



Digital Historical Maps

Report from WP7

Production and Production Methods

Final version 1.0

Deliverable 22

April 06, 2001

Harry Hietanen

National Land Survey of Sweden

Table of Contents

1	Purpose of this Workpackage	2
1.1	Denmark	2
1.1.1	Method	2
1.1.2	Workflow	2
1.1.3	Comments	2
1.1.4	Expansion of scanning	3
1.2	Germany	3
1.2.1	Method	3
1.2.2	Workflow	4
1.2.3	Comments	5
1.3	Sweden	5
1.3.1	Method	5
1.3.2	Workflow	5

Enclosures

- Appendix 1 Production report
- Appendix 2 Workflow Denmark

1 Purpose of this Workpackage

The purpose of this workpackage is to plan, follow up and report the production of digital files to the prototype. In this report an outline of methods used in the production is made.

1.1 Denmark

1.1.1 Method

The maps were scanned in the Technical Department of KMS in a flat bed scanner in 508-dpi TIFF. The files were burned upon CD-ROMs, normally uncompressed, but when the file size exceed the available space, it was compressed with LZW.

The naming/indexing was done beforehand by the archive section, thus from the beginning the file name signified a certain parish making it easy to check whether a map file was missing.

1.1.2 Workflow

The CD-ROMs were taken to the archive section, which did the rest of the work. Later in the project the Technical Department put the files on a server, from which the archive section could copy them. The files stayed on the server until the archive section had checked the quality, then they were transferred to tape.

In the archive section the files were converted by Image Alchemy to .raw files and thumbnails in jpg were made. From the thumbnails were made a quality check in order to spot obvious mistakes as wrong file names or bad scanning. File names were corrected and bad files deleted. Then the .raw files were down-sampled by Image Alchemy to 254-dpi TIFF. From these sid files were made. Again a quality check was made.

The 254 dpi TIFF files and the sid files were burned (2 sets of each) on CD-ROMs. A quality check of this process was also made.

Workflow overview, see appendix 2.

1.1.3 Comments

The decision to take a copy of the original scanning in 508 dpi TIFF before it was converted to 254 dpi was done for several reasons. One of them was that the production flow from the beginning not was optimized, another that a decision was made after the project had started to make the archive copy so good that very few new scannings were needed.

Another factor influencing the work flow of the project was that only late in the production process server capacity was made available, making the process dependent upon CD-ROMs.

These circumstances influenced the workflow and gave us some experience in what to avoid.

The workflow is described in the diagram added to this part of the report. Noticeable is that the workflow is decided by two factors:

- the decision of having the archive file of 508 dpi combined with other things concentrated a lot of work in the archives for further handling of the file. Eventually we succeeded in automatization of the greater part of the process, but it took some time.

Digital Historical Maps

Report	Status:	Version:	Date
WP 7	Final	1.0	2001-04-06
Production and production methods			Page 3

- the check points for quality control (bad scanning, bad file names etc.) was built into the work flow. As in the beginning up to 10% of the files had problems this was very important. The files were rather big and sometimes took some time to open. Thus it in the long run was better not to check at once but let the computer do the work, even though we thus had to rename or delete more files. The recommendation could even be that if the workflow is fully automatized, creating all files at once and then check - so it might be better to wait with the quality control until the end. The computer works overnight, the people do not.

It can be strongly recommended to invest some money in hardware and software. This was done during the process and it saved a lot of time.

Workflow and automatization are key words. It takes a lot of time if things are not synchronized, too. We had only one production machine and found out how much to put through the system overnight so the machine was not locked producing sid-files when we came in the morning and needed it for another task.

For future use it is very important that the workflow is not dependent on CD-ROMs, but on server capacity. The need for changing a file name or for deleting a bad file is great. By keeping the files archive on CD-ROMs we shall have to make a new edition of the 508-dpi scanning in order to make a workable file archiving system. When we got server capacity we were able to do the editing along the road, thus making the archive system in order.

1.1.4 Expansion of scanning

At first, we thought that there only was time to scan one map for each parish. We would like to make the production cover as much of Denmark as possible. But as we found out that we could not get a synchronized coverage of Denmark from these, and that we could obtain the amount of scanner time needed, the decision was to take all parish maps, making the result 2964 instead of 1800. One experience similar to the ones in point 3 is that it often takes much time for the editing work. What you save in scanner time and file handling time is partially used in editing time. This should also be taken into account when deciding scanning policy. We have later done this and scanned all of a map set of app. 500 maps covering parts of the same area in several editions. It would take a long time to sort things out, thus it was easier and better services to scan the whole lot and let the customers decide which parts they want. Especially as scanner capacity was one of the bottlenecks.

1.2 Germany

1.2.1 Method

Scanning with:

- Line scanning photographic digital camera Leica S 1 pro
- Triple linear colour CCD line with 5140 pixels; 36 bit technology
- Max. image size: 76 MB
- Resolution: 5140 x 5140 pixels
- Optics: Leica, APO-Macro-Elmarit-R
- Light: HEDLER-System-Light (HMI-Licht, Daylight - Lamps: 5200 K)

Digital Historical Maps

Report	Status:	Version:	Date
WP 7	Final	1.0	2001-04-06
Production and production methods			Page 4

- Colour: Colour system: CIE Lab
- Colour depth: 36 bit RGB Colour Space
- Reference: AGFA Colourtable

Digital photo:

For 139 scanned maps: storage of TIFF-files (appr. 60 MB each) on CD-ROM and hard disk drive of NT4.0-PC

Conversion:

Compression of the TIFF-files with MrSid-Software to MrSid-Files (appr. 2.5 MB each), storage on CD-ROM and hard disk drive of NT4.0-PC

Naming/Indexing

File naming at the same time as photo scanning: storing of each file during the scanning process using the file name; Example: DM_AF_ABW_121_0.TIFF

DM = Deutschland, Matrikelkarte (Germany, Matrikel map)

AF = Amt Franzburg (Shortening of old county name, county of Franzburg)

ABW = Shortening of Map name, which is the name of the village(s) shown on the map

121 = Map number (internal number for this project – german part)

o = original data (before any subsequent data processing, which may follow – e.g. geocoding)

TIFF = File format TIFF

1.2.2 Workflow

- **Attach** the map at panel, **arrange** light and distance between camera and panel.
- **Scan** with photo scanning camera.
- **Store** original TIFF-file on PC hard disk drive during scanning process using the final file name for TIFF-files (*.tif)
- Do the **quality check** visually on screen of PC; in case of unsatisfying result: re-arrange light and other parameters and repeat scanning.
- **Import** the TIFF-file to MrSid software on PC.
- **Convert** the TIFF-file to MrSid compressed format using the final file name for MrSid-files (*.sid)
- Make a visually **check** of the result using MrSid Viewer; in case of unsatisfying result: re-arrange compression parameters and repeat compressing.
- **Copy** TIFF-files on CD-ROM; make a table of contents in Excel for all CD-ROMs with TIFF-files
- **Copy** MrSid-files on CD-ROM; make a table of contents in Excel for the CD-ROM (just one CD) with MrSid-files

Digital Historical Maps

Report	Status:	Version:	Date
WP 7	Final	1.0	2001-04-06
Production and production methods			Page 5

- **Attach Labels** to CDs, showing their contents.

1.2.3 Comments

Why the method was used

The production method was used, because with photo scanning

- it was possible to scan all maps in the office; it was not necessary to bring the maps to another place; instead the equipment for photo scanning was brought to the office in Greifswald;
- it was possible to do the scanning most carefully with respect to the handling of vulnerable and valuable originals of historical maps;
- it was possible to get high quality results, as it was shown before in another project in Mecklenburg-Vorpommern;
- the costs were lowest in comparison with other offers

Comments about text scanning

Scanning of additional texts was carried out using an ordinary A4-Scanner (colour scanning). It was decided late in the project to include (German) texts as an additional information in order to increase the information content and quality of the prototype.

1.3 Sweden

1.3.1 Method

Several methods are used depending on the material and the costs for each production line.

- Photographing to colour diapositive
- Scanning of colour diapositive
- Scanning of microfilm
- Transfer data from tape and CD to server
- Conversion of files
- Compression of files
- Geocoding

1.3.2 Workflow

Selection

A representative selection of the archives contents is made

- Geographically to get examples from all parts of Sweden
- From a historic perspective to find areas that are of special interest and where it has been possible to find maps from different times.
- From a “artistic” perspective to make it possible to present beautiful examples from the archives

Photographing to diapositive

Digital Historical Maps

Report	Status:	Version:	Date
WP 7	Final	1.0	2001-04-06
Production and production methods			Page 6

See the specification in WP2

Scanning of colour diapositive

The size of the diapositive decides the resolution to be used at the scanning (ten pixels per mm in original scale).

In the project the diapositives of the cadastral maps have been scanned at a resolution of 1270 dpi to meet the specification. The diapositives of the geometric landbooks have been scanned at 2189 dpi due to use of another film. The diapositives are scanned in 48-bit colour depth to make it possible to catch the colour shades in the map.

The TIFF image is finally converted and saved in 254 dpi, 24 bits colour depth.

Scanning of microfilm

Almost all archives are microfilmed in B/W. We have chosen to scan the microfilms based on the selection mentioned above. This method has, at a low cost, resulted in a rather large amount of image files. Only a limited part of those are available in the prototype.

The microfilms are scanned in B/W (1 bit) colour depth and saved in TIFF group 4 (254-dpi)

Transfer of data from tape and CD

Parts of the material was earlier digitised using the same specifications and stored on different types of tape and CD-ROM. This material was transferred, after indexing and quality control, to the storage system.

Conversion of files

To handle image files from different sources and of different types the file name convention is created. At transferring of data the file was given its new name.

Compression of files

At the moment the image files are stored in the final TIFF storage also automated processes convert the files to sid format and store them on a sid version on tapes and, once a week, to the web-server disks.

Geocoding

At transfer of the files to the webserver a logfile is created. This logfile is used to update the database.

A few files have been geocoded to the C-level to show the possible use of the material.

Digital Historical Maps

Report

Status:

Version:

Date

WP 7

Final

1.0

2001-04-06

Production and production methods

Page 7

1.3.2.1 Production report

Table A: Searchable files in Prototype Mars 2001.

	Acts	Maps	Map-files	Text-files	Total files
Denmark		2925	2925		2925
Germany		131	131	147	278
Swedish cadastral maps	1422	1653	2190	2564	4754
Swedish land books	1335	1279	2481		2481
Total	2757	5988	7727	2711	10438

Table B: Further image files that will be available in April 2001.

	Map-files	Text-files
Swedish cadastral maps	200-1900	43
Swedish printed maps, Häradskartan	~500	
Swedish printed maps, Generalstabskartan	~180	
Total	~2500	43

PUB 1178 – DHM/25262

Production report

Production Report Final (5)

Reporting period 2000-08-28—2001-03-31

2001-04-06

Harry Hietanen

NLS

Project duration: 1999-01-01 - 2001-03-31

WP7	Activity	Time schedule		
		Completed 2000-08-28	Completed 2001-03-31	Planned 2000-12-31
Sweden				
Map files	Transferring of data	2750	4259	~ 3000
	Scanning	2342	2342	~ 3000
	Compression	5092	6601	6000
	Name Conversion	5092	6601	6000
	Geocoding (A-level)	4576	4671	6000
	Geocoding (C-level)	0	15	< 100
	Rectifying	0	15	< 100
Text files	Scanning	4000	4115	4000
	Transferring of data	4000	4115	4000
	Compression	265	4115	4000
	Name Conversion	265	4115	4000
	Geocoding (A-level)	265	2564	4000
In the application		4841	7235	10000
Denmark				
Map files	Scanning	1950	2925	1800
	Compression	1700	2925	1800
	Geocoding (A-level)	1550	2925	1800
In the application		1454	2925	1 800
Germany				
Map files	Scanning	103	131	127
	Compressing	103	131	127
	Geocoding (A-level)	103	131	127
Text files	Scanning		333	
	Compression		333	
	Geocoding (A-level)		147	
In the application		103	278	127

Total	Completed	Planned
In the application	10438	11 927

