

Project Plan

State of Affairs

Conferences

Publications

Presentations

Credits

hsn@iisg.nl

Designed and hosted by IISH

Life courses in context

A collaboratory based on Dutch population registers and censuses (19th and 20th century)

The programme's objective is to develop a database with about 40.000 individual life courses of people born in the period of 1863-1922. This database with micro-data will be supplemented with core data on the level of the municipalities. This will be done by digitalizing the results of the ten-yearly censuses as they were taken between 1859 and 1947.

The programme builds on work already carried out by the Historical Sample of the Population of the Netherlands (HSN) and the Data Archiving and Networked Services (DANS) for the digitised census data.



Census 1947, the calculating machine Photo Ben van Meerendonk, coll. IISH

Life Courses in Context Cruquiusweg 31 1019 AT Amsterdam 020-6685866 hsn@iisg.nl



Project Plan

State of Affairs

Conferences

Publications

Presentations

Credits

hsn@iisg.nl

Designed and hosted by IISH

Project Plan

(Short version, see also the extended version)

From the programme for large investments from NWO (Dutch National Fund for Scientific Research) the Historical Sample of the Netherlands (HSN) in collaboration with the DANS did get a grant of 3,085 million EURO for the programme 'Life courses in context'. It is the first time that Humanities gets a grant from this fund for large investments. The programme's objective is to develop a database with about 40.000 individual life courses of people born in the period of 1863-1922. This database with micro-data will be supplemented with core data on the level of the municipalities. This will be done by digitalizing the results of the ten-yearly censuses as they were taken between 1859 and 1947. Above this amount the Royal Netherlands Academy of Arts and Sciences has granted 600.000 Euro.

The programme will build on work already carried out by two KNAW institutes - the International Institute of Social History (IISH) for the Historical Sample of the Population of the Netherlands (HSN) and the Data Archiving and Networked Services (DANS) for the digitized census data - in collaboration with scholars active in a broad group of research institutes.

The database of the Historical Sample of the Netherlands (HSN) will cover the entire country and contain micro-level data on the life courses of over 40,000 individuals born between 1863 and 1922. These life courses are to include data on each successive family situation in which the individuals lived, as well as data on the religion and occupational title of each subject and of every person with whom they co-resided (and, for married subjects, data on the occupational title and place of residence of family members of the subject's spouse). Because the population registers have recorded details of every change of residence for every individual since 1850, investigators can have access to research populations that are not limited to persons who stayed put in just one municipality. The HSN database will cover the entire period from 1863 to 2000 and can be regarded as a chronological expansion of the retrospective databases currently in use by the social sciences which database consists of data form persons only born in the period after 1900 and which have very scarce information about the parental situation for the earlier cohorts.

Individual life courses have to be analysed in a rapidly changing environment of a industrializing and modernizing Dutch society. The Dutch national censuses form a fundamental source of information for conditions on the level of the municipality. In addition to the population size, population censuses contain information on the structural characteristics of the population, such as age, gender, marital status, religion, household status, occupational activity, and nationality. In some years the censuses were combined with an occupational census and a housing census. The present grant application includes the digitisation of the population and occupational censuses of the Netherlands for the full period 1859 to 1947. The data will be linked and integrated with those from the Historical Database of Dutch Municipalities and will build on those censuses already compiled by the Data Archiving and Networked Services (DANS) and Statistics Netherlands (1795-1859 and 1899).

The expansion of the existing databases of HSN and DANS into complete systems will be of major significance in the long term. In the short term it will provide a strong stimulus to academic research into social, demographic and economic developments over the past two centuries. Based on these databases, a research programme has been drawn up involving the principal research groups using quantitative material on the Netherlands in the nineteenth and twentieth centuries. This programme will result in a series of studies in historical demography, social and economic history, human geography, sociology and epidemiology. The work will be performed in stages so that the data will become available for research before the project is completed.

Another aspect of methodological innovation envisaged by this programme lies in the scope for connecting data from the micro and meso/macro levels. The population censuses and other municipal data offer a context for the individual-level and family-level data. The combination of the different sources will create new opportunities for multi-level or cross-level analysis.

Project Plan

State of Affairs

Conferences

Publications

Presentations

Credits

hsn@iisg.nl

Designed and hosted by IISH

State of affairs

Here, we present a short overview of activities since the start of the project at 1 March 2003. This will be done separately for each project.

A. Life Courses B. Census data

A. Life courses (HSN)

The database of the Historical Sample of the Population of the Netherlands (HSN) covers the entire country and will contain micro-level data on the life courses of over 40,000 individuals born between 1863 and 1922. These life courses are to include data on each successive family situation in which the individuals lived, all the addresses they lived at, as well as data on the religion and occupational title of each subject and of every person with whom they co-resided (and, for married subjects, data on the occupational title and place of residence of family members of the subject's spouse). The project will be carried out during the period March 2003 - February 2008.

Organization and personnel

In the months before the 1st of March 2003 the whole organization of the HSN was discussed and it was decided to make a more sharp distinction between the different parts of the work. The scope of the new project made it possible to distinguish between A) The collecting of the data in the region of birth of the Research Person (RP), B) Evaluating and coding the incoming copies of population registers including the mailing process necessary to get information about RP's who migrated out of the region of birth, C) Entering the data into the database and D) Semi-automatic checking of the data on errors and inconsistencies and eventually correcting them and in the end creating data-output.

In the period January-February 2003 the data-collection process was started. We appointed personnel in all parts of the country to start the data gathering process. Secondly we organized the handling of the incoming copies of population registers, a.o. by developing software to manage the routing of the material and to collect management information on the progress of the data stream. This program called 'Management and mail' became operational in July. Two data typists had been contracted to start with data entry. By 1 October this number was increased to seven (all part-time). Data entry works with a new data entry program which had already been implemented in the 2002. This data entry program was revised to overcome some bugs and to further sophisticate entry of some variables (HSN 4.04). In June 2006, a renewed version was launched: HSN 4.05.

At the end of 2002 the HSN-organization numbered about 25 employees of whom about 80% was financed by way of unemployment schemes. At the moment, this number has been enlarged to almost 40 persons of whom about 50% are financed by unemployment schemes. Most of the extra 25 staff have been hired on part-time basis in order to spread data collectors as much as possible across the country and to diminish risks of RSI for data entry typists. On full time basis the HSN numbers about 25 employees.

The work is directed by a staff of five: a general manager, two managers of the archival work, a database-manager, and an office-manager.

Situation of the data-set on 1 September 2006

To get some substantial intermediate results the project has been subdivided into three parts:

- Friesland, Utrecht, Zeeland and Rotterdam. In these regions work had already started in the context of other research, esp. the Early Childhood, Social Mobility and Longevity (ESMproject) and the project Mobility of Workers in Rotterdam (RAM-project), period until March 2005 (n=10.000)
- 2. All persons born in the period 1883-1922, period until March 2006 (n=19.000)
- 3. All persons born in the period 1863-1882, period until March 2008 (n=13.000)

The situation of data collecting and entry by 1 September 2006 have been given in two tables. The first table (see table 1) shows the result for the regions mentioned under A), the second one (see table 2) for the total of the country, divided in the two relevant sub periods. The following explains in detail the different steps in the data process, which are distinguished in the tables.

The production process of collecting and entering data is divided into four parts:

- 1. Data gathering Creating the files of the Research Persons (RP's) and distributing them over all the employees in the country [RP's had already been sampled from the certificates of births]. All information will be collected as far as RP's stayed in the region of birth.
- 2. Mailing and Coding As soon as the files of an RP are completed and all copies are at the office the results are evaluated by the archival coordinator. He or she will decide whether the information about the RP is complete or whether if persons need to be followed in other parts of the country. In case they are complete, all copies get some special codes like source-numbers and are transported to the data entry environment. The mailing process is also handled by the archival manager or assistants.
- 3. Data entry As soon as a file is complete all information of an RP can be entered. The software facilitates repeated data entry of population registers with makes it easier to enter registers with more or less the same information as the foregoing.
- 4. Checking data This phase starts with an automatic check on mistakes and inconsistencies in the data. This is a repeating process until the examiner is satisfied. In the end the data will appear in a new data format.

The last two lines in the tables show how much of the production is realized:

- 5. Number of marriage certificates.
- 6. Number of completed life courses.

B. Digitization of Dutch Population Censuses

The digitisation of the censuses started in mid March 2003.

The project proposal discerns the following groups of activities:

- 1. Preparatory work
- 2. Technical research
- 3. Researching/preparing data storage
- 4. Users/market research
- 5. Media conversion (imaging)
- 6. Content conversion
- 7. Digital conversion
- 8. Post processing
- 9. Implementing retrieval/access
- 10. Analysis
- 11. Testing and installation, production
- 12. Project management and organisation
- 13. Digital archiving
- 14. Operation, maintenance and support

Only those activity groups are mentioned, on which work was done in the report period so far.

Ad 1. Preparatory work

The selection of the materials to be digitised and its pre-processing, including microfilming, have already been carried out in an earlier project and in the preparation of this proposal.

Ad 2. Technical research table recognition

This activity has been carried out by two digitization specialists of DANS, who partly based themselves on earlier work carried out in 1997-1998, when the digitization of the census of 1899 took place. A score of commercial OCR packages has been compared: Finereader (ABBYY Bit Software Inc.), Typereader (ExperVision), Omnipage Pro (Caere), Readiris Pro 8 (I.R.I.S.). The best results in recognising tables are reached using FineReader (6.0 corporate edition) and Omnipage Pro.

All tested software works better on small tables than on large ones. If tables comprise two pages, a shift in the horizontal rows can cause problems. Spots are sometimes recognised as commas. In badly printed pages, it occurs that 6's are recognised as 0's. On the basis of the research it was decided not to pursue the possibility of automatic structure analysis of the tables any further. For the textual parts of the censuses (introductions, appendixes, etc.) experiments with PDF-files containing the original images and OCR'd text in the background gave satisfactory results. The application of OCR will be restricted to pages with mainly text, because it appeared not to be cost effective to use OCR for the large tables. It is not yet certain whether all small tables in the introductions can be processed.

Ad 3. Researching/preparing data storage

The data entry is being carried on workstations with two screens. On one screen the image of a table is projected, on the other a prepared Excel worksheet. For every table an Excel worksheet is prepared, in which the structure of the printed table is copied as well as possible. Per table an instruction is made on the particularities of the data entry. By using Excel formulas, copying and macros the amount of keying is reduced where possible. Formulas of totals and percentages are used to check for data entry mistakes during data input. The data-entry instructions have been made for all tables for the years 1869, 1879, 1889 and 1919.

Ad 6. Data-entry and OCR

All years except 1889 are being typed in by seven data-entry workers at DANS. The maximum number of working hours per employee per week on this job is 24. The data-entry work for the 1889 census, which is the largest census in terms of published pages, is out-sourced to a specialised company.

The progress of the data entry is followed in a simple system and in two-weekly work meetings. The data-entry of the census of 1919 is finished, 1869 is almost finished and 1879 is well under way. The data entry speed varies considerably and is between 30 minutes and 2 hours per image (two pages of a table). The speed mainly depends on the complexity of the table structure and on the number of empty cells. Data entry statistics show that the available budget should be just enough for the work to be finished.

Ad 8. Post processing

Part of the checking and correction rakes place during data entry by comparing calculated totals in the spreadsheets with printed totals in the sources. This method is very useful to find data entry errors and mistakes in the sources. A special meeting was organised to discuss the possibilities to annotate and/or correct source errors.

Because the structures of the tables vary and because the data entry for the digitisation of the censuses has taken place since 1997 in three projects and by different institutions and companies, the possibilities for error checking are not uniform throughout. Tables can have row totals, column totals, subtotals and other calculated numbers such as percentages and ratios. Totals are not always based on all rows or columns. The row and column headings consist of hierarchically subdivided texts that can only be checked visually. The small tables in the introductions offer additional information that could be used for checking purposes. Sometimes, the information in two or more tables overlaps, so that additional checking is possible. Finally, printed errata are published for several tables and years.

The experiences, also from earlier projects, are that a thorough checking and correction is very labour intensive and difficult to plan ahead. Therefore, the principles of checking and correction have been formulated cautiously as follows:

- Data-entry mistakes are corrected during data entry on the basis of the comparison of printed with calculated totals.
- Printed errata are processed in the digital tables.
- Source errors are recorded in the total cells. Standard, the automatically calculated total is included in the table, and in a comment the printed total is mentioned. In this way, the consistency

of the data is enhanced. This work method is preferred, because it appears that many source errors are calculation mistakes.

- When time permits, deviations of printed totals from calculated totals larger than 1000 are inspected more closely.
- No checking takes place involving other tables.
- All checks and corrections are documented.

Ad 9. Implementing retrieval/access

An initial design for a Website on which the data of the censuses will be published has been made. The website makes use of an open-source content management system developed by DANS. The design of the website contains a number of interactive elements (possibility of comments, discussions, additions, suggestions, links, etc.), that will gradually give body to an evolving collaboratory. Work files of censuses for which the data entry is ready will be made accessible as soon as they become available. The website can be seen at www.volkstellingen.nl/en/.

Project Plan

State of Affairs

Conferences

Publications

Presentations

Credits

hsn@iisg.nl

Designed and hosted by IISH

Conferences and workshops

2008

Life Courses in the 19th and 20th Century
 Concluding congress of the NW0-Groot program 'Life courses in context'
 Amsterdam IISH, 11 April 2008
 Introduction and programme

Past Conferences

- Social Science History Conference, Minneapolis, 2-5 November 2006
 Papers and presentations: see http://www.ssha.org/
- Conference 'Analysis of the Dutch Population Censuses 1795-2001', The Hague 29 September 2006
- World Congress Economic History, Helsinki, August 2006
 First call for papers was 30 April 2003
 Second call for papers: 31 January 2005



Project Plan

State of Affairs

Conferences

Publications

Presentations

Credits

hsn@iisg.nl

Designed and hosted by IISH

Publications

2005

- Hans Knippenberg, 'The Netherlands: Selling churches and building mosques', in: Hans Knippenberg (ed.), The changing religious landscape of Europe (Amsterdam 2005), pp. 88-106
- Jan Kok & Marco H.D. van Leeuwen (ed.), Genegenheid en gelegenheid. Twee eeuwen partnerkeuze en huwelijk (Amsterdam 2005) Hierin de volgende bijdragen:
 - M. van Leeuwen, I. Maas en K. Mandemakers, 'Het kiezen van een huwelijkspartner in Nederland 1840-1940. De rol van de familie', 63-84.
 - M. Schrover, 'Huwelijk, gender, migratie en integratie. Partnerkeuze van Duitsers in Utrecht in de negentiende eeuw', 135-158.
 - J. Kok & K. Mandemakers, 'Vrije keuze uit een beperkt aanbod. De huwelijksmarkt in Utrecht en Zeeland 1850-1940', 213-230.
 - D. Damsma & J. Kok, 'Ingedroogde harten? Partnerkeuze en sociale reproductie van de Noord-Hollandse boerenstand in de negentiende en vroeg-twintigste eeuw', 285-308.
- J. Kok, K. Mandemakers & H. Wals, 'City nomads: changing residence as a coping, Amsterdam, 1890-1940', Social Science History 29 (2005) 1, 15-44.
- F. van Poppel, M. Jonker & K. Mandemakers, 'Differential infant and child mortality in three Dutch provinces, 1812-1909', *Economic History Review* 58 (2005), 272-309.

- H. Bras, 'Social change, the institution of service and youth: the case of service in the lives of rural-born Dutch women, 1840-1940', in: *Continuity and Change* 19 (2) 2004, 241-264.
- J. Kok, Th. Engelen, A. Knotter en R. Paping, 'Labor strategies of families: An introduction', in: *History of the Family. An International Quarterly* 9 (2) 2004, 123-135.
- J. Kok, 'Choices and Constraints in the Migration of Families. Central Netherlands 1850-1940', in: *History of the Family. An International Quarterly* 9 (2) 2004, 137-158.
- J. Kok, Th. Engelen en R. Paping, 'The family strategies concept: An evaluation of four empirical case studies', in: *History of the Family. An International Quarterly* 9 (2) 2004, 239-251
- J. Kok en J. van Bavel, 'Birth spacing in the Netherlands. The effects of family composition, occupation and religion on birth intervals, 1820-1885', in: *European Journal of Population* 20 (2) 2004, 119-140.
- M. van Leeuwen, I. Maas en A. Miles, 'Creating a Historical International Standard Classification of Occupations: an exercise in multi-national, interdisciplinary co-operation', in: *Historical Methods* 37 (2004) 186-197.
- L. Lucassen, 'De selectiviteit van blijvers: een reconstructie van de sociale positie van Duitse migranten in Rotterdam (1870-1885)', in: *Tijdschrift voor sociale en economische geschiedenis* 1 (2) 2004, 92-115.
- J. van Maarseveen, 'The Dutch Virtual Census of 2001 compared to previous censuses', in: Eric Schulte Nordholt et al., *The Dutch Virtual Census of 2001. Analysis and Methodology* (Voorburg / Heerlen 2004), 119-142.
- K. Mandemakers en L. Dillon, 'Best practices with large databases on historical populations', in: *Historical Methods* 37 (1) 2004, 34-38.
- K. Mandemakers, 'De Historische Steekproef Nederlandse bevolking (HSN) en het project Life Courses in Context', in: Bevolking en Gezin 33 (1) 2004, 91-114.
- K. Mandemakers en L. Dillon, 'Best practices with large databases on historical populations', in: *History and Computing* 13 (2) 2001, gepubliceerd 2004, 207-216.
- F. van Poppel, 'How effective was nineteenth-century knowledge in fighting infant mortality?

- Evidence on infant survival among the Dutch medical profession', in: Marco Breschi and Lucia Pozzi (eds.), *The determinants of infant and child mortality in past European populations* (Udine: Forum, 2004) 191-212.
- M. Schrover, "De Duitsche is immers tegenwoordig ook zeer galant voor dames". Huwelijksgedrag van migranten in Nederland in de negentiende eeuw', in: Bevolking en Gezin 33 (2004) 2, 103-125.
- M. Schrover, 'De grenzen van het Deutschtum. Duitse immigranten in het negentiendeeeuwse Nederland', in: Leo Lucassen (ed.), Amsterdammer worden. Migranten, hun organisaties en inburgering, 1600-2000 (Amsterdam 2004) 127-145.

2003

- Th. Engelen & J. Kok, 'Permanent celibacy and late marriage in the Netherlands, 1890-1960', *Population-E* 58 (2003), 1, 67-96.
- Th. Engelen & J. Kok, 'Célibat définitif et mariage tardif aux Pays-Bas, 1890-1960', Population-F 58 (2003), 1, 69-102.
- H. Bras, 'Maids to the city: migration patterns of female domestic servants from the province of Zeeland, The Netherlands (1850-1950)', The History of the Family 8 (2003), 217-246.
- K. Mandemakers, 'Historische Steekproef Nederlandse bevolking. Een database van negentiende- en twintigste eeuwse levensgeschiedenissen', *Archievenblad* 106 (2003), 18-21.
- M. Schrover, 'De winkel van Sinkel. Duitse migranten in de 19e eeuw', *Demos, bulletin over bevolking en samenleving*, jrg. 19, februari 2003, 17-20.
- M. Schrover, 'Wat is een Duitscher voor een vrouw? Gemengde huwelijken in de 19e eeuw',
 Demos, bulletin over bevolking en samenleving, jrg. 19, augustus 2003, 58-59.
- H. Delger, 'Chancen in Rotterdam. Eine Untersuchung der sozialen Mobilität zweier deutscher Migrantengruppen (1870-1879 und 1920-1929)', in: Angelika Eder (ed.), "Wir sind auch da!". Über das Leben von und mit Migranten in europäischen Grobstädten (Hamburg 2003), 47-72.
- J. Kok, "Liever dood dan getrouwd". Verbroken verkeringen en verboden verhoudingen in Drenthe, 1909-1940', Ex Tempore. Verleden Tijdschrift 22 (2003) 3, 173-182.
- H. Bras and J. Kok, "Naturally, Every Child Was Supposed to Work". Determinants of the Leaving Home Process in The Netherlands, 1850-1940', in: F. Van Poppel, M. Oris and J. Lee (ed.), *The Road to Independence. Leaving Home in Western and Eastern Societies,* 16th-20th Centuries. Population, family and society (Bern 2003).
- J. Kok, K. Mandemakers and H. Wals, "Toen scharrelde ze met haar hele zoodje naar een derde-achterkamer". Verhuizen als bestaansstrategie, Amsterdam 1890-1940', *Tijdschrift voor Sociale Geschiedenis* 29, 333-360.
- F. van Poppel, A.C. Liefbroer, J. Schellekens, 'Religion and Social Mobility in Nineteenth-Century The Hague', *Sociology of religion: a quarterly review*, vol. 64 (2003), 2, pp. 247-272.
- K. Mandemakers, 'Historische Steekproef Nederlandse bevolking (HSN)', NEHA-bulletin 17, nr. 2, 22-30
- M. Schrover, 'Living together, working together: concentrations amongst German immigrants in the Netherlands in the nineteenth century, *Continuity and Change* 18: 2 (2003) 263-285.

- L. J. Smits, G. A. Zielhuis, P. H. Jongbloet & F. W.A. Van Poppel, 'Mother's age and daughter's fecundity. An epidemiological analysis of late 19th to early 20th century family reconstitution', *International Journal of Epidemiology* 30 (2002), 31, 349-358.
- F. van Poppel & K. Mandemakers, 'Vaders op leeftijd', *Demos. Bulletin over bevolking en samenleving* 18 (2002), febr., 13-16.
- M. H. D. van Leeuwen, I. Maas & A. Miles, HISCO. Historical International Standard Classification of Occupations (Leuven University Press Leuven 2002), 441 p.
- H. Bras, Zeeuwse meiden. Dienen in de levensloop van vrouwen, ca. 1850-1950 (Aksant Amsterdam 2002), 262 p. Dissertatie Universiteit Utrecht.
- L. Heerma van Voss, Why is there no socialism in The Netherlands? De Nederlandse

- arbeidersklasse in de twintigste eeuw. (Aksant Amsterdam 2002), 27 p. Inaugurale rede Universiteit Utrecht.
- M. Schrover, Een kolonie van Duitsers. Groepsvorming onder Duitse immigranten in Utrecht in de negentiende eeuw (Aksant Amsterdam 2002), 416 p.
- J. Kok, 'The challenge of strategy: A comment', International Review of Social History 47 (2002), 465-485.
- J. Kok, J. W. Adams & A. Kasakoff, 'Migration over the life course in XIXth century Netherlands and the American North. A comparative analysis based on genealogies and population registers', *Annales de Démographie Historique* (2002) 2, 5-27.
- C. Lesger, L. Lucassen & M. Schrover, 'Is there life outside the migrant network? German immigrants in 19th century Netherlands and the need for a more balanced migration typology', *Annales de Démographie Historique* (2002) 2, 29-50.
- J. Kok, L. Lucassen, A. B. Kasakoff & R. M. Schwartz, 'Autour du livre de C. Pooley et J. Turnbull. Migration and mobility in Britain since the XVIIIth century', *Annales de Démographie Historique* (2002) 2, 101-123.
- J. Kok, J., J.W. Adams, T. Ericsson and L.P. Moch, 'Autour du livre de Paul-André Rosental. Les sentiers invisibles: espaces, familles et migrations dans la France du XIXe siècle', Annales de Démographie Historique (2002) 2, 129-144.
- F. van Poppel, J. Schellekens & A. C. Liefbroer, 'Religious differentials in infant and child mortality in Holland', *Population Studies* 56 (2002) 3, 277-290.
- M. Schrover, 'Registratie van vreemdelingen in het bevolkingsregister (1850-1920)', in: M. Schrover (red.), Bronnen betreffende de registratie van vreemdelingen in de negentiende en twintigste eeuw, (Den Haag 2002), 93-111.
- K. Mandemakers, 'De Historische steekproef Nederlandse bevolking (HSN) en haarprojecten te Rotterdam', in: S. Boef-van der Meulen & W.J.H.M. van de Laar e.a. (red.), Migranten en passanten. Voorbeelden uit een eeuwenlange geschiedenis, (Rotterdam 2002), 93-105.
- F. van Poppel & K. Mandemakers, 'Het geluk van borstvoeding. Regionale en sociaaleconomische verschillen in kindersterfte, 1812-1912', *Index* 9 (2002), 10, 28-31.
- F. van Poppel & K. Mandemakers, 'Social-economische verschillen in zuigelingen- en kindersterfte in Nederland, 1812-1912', *Bevolking en Gezin* 31 (2002), 2, 5-40.

- M. Schrover, 'Immigrant business and niche formation in a historical perspective. The Netherlands in the nineteenth century', *Journal of Ethnic and Migration Studies* 27 (2001), 295-311.
- L. Lucassen, 'Het einde van een migratieregime: buitenlanders in Holland gedurende de negentiende eeuw', *Holland. Regionaal Historisch Tijdschrift* nr. 2 (2001).
- M. Schrover, 'Maria en Rocco. Gemengde huwelijken en integratie in de negentiende eeuw.' Jaarboek voor Vrouwengeschiedenis 21 (2001) 81-100.
- K. Mandemakers, 'The historical sample of the Netherlands (HSN)', Historical Social Research 26 (2001), 4, 179-190.
- M. Schrover, 'Women and long distance trade migration in the nineteenth-century Netherlands' in, Pamela Sharpe (red.), Women, gender and labour migration (Londen 2001), 85-107.
- L. J. Smits, P. H. Jongbloet, G. A. Zielhuis. 'Season of birth and reproductive performance. An analysis of family reconstitutions of 800 women born in the Netherlands at the end of the 19th century', *Chronobiology International* 18 (2001), 3, 525-539.
- J. Frieswijk, L. Jansma, K. Mandemakers & B. Oosterhaven (red.), *Oer Fryske minsken / Over Friese mensen.* Themanummer *It Baeken* 63 (2001), nr. 3/4. Inleidingen HSN-congres Leeuwarden, 23 februari 2001, 96 p.'s, hierin de volgende artikelen:
 - J. Frieswijk, L. Jansma, K. Mandemakers & B. Oosterhaven, 'Over Friese mensen', 137-138.
 - K. Mandemakers, 'Introductie op de Historische Steekproef Nederland', 139-145.
 - F. van Poppel & K. Mandemakers, 'Sterven in Friesland: Levensduur en sociale ongelijkheid in Friesland in de negentiende en vroeg-twintigste eeuw', 146-163.
 - I. Maas & M. H. D. van Leeuwen, 'Huwelijksmobiliteit in Friesland tussen 1850 en 1929', 164-178.
 - D. Gerritzen, 'Voornamen in Friesland. Over Friese namen, patroniemen en

meernamigheid', 179-189.



Project Plan

State of Affairs

Conferences

Publications

Presentations

Credits

hsn@iisg.nl

Designed and hosted by IISH

Presentations

- Jan Kok, 'New insights in demography using a database with historical life courses, The Netherlands 19th and 20th centuries', Institute of Ethnology of Academia Sinica, Taipei, Taiwan, 16 January 2003.
- Jan Kok, 'Geboortebeperking in het Laat Keizerlijke China. Felle debatten en nieuwe feiten', International Institute for Social History, 25 February 2003.
- Margaret Chotkovsky, 'Colleagues, friends and partners? The composition of the personal relationships of the Italians migrants in the Netherlands, 1860-1936', Netherlands Institute for Advanced Study in the Humanities and Social Sciences, April 2003.
- Frans van Poppel, 'The Historical Sample of the Population of the Netherlands', Workshop Genetics, Genealogies and Family Databases, Umea (Sweden), 4-6 June 2003.
- Frans van Poppel and Aart Liefbroer, 'Family structure and socio-economic position: survival effects in later life. First results of the Historical Sample of the Netherlands', Geneve Early life conditions and longevity: reconstructing lives from cradle tot grave, University of Geneva, 12-14 June 2003.
- Kees Mandemakers and Lisa Dillon, 'Best practices with large databases on historical populations', XVth International Conference of the Association for History and Computing, Tromsø (Norway), 6-9 August 2003.
- Kees Mandemakers and Peter Doorn, 'Life courses in context. A collaboratory based on Dutch population registers and censuses in the 19th and 20th century', XVth International Conference of the Association for History and Computing, Tromsø (Norway), 6-9 August 2003.
- Marco van Leeuwen, 'Invited comments on the paper 'Best practices with large databases on historical populations', International Association for History and Computing Conference, Tromsø (Norway), August 2003.
- Marco van Leeuwen, 'HISCO and the History of Work Information System', International Association for History and Computing Conference, Tromsø (Norway), August 2003.
- Gerard van der Harst, 'De vestiging van Zeeuwse migranten in Rotterdam e.o. aan het einde van de negentiende eeuw', Zeeuws Archief, 6 September 2003.
- Kees Mandemakers, 'De Historische Steekproef Nederlandse bevolking', Symposium 'Mensen achter het beleid. Persoonsdossiers als bron voor het historisch-wetenschappelijk onderzoek', Den Haag, 26 September 2003.
- Hilde Bras, 'Zeeuwse meiden. Dienen in de levensloop van vrouwen, 1850-1950', Reunion (subdepartment of the Nederlandse Vereniging voor Genealogie), Utrecht, 11 Oktober 2003.
- Hilde Bras, 'Dienstbodes: werk en levensloop', N.W. Posthumus Instituut, 31 Oktober 2003.
- Hilde Bras, 'The institution of service, youth and social change', American Social Science History Conference, Baltimore (USA), 7 November 2003.
- Kees Mandemakers, 'HSN expertise-software for population registers and data handling', IMAG (International Microdata Access Workgroup)-workshop 'Longitudinal and Cross-Sectional Historical Data: Intersections and Opportunities', Montréal (USA), 10-11 November 2003.
- Kees Mandemakers, Marco van Leeuwen en Ineke Maas, 'Familial influence on partner choice: the Netherlands 1830-1940', 28th Social Science History Conference, Baltimore (USA), 13-16 November 2003.
- Jona Schellekens and Frans van Poppel, 'Religious differentials in marital fertility in Holland, 1860-1910', Annual Meeting of the Social Science History Association, Baltimore (USA), 13-16 November 2003.
- Jan van Bavel and Jan Kok, 'The Role of Religion in the Dutch Marital Fertility Transition.

- Starting, Spacing and Stopping in the heart of the Netherlands, 1845-1945', presented by Jan van Bavel at 28th Social Science History Association, Baltimore, 16 november 2003.
- Kees Mandemakers, 'The Historical Sample of the Netherlands (HSN) and other Dutch databases as new instruments for the historical study of intermediary structures', ODIS-International Symposium 'Intermediary Structures and Democracy. An historical Approach', Leuven (Belgium), 28 November 2003.
- Hilde Bras and Muriel Neven, 'Rivals, substitutes or resources? The influence of siblings on the life course of rural-born women from Zeeland, the Netherlands and the Pays de Herve, Belgium (1850-1950)', Seizièmes Entretiens du Centre Jacques Cartier, colloquium 'Frères, soeurs et jumeaux: passé et present des fratries', Lyon (France), 1 december 2003.
- Jan Kok, 'Burden or opportunity? Illegitimate births in The Netherlands and Taiwan' (with Hill Gates and Sping Wang), workshop Positive or Preventive? Reproduction in Taiwan and The Netherlands, 15-17 December 2003, Academia Sinica Taipei, Taiwan.
- Jan Kok, 'Marital fertility and birth control in rural Netherlands and Taiwan, 19th and early 20th centuries' (with Wen-shan Yang and Ying-hui Hsieh), workshop Positive or Preventive? Reproduction in Taiwan and The Netherlands, 15-17 December 2003, Academia Sinica Taipei, Taiwan.



Home Credits

Project Plan

Life Courses in Context is a project of the Royal Netherlands Academy of Arts and Sciences

(KNAW)

State of Affairs Participating institutes

Conferences International Institute of Social History (IISH)
Data Archiving and Networked Services (DANS)

Data Archiving and Networked Services (DAN

Publications In collaboration with

Foundation Historical Sample of the Population of the Netherlands (HSN)

Presentations Foundation Historical Database of Dutch Municipalities (HDNG)

Statistics Netherlands (CBS, Central Bureau of Statistics)

Credits Executive management

Dr P.K. Doorn, Head of Data Archiving and Networked Services (DANS)

hsn@iisg.nl Dr C.A. Mandemakers, Head of the Historical Sample of the Population of the Netherlands (HSN),

International Institute of Social History (IISH)

Email: kma@iisg.nl

Designed and hosted by IISH

Funding

The project is funded by grants from the Netherlands Organisation for Scientific Research, the Royal Netherlands Academy of Arts and Sciences (KNAW) and the participating institutes.

The Life Courses in Context website was established in a joint effort of DANS and IISH.

Design: Eric de Ruijter (IISH)

Hosting: International Institute of Social History Maintenance: C.A. Mandemakers (IISH)

Project Plan

State of Affairs

Conferences

Publications

Presentations

Credits

hsn@iisg.nl

Designed and hosted by IISH

Life courses in context

A collaboratory based on Dutch population registers and censuses (19th and 20th century)

The programme's objective is to develop a database with about 40.000 individual life courses of people born in the period of 1863-1922. This database with micro-data will be supplemented with core data on the level of the municipalities. This will be done by digitalizing the results of the ten-yearly censuses as they were taken between 1859 and 1947.

The programme builds on work already carried out by the Historical Sample of the Population of the Netherlands (HSN) and the Data Archiving and Networked Services (DANS) for the digitised census data.



Census 1947, the calculating machine Photo Ben van Meerendonk, coll. IISH

Life Courses in Context Cruquiusweg 31 1019 AT Amsterdam 020-6685866 hsn@iisg.nl

Project Plan

State of Affairs

Conferences

Publications

Presentations

Credits

hsn@iisg.nl

Designed and hosted by IISH

Project Plan

(Extended version, see also the short version)

Since 1991 the Historical Sample of the Population of the Netherlands (HSN) has been working at the International Institute of Social History to construct a database containing micro-level data on the Dutch population. And since 1997 the Netherlands Institute for Scientific Information Services (NIWI), in collaboration with Statistics Netherlands (and more recently with the Historical Databank of Dutch Municipalities), has been digitising the census data published in countless volumes of official statistics. The work on Life Courses in Context is thus being carried out under the auspices of two KNAW institutes. Through these projects, the institutes have built up substantial relevant experience in this field in recent years, and both institutes have been able to interest or recruit researchers from various universities to take part in these projects.

A. Life Courses B. Census data

Life courses (HSN)

The life-history database is based on the existing HSN database. The HSN aims to reconstruct life courses as completely as possible for a representative segment of the nineteenth and twentieth century population. The sample required for this purpose (N=77,000) has been derived from the birth registers for the period 1812-1922. Earlier grants from the Programme for NWO Medium-Sized Investments have already enabled this base sample to be input. A large number of death certificates, marriage certificates and all the personal record cards of the subjects in the sample have also been added to the database. Much diverse research has been and is still being carried out using the database. Research based on the HSN has so far resulted in a large number of publications in the Netherlands and abroad including two PhD theses.

The dynamic population registration system introduced in the Netherlands in 1850 has resulted in an archive that is unique in international terms. It is unique because it recorded details of where inmigrants came from and where outmigrants went to, thus enabling the inclusion of the complete life courses of migrants in the dataset. The project thus involves the systematic retrieval of the data on life courses from the population registers. Although the population registers were already in use in 1850, we will input life-course data only for subjects born in 1863 or later. There are two reasons for this: first, the registers did not function properly everywhere in the Netherlands during the first few years of their existence. Second, new regulations concerning population registration were introduced in the course of 1862. As a result, the design of the system was modified, and every household was then re-registered.

Once the population registers have been input, the HSN database will contain the following information for the 40,000 subjects born after 1862:

- composition of the family into which the subject is born, and the changes in that family before the subject ultimately leaves home
- illiteracy of the subject's father (evidenced by the absence of signature on birth certificate or on death certificate of the subject)
- migration history, including information on boarding houses, etc.
- occupational title, marital status and religion of all the relatives with whom the subject coresides
- occupational title of the parents, parents-in-law, four witnesses, subject and partner of the subject (in the event of a marriage certificate)

- literacy in the subject's social environment (signature of parents, parents-in-law, four witnesses, subject and partner of subject (in the event of a marriage certificate))
- composition of the subject's own nuclear family and the changes in that family prior to the subject's death
- relief or care arrangements for the subject in old age.

These features will make the dataset a fundamental source for scholars investigating historical issues in demography, sociology, epidemiology, social economics and human geography. The nationwide coverage of the dataset ensures that regional variations can be identified, whereas current research involving such topics is usually local in nature and necessarily excludes migrants. *The database*

For the HSN database, data for each individual are systematically collected from the records kept in the public archives. Primarily, these are birth certificates, death certificates, and personal record cards. The birth certificates include information on the person born, as well as the names, addresses, ages and occupations of the parents. The death certificates include the most recent place of residence and most recent occupation of the deceased, and information on his/her spouse(s); certificates of deceased children provide a second indication of the occupational title of the father (the person reporting the death) as well as a double-check on illiteracy. The personal record cards for all subjects who died between 1 January 1940 and 1 October 1994 have now been input. The cards include data on occupation (from 1940), cause of death (up to 1953), a full migration history (all addresses), family composition, and religion. On the basis of these data, it is now already possible to research topics such as childhood mortality and migration patterns for the whole of the Netherlands.

The above data are now to be supplemented by information from marriage certificates. These give details of the occupational titles, literacy (signature), and place of residence of the bride and bridegroom, their parents and the witnesses (usually friends or family of the couple). These certificates will enable scholars to research topics such as social and geographical mobility, marital mobility and literacy. A substantial proportion of the marriage certificates in the provinces of Utrecht, Friesland, Limburg, Gelderland, Groningen and Zeeland, have already been entered. Much useful "pilot" experience has been gained through the data input. Furthermore, various studies based on the certificates have already been published.

At a later stage, information from the population registers, land registers and tax records will be input. These sources are extremely rich, providing information on the family structure, pattern of migration, further occupational history, and the income and wealth of the subject (and sometimes of his or her relatives). From 1850 the Dutch population registers were maintained as dynamic records. By this we mean that the registers did not merely record a situation at a particular moment in time (a snapshot), but that all changes in a subject's address, family size and migration are noted, creating a longitudinal record. From 1870 onwards the records are actually fairly accurate. Many subjects born after 1870 can also be found in the personal record cards archive at the Central Bureau for Genealogy, so that their migration history can be traced in the reverse direction, thus minimising the risk of "losing" a subject.

The subjects were selected by taking a simple random sample from the birth registers for 1812-1922. The aim was to secure a sample size of 77,000. This is just over half of one per cent of the total number of births, assuming around 14.5 million people were born in the Netherlands in this period. A sample of 77,000 is sufficient for drawing statistically reliable conclusions on subpopulations of two per cent or more of the population born in the Netherlands during the period. So far (at date 1st June 2003) the following have been entered: birth certificates for the entire group, 22,000 death certificates, 16,000 personal record cards (available only for persons alive on 1 January 1940), 9000 marriage certificates, and details of 4000 initial registrations in the population register.

Through the collaborative projects with other researchers, the HSN database has been further enriched with around 20,000 birth certificates, 4000 marriage certificates, 4000 personal record cards and information from the population registers for 5000 individuals.

In addition to being an important source for research and a control database that can be used in interpreting findings on specific groups, the HSN database also acts as a foundation for the collection of new data. In practice, this is achieved by maintaining a data structure that can be used

by individual researchers, and by consistently using the database as a starting point in subsequent research, both by expanding the number of subjects included (oversampling) and enriching the database by introducing supplementary data for specific groups of subjects. For researchers, it cuts both ways. Not only can they use the material already input, they also have access to the software and expertise developed by the HSN. This expertise can be seen as an important byproduct of the data-entry work carried out over the past ten years. In return for the use of the software and the data already recorded, the HSN requires researchers to add to the dataset any new data they collect in the course of their research, thus ultimately making it available to other researchers too.

Public access to the data is subject to a dedicated set of privacy regulations (Dutch Data Protection Authority, number O-0030426; Law for the Protection of Personal Data: Wet Bescherming Persoonsgegevens, 6 juli 2000, Stb. 302). In accordance with the Personal Data Protection Act, the premise embodied in these regulations is that public access to data for research purposes is governed by the same arrangements as those of the archive from which the data were derived. This might mean that some data can only be made available in an anonymised form. The work on the database is carried out at the International Institute of Social History; ownership of the data rests with the HSN Foundation.

Literature:

- T. van den Brink, 'The Netherlands population registers', *Sociologia Neerlandica* 3, 51-63 C. Gordon, *The Bevolkingsregisters and their use in analyzing co-residential behaviour of the elderly* (NIDI report no. 9: Den Haag 1989)
- A. Knotter and A.C. Meijer (ed.), *De gemeentelijke bevolkingregisters, 1850-1920* (Den Haag 1995)
- K. Mandemakers, 'The Netherlands: Historical Sample of the Netherlands', P.K. Hall, R. McCaa and G. Thorvaldsen (ed.), *Handbook of international historical microdata for population research* (Minneapolis 2000)
- R.F. Vulsma, Burgerlijke stand en bevolkingsregister (Den Haag 2002)

Census data (NIWI)

The population census

National population censuses are one of the fundamental sources of information on conditions in a country. In addition to the population size, the population census generally contains information on the structural characteristics of a population, such as age, gender, marital status, religion, household status, occupational activity and nationality. In some years the Dutch censuses were combined with an occupational census and a housing census.

The first general Dutch population census was held in 1795 under the Batavian Republic. From 1829 onward, censuses were held every ten years. The 1940 census was postponed until 1947 because of the war. No population census has been held in the Netherlands since 1971 because of growing privacy consciousness (and refusal to take part) among the general public.

Only a limited number of original copies of the 200 or so published volumes of the 1795-1971 Dutch population censuses have survived. These censuses have always played an important part in historical and social-science research. Many of the published census volumes are now in poor condition. Digitisation can help preserve this material while also increasing access to it. The c. 42,500 pages contained in these volumes have now been scanned and are available digitally on CD-ROM. About 35,000 pages relate to the period 1859-1947. About 10,000 pages of the 1899 census have been published electronically and about 5,000 more pages have been converted to digital form but are not yet available. About 20,000 pages of tables still need to be converted in order to make all published population censuses until 1947.

The digital census databases will be important resources for historical and social-science research. National historical census projects have been, or are being, carried out in a number of countries, including the US, the UK, Ireland, France, Norway, Denmark, Germany, Russia and Austria. Some of these projects are based on the original source material, which means the databases can be constructed at the level of the individual.

Full digitisation

The purpose of this part of the programme is to construct and make accessible a comprehensive database containing the 1859-1947 national population censuses (including occupational and housing censuses). The population censuses, which form the basis for so many other statistics, will be a key research resource for social and economic historians, demographers, social scientists and epidemiologists. Secondary target groups are amateur historians, local authorities, the media and education. The project will comprise the following elements:

- Making accessible the figures from the 1859 and 1930 census publications that have already been input, and publishing them digitally on CD-ROM and the Internet and publishing them digitally on CD-ROM and the Internet; the software for retrieval and access to the data will be StatLine, a package developed by the CBS (Statistics Netherlands).
- Content conversion (by manual data entry and/or optical character recognition) of the figures from the 1869, 1879, 1889, 1909, 1920 and 1947 censuses (ca. 20,000 pages) and publishing them in digital form. The material needs to be checked and, where necessary, corrected, documented and converted to the retrieval StatLine. Wherever possible, the data will be standardised or uniformalised to enable comparisons over time.
- A historical database of co-ordinates is available, enabling data at the municipal level to be represented visually for any desired point in time during the past two centuries. The project will also provide additional information for the Historical Database of Dutch Municipalities (HDNG), which is currently being developed and in which the universities of Nijmegen and Amsterdam, the Netherlands Interdisciplinary Demographic Institute and the Historical Sample of the Population of the Netherlands collaborate. A number of existing historical-statistical and demographic databases (the historical-ecological database at the University of Amsterdam, the 'Hofstee' database at the NIDI, and databases at the University of Nijmegen) will be integrated with the HDNG.



Table 1: The total number of research-persons by phase of collecting and entry for the main regions

See also the numbers for the Netherlands, or go back to the state of affairs.

Provinces of Friesland, Utrecht and Zeeland and the city of Rotterdam*	1850-1862		1863-1922		Total birth period 1850-1922			
	N % of total	% of	N	% of	1 Sept	2006	1 Jan	1 Sept
N.		1	total	N	%	2005	2004	
Total number of research-persons	2544		10125		12669			
1. Data-collection 1: persons followed in region of birth	2544	100.0	10125	100.0	12669	100.0	100.0	100.0
2. Data-collection 2: a.o. migration trajectories	2544	100.0	10125	100.0	12669	100.0	100.0	95.8
3. Data entry	2100	82.5	9000	88.9	11100	87.6	73.2	56.8
4. Checking of data entry	2028	79.7	8216	81.1	10244	80.9	52.0	46.0
5. Finished: number of marriage-certificates	930	^a 59.9	4472	^c 56.8	5402	^d 57.3	51.9	51.9
6. Finished: Life courses (population register)	1800	70.8	6720	66.4	8520	67.3	43.9	37.1

Explanation: the numbers of phase 3, 4 and 6 are estimations and not exact counts.

^a Calculation was based on the total amount of 1552 expected marriage-certificates.

^c Calculation was based on the total amount of 5548 expected marriage-certificates.

^d Calculation was based on the total amount of 9429 expected marriage-certificates.

^{*}Rotterdam includes the surrounding villages



Table 2: The total number of research-persons by the phase of collecting and entry for the Netherlands

See also the numbers for the main regions, or go back to the state of affairs.

Netherlands	1863-1882		1883-1922		Total birth period 1863-1922				
	N % of	% of	N	% of total	1 Sept 2006		1 Jan	1 Sept	
IX.	'	total			N	%	2005	2004	
Total number of research-persons	16499		25324		41823				
Data-collection 1: persons followed in region of	40400	400	05004	400	44000	400	00.7	00.0	
birth	16499	100	25324	100	41823	100	86,7	80.0	
2. Data-collection 2: a.o. migration trajectories	15500	93.3	25125	99.2	40625	97.1	75.6	68.7	
3. Data entry	10600	64.2	19500	77	30100	72	31.2	23.1	
4. Checking of data entry	6735	40.8	15402	60.8	22137	52.9	13.7	12.5	
5. Finished: number of marriage-certificates	6040	^a 50.9	6901	b	15641	^c 47.4	38.7	33.8	
6. Finished: Life courses (population register)	2460	14.9	4800	>19	7260	17.4	11.0	9.2	

Explanation: the numbers of phase 3, 4 and 6 are estimations and not exact counts.

^a Calculation was based on the total amount of 11,999 expected marriage-certificates.

^b Calculation was based on the total amount of 20,961 expected marriage-certificates.

^c Calculation was based on the total amount of 32,960 expected marriage-certificates.