from GW* – Records if coordinates could be attributed to the place name found in source GW.
- E. *Accepted from POL* – Records if coordinates could be attributed to the place name found in source Piccard.
- F. *Accepted from BPC* – Records if coordinates could be attributed to the place name found in source Briquet.
- G. *Acceptance status* – Records if the person processing the record has found the place name to be ambiguous (ex.: several places with the same name) or couldn't locate it geographically.
- H. *Name in the source* – Name of the place as given in the source.
- I. *Normalized name* – When several names are available for the same location, one name, is chosen to be displayed on the Bernstein maps.
- J. *Latitude* – Decimal latitude coordinate of the place.
- K. *Longitude* – Decimal longitude coordinate of the place.
- L. *Country* – Country in which the place is located at the present time.
- M. *Admin region 1* – 'First order' administrative unit (such as 'Land' in Germany) in which the place is located at the present time.
- N. *Admin region 2* – 'Second order' administrative unit (such as 'county' in England) in which the place is located at the present time.

		A	B	C	D	E	F	G	H	I	J	K	L	M	N
Bernstein Geographical Data on Watermarks and Paper features (Sample)															
1	2	Accepted from:													
		Acceptance													
3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
id nr.	Source	Place type	GW	POL	BPC	status	Name in the source	Normalized name	Lat	Lon	Country	Admin region 1	Admin region 2	Admin region 1	Admin region 2
1	1 GW	incunabula location	y				's-Gravenhage	's-Gravenhage	51,52	5,92	NL	Zuid-Holland			
2	953 GW	edition place	y				Den Haag	's-Gravenhage	51,52	5,92	NL	Zuid-Holland			
3	2260 GW	incunabula location	y				's-Heerenberg	's-Heerenberg	51,88	6,25	NL	Gelderland			
4	2 GW	incunabula location	y				's-Hertogenbosch	's-Hertogenbosch	51,70	5,32	NL	Noord-Brabant			
5	4578 Piccard Findbuch	place of paper use		n			Aach	Aach			D	Baden-Württemberg	Freiburg		
6	3 GW	incunabula location	y				Aachen	Aachen	50,46	6,06	D	Nordrhein-Westfalen	Köln		
7	4580 Piccard Findbuch	place of paper use	y				Aalen	Aalen	48,83	10,10	D	Baden-Württemberg	Stuttgart		
8	4 GW	incunabula location	y				Aalst	Aalst	50,56	4,02	B	Oost-Vlaanderen			
9	6 GW	incunabula location	y			y	Aarau	Aarau	47,38	8,05	CH	Aargau	Aargau		
10	4582 Piccard Findbuch	place of paper use	y			y	Aardenburg	Aardenburg	51,17	3,27	NL	Zeeland	Zeeland		
11	7 GW	incunabula location	y				Aarhus	Aarhus	56,15	10,22	DK	Århus Amt			
12	4583 Piccard Findbuch	place of paper use		y		y	Abbach	Abbach	48,93	12,05	D	Bayern	Oberpfalz		
13	3104 Briquet	place of paper use					Abbaye de Bonne-Espérance	Vellerelle-les-Brayeux							
14	4584 Piccard Findbuch	place of paper use		n		not identified	Abbaye de Lyrn								
15	3105 Briquet	place of paper use				y	Abbaye de Saint-Jacques	Abbaye de Saint-Jacques	50,38	5,34	B	Liège			
16	9 GW	incunabula location	y				Abbeville	Abbeville	50,06	1,50	F	Picardie	Somme		
17	3107 Briquet	place of paper use				n	Abdinghof				D				
18	10 GW	incunabula location	y				Aberdeen	Aberdeen	57,13	-2,10	UK				
19	11 GW	incunabula location	y				Aberystwyth	Aberystwyth	52,42	-4,07	UK				
20	4585 Piccard Findbuch	place of paper use		y		y	Abiate	Abbiategrosso	45,40	8,90	I	Lombardia	Milano		
21	2522 GW	incunabula location	y				Åbo	Turku	60,45	22,28	FI	Lansi-Suomen			
22	4586 Piccard Findbuch	place of paper use		n		ambiguous (2)	Åbo				DK	Århus	Århus		
23	3108 Briquet	place of paper use				n	Abstberg								
24	3109 Briquet	place of paper use				y	Acey	Vitreux	47,14	5,41	F	Franche-Comté	Jura		
25	12 GW	incunabula location	y				Achel	Achel	51,15	5,29	B	Limburg			
26	4587 Piccard Findbuch	place of paper use		y		y	Achern	Achern	48,63	8,07	D	Baden-Württemberg	Freiburg		
27	13 GW	incunabula location	y				Acireale	Acireale	37,62	15,17	I	Sicilia	Catania		
28	4588 Piccard Findbuch	place of paper use		n		ambiguous (5)	Acken								
29	14 GW	incunabula location	y				Acquapendente	Acquapendente	42,73	11,87	I	Lazio	Viterbo		
30	15 GW	incunabula location	y				Adelaide	Adelaide	in area	in area	AU				
31	4589 Piccard Findbuch	place of paper use		n		ambiguous (2)	Adelberg	Adelmannsfelden	48,80	9,60	D	Baden-Württemberg	Stuttgart		
32	4590 Piccard Findbuch	place of paper use		y		y	Adelmannsfelden	Adelmannsfelden	48,57	10,01	D	Baden-Württemberg	Stuttgart		
33	4591 Piccard Findbuch	place of paper use		n		not identified	Adelsbach								
34	4592 Piccard Findbuch	place of paper use		y		y	Adelsheim	Adelsheim	49,40	9,39	D	Baden-Württemberg	Karlsruhe		

		A	B	C	D	E	F	G	H	I	J	K	L	M	N	
Bernstein Geographical Data on Watermarks and Paper features (Sample)																
Accepted from: Acceptance																
Source																
id nr.	Source	Place type	GW	POL	BPC	status	Name in the source	Normalized name	Lat	Lon	Country	Admin region 1	Admin region 2			
38	16 GW	incunabula location	y				Admont	Admont	47,34	14,27	AT	Steiermark				
39	17 GW	incunabula location	y			ambiguous (3)	Adria	Adria	45,06	12,06	I	Veneto			Rovigo	
40	3111 Briquet	place of paper use			n	ambiguous (3)	Adria									
41	4593 Piccard Findbuch	place of paper use		y			Aelst	Aalst	50,46	4,52	B	Vlaams-Brabant			Vlaams-Brabant	
42	3112 Briquet	place of paper use			y		Aerdenbourg	Aardenburg	51,17	3,27	NL	Zeeland			Zeeland	
43	3113 Briquet	place of paper use			n	not identified	Aeren									
44	3114 Briquet	place of paper use			y		Aernen	Aernen	46,40	8,15	CH	Valais				
45	4594 Piccard Findbuch	place of paper use		n		not identified	Afflengt									
46	18 GW	incunabula location	y				Affligem	Affligem	50,50	4,20	B	Vlaams-Brabant				
47	19 GW	incunabula location	y				Afflingen	Affligem	50,50	4,20	B	Vlaams-Brabant				
48	20 GW	incunabula location	y	y			Agen	Agen	44,22	0,63	F	Aquitaine			Lot-et-Garonne	
49	21 GW	incunabula location	y				Agira	Agira	37,65	14,52	I	Sicilia			Enna	
50	22 GW	incunabula location	y				Agnone	Agnone	41,80	14,37	I	Molise			Isernia	
51	2772 GW	incunabula location	y				Agram	Zagreb	45,80	16,00	CR	Grad Zagreb				
52	23 GW	incunabula location	y				Agrigento	Agrigento	37,32	13,57	I	Sicilia			Agrigento	
53	4596 Piccard Findbuch	place of paper use		n		not identified	Ahanau									
54	4597 Piccard Findbuch	place of paper use		n		ambiguous (6)	Ahausen									
55	4598 Piccard Findbuch	place of paper use		n		not identified	Ahirch									
56	4599 Piccard Findbuch	place of paper use		n		ambiguous (3)	Ahlen									
57	4600 Piccard Findbuch	place of paper use		y			Ahus	Ahus	51,75	7,92	D	Baden-Württemberg			Tübingen	
58	4601 Piccard Findbuch	place of paper use		n		ambiguous (31)	Aich									
59	4602 Piccard Findbuch	place of paper use		n		ambiguous (2)	Aichach									
60	4603 Piccard Findbuch	place of paper use		n			Aichbach									
61	4604 Piccard Findbuch	place of paper use		y			Aichbach	Aichbach	48,28	11,08	D	Bayern			Schwaben	
62	3117 Briquet	place of paper use		n		ambiguous (3)	Aichelberg									
63	25 GW	incunabula location	y			ambiguous (2)	Aichstätten									
64	3118 Briquet	place of paper use			y		Aidona	Aidone	48,27	9,38	D	Baden-Württemberg			Stuttgart	
65	3119 Briquet	place of paper use			y		Aigle	L'Aigle	37,42	14,45	I	Sicilia			Enna	
66	3120 Briquet	place of paper use			y		Aigueperse	Aigueperse	48,75	0,63	F	Basse-Normandie			Orne	
67	3121 Briquet	place of paper use			y		Airasca	Aigueperse	46,02	3,20	F	Auvergne			Puy-De-Dôme	
68	26 GW	incunabula location	y		n	not identified	Aisey-Le-Duc	Airasca	44,92	7,48	I	Piemonte			Cuneo	
69	29 GW	incunabula location	y		y		Aix-en-Provence	Aisey-sur-Seine	47,75	4,58	F	Bourgogne			Côte-d'or	
70	3123 Briquet	place of paper use		y			Aix-La-Chapelle	Aix-en-Provence	46,30	23,72	RO	Alba				
71	30 GW	incunabula location	y		y		Ajaccio	Aix-en-Provence	43,32	5,26	F	Provence-Alpes-Côte [	Bouches du Rhône			
							Ajaccio	Aachen	50,46	6,06	D	Nordrhein-Westfalen			Köln	
							Ajaccio	Ajaccio	41,55	8,44	F	Corse			Corse-du-Sud	

## 2.3 Generation process of the geo-list

The generation of the geo-list delivered here was not straightforward: it implies several steps in a lengthy process, as detailed hereafter.

1. *Extract place names* – Various methods have been used to extract the data, from a simple download of the appropriated field in the case of digital databases (IPB) to typing in the computer list the information found in a printed source (BPC).
2. *Identify locations* – Using various digital and printed maps and atlases the place names given in the sources had to be identified. For each record several sources were checked each time to minimize the risk of excluding homonym places.
3. *Disambiguation* – In many cases the same place name is given to distinct location ('Paris' in France and 'Paris' in Texas, USA). Following rules have been followed to decide to which location to attribute the place name (if possible):
  - a. *Additional data in the source* – Sometimes the source provides clues to the location of the place name, such as mentioning a nearby city or river.
  - a. *Geographical likelihood* – A bigger city is more likely to be meant than a smaller city.
  - a. *Historical likelihood* – Paris, Texas, USA is not a likely location of a place name of a book printed during the time of the Renaissance.
4. *Find coordinates* – Once it was established which location the source means, its geographical coordinates had to be found using Internet and printed resources. Giving the various coordinates systems used by the reference materials, some had to be converted into the decimal system used in this geo-list.
5. *Attribute administrative units* – Providing the administrative unit to which a place belongs allows doing geographical statistics by clustering point-locations into areas. The overview thus gained helps the historian in his investigative work of the past. This is the reason why along with the coordinates information on the administrative units was provided.

Because administrative units are dynamical historical entities – created, abolished, their boundaries changed – it was decided to use the administrative units as they are at the present time.

## 2.4 Amount of data the geo-list

The geo-list contains 3.000 place names ordered alphabetically. 250 (A-Brest) have been processed, i.e. name-normalized, matched with coordinates, attributed to administrative units, place type identified and given identification references.

There are two reasons why not more data was processed: work on the GIF functionality is scheduled for the second year of the Bernstein project and much time was spent in extracting the place names from the sources.

The remaining work (Brest-Z) is estimated at one month for one person.

## 3. Collaborations initiated

---

1. *ECAI* – Through participation in an Electronic Cultural Atlas Initiative (ECAI) conference, contact has been established with an international community active in the area of historical GIS research and development. The core of this community is the so-called 'clearing house' (<http://ecaimaps.berkeley.edu/clearinghouse/>), a software that interconnects various projects with geographical content so that datasets from a project can be used by other projects. The Bernstein GIS developing team intends to take advantage of this opportunity and contribute with its work to the geographical information environment managed by ECAI.

Noteworthy, some of the participants at the ECAI meeting were projects funded by the European Commission, such as Query and context based visualization of time-spatial cultural dynamics (QVIZ, <http://qviz.eu>), Great Britain Historical Geographical Information System (GBHGIS, <http://www.gbhgis.org>), Server for digital historical maps at the Institute of European History Mainz (<http://www.ieg-maps.uni-mainz.de>).

2. *CERL* – The Consortium of European Research Libraries (CERL, <http://www.cerl.org>) maintains a database of historical variants of place names, which is a useful extension to the Bernstein GIS in so far as (1) it allows users a wider range of search criteria and (2) it provides a greater flexibility of integrating the Bernstein digital workspace with other databases and Internet services. Discussions are underway on providing CERL with the place coordinates in exchange for variants of place names.

3. *BKG* – In order to solve the issue of digital boundary coordinates for administrative units it is intended to contact the German Federal Agency for Cartography and Geodesy (*BKG*, <http://www.bkg.bund.de>), which has work experience in spatial data projects of the European Commission. The contact was suggested by the Bernstein Project Officer, who we thank for providing us with the name of the persons to write to in this matter.

## 4. Future steps

---

The next steps to complete the geographical data stock for the Bernstein GIS, are:

1. Complete the matching of place names in the geo-list.
2. Incorporate the place names as given by three Bernstein databases (*NIKI*, *WILC*, *WZMA*, c. 200 place names).
3. Acquire boundary coordinates for the administrative units in the geo-list.