## Contents

1. The HSN in 2015 (summary) ........................................................................................................... 5  
2. The HSN ........................................................................................................................................ 7  
   2.1 Organizational Structure ............................................................................................................. 7  
   2.2 Data Collection: Starting point and sources ............................................................................. 8  
   2.3 Content of the HSN Database .................................................................................................. 9  
   2.4 Promotional activities, lectures and publications ................................................................. 13  
3. Projects .......................................................................................................................................... 15  
   3.1 Genes, Germs and Resources .................................................................................................. 15  
   3.2 LINKing System for historical family reconstruction (LINKS) .............................................. 16  
   3.3 European Historical Population Sample Network (EHPS-Net) ................................................ 18  
   3.4 CEDA_R project ....................................................................................................................... 20  
   3.5 IDS meso .................................................................................................................................. 21  
4. Staff and Boards HSN ..................................................................................................................... 22  
   4.1 Staff HSN .................................................................................................................................. 22  
   4.2 Board Foundation HSN ............................................................................................................. 23  
   4.3 Steering Committee HSN .......................................................................................................... 23  
   4.4 Scientific Council of Advice ..................................................................................................... 23  
   4.5 International Advisory Board ................................................................................................. 24  

Appendix A  Publications .................................................................................................................... 25  
Appendix B  Lectures, symposia and other promotional activities ................................................... 27  
Appendix C  Reports and Working papers ........................................................................................ 30  
Appendix D  Releases .......................................................................................................................... 31  
Appendix E  Project history ............................................................................................................... 32
The Historical Sample of the Netherlands (HSN) is an initiative of researchers from different disciplines within the social and historical sciences. The goal of the HSN is to create a representative database of nineteenth- and twentieth-century life courses. A sample of the birth certificates from the period 1812-1922 serves as the basis for the HSN database.

1 The HSN in 2015 (summary)

Highlights of the year 2015 were the first IDS-school at the IISH in Amsterdam, the successful defences of theses by four PhDs using HSN data, the publication of the conference book *Population Reconstruction*. And we were notified that €200,000 euro would become available from the CLARIAH project for the curing and completion of HSN data on life courses.

The thesis of Antonie Knigge, *Sources of sibling similarity: Status attainment in the Netherlands during modernization*, was based on a LINKS dataset with linked marriages. Małgorzata Popławska and Ruben Schalk used the HSN life course dataset for their theses respectively *Essays On Insurance And Health Economics* and *Splitting the bill: matching schooling to Dutch labour markets*. Paul Puschmann used a dataset on life courses of immigrants arriving in Rotterdam around 1870 and 1920 for his comparison of three harbor cities: *Social Inclusion and Exclusion of Urban In-Migrants in Northwestern European Port Cities. Antwerp, Rotterdam & Stockholm ca. 1850-1930*.

Besides the continuous work on the extension of the HSN and LINKS database, the HSN staff was engaged on four projects during 2015. One small project, the building of a demonstrator for an Intermediate Data Structure of organizations (IDS Meso) started in November. The work on three other projects was continued: the *Genes, Germs and Resource*, the EHPS-Net and the CEDA-R project.

The project, *Genes, Germs and Resources*, studies the phenomenon of familial influences on early death and exceptional survival in the Netherlands between 1812 and 2015. It takes into account the simultaneous effects of resources, germs and genetic influences. The HSN builds two datasets for this project. One starts with the HSN basic set from the birth period of 1860-1875, selecting all children of these research persons and collecting their personal cards. In a second step all personal cards of the third generation are collected and entered. The second database starts with the LINKS dataset and reconstructs the families for provinces with relatively rich indices of the civil registers.
For quite some time several important databases with historical life course data have been working together to develop comparable datasets and joint software. In 2011, a grant of the European Science Foundation (ESF) gave this cooperation a strong impulse when the European Historical Population Sample Network (EHPS-Net) was founded. Within this network, we concentrate not only on the creation of common data structures and software, but also on education by way of summer schools, on developing new databases and on the publication of results in an E-journal. The HSN is chair of this project in which over ten countries and twenty databases are cooperating. In 2015 several working groups were engaged on all kind of topics relevant for our kind of research and on data gathering and mining. We had summer courses in Cluj, Umeå, Nijmegen and Lund and the E-journal, Historical Life Course Studies continued and reached the goal of minimal five articles per year.

The CEDA_R project aims at the conversion of about 2,288 spreadsheets in which the Dutch censuses 1830-1947 were transcribed. For this period in the history of the census the results are available at the aggregative level of municipality or province. All spreadsheets were converted into one database in order to convert the data into the RDF structure. The team concentrated on the testing and developing the RDF system and the harmonization of those spreadsheets that contain population and dwellings for each location in the Netherlands.

In 2015 17 publications and 3 working papers in relation to the HSN database or using data from the HSN database were published. The number of lectures, presentations, interviews, and other promotional activities amounted to 29. In total 46 different researchers were involved in these activities.

Work on the HSN database itself has continued throughout the year: collecting and entering life courses and complementing the sample of birth certificates. We added 2,500 birth certificates, mounting to 85,000, 2,500 personal cards and lists and entered another 500 life courses and gathered data of another 1,500 life courses.

The number of HSN employees including volunteers lowered from 16 to 15 at the end of the year. Part of the employees work part-time and some of them work away from the institute, to collect data in various archives.

Chapter 2 of this report gives an overview of the HSN organization, of the development of the database during 2015 and of the outreaching activities. Chapter 3 contains a more detailed account of the five projects that we have worked on. Chapter 4 presents the composition of the staff and the several boards of the HSN.

An overview of the publications, presentations, working papers and data releases of 2015 is presented in respectively appendix A, B, C and D. Appendix E contains an overview of all projects undertaken by the HSN since the start in 1991.
HSN staff (21 May 2015)

From top to bottom and left to right: Ben Mowues, Walfried Commandeur, Kees Mandemakers, Marja Koster, Huub van Eijden, Behice Gül, Jan Bartman, Frits Nijstad, Rolf Wasser and Bert Schijf. Not in this picture: Jos van Hees Tatiana Moisseenko and Inez de Vries.

2 The HSN

2.1 Organizational Structure

The HSN is governed by the HSN foundation. The members of the Board work at several Dutch universities. The purpose of the foundation is the construction of the HSN database and to make the HSN data available to scientific researchers in the Netherlands and abroad. The only restrictions concern preventing overlap of the research inquiries in question and the protection of data confidentiality.

Although the database of the HSN is a historical database of which most part of the included individuals is no longer alive, some still are. This implies that the HSN is bound to the regulations of the Dutch Personal Data Protection Act (Wet Bescherming Persoonsgegevens). Secondly, although most of the data are taken from records which are open to the public, some of the data have been made available by the archives for the HSN-database only for scientific research and under the condition of anonymous use of the data. The HSN privacy regulations (see https://socialhistory.org/en/hsn/hsn-privacy-statement) determines that the HSN data are only available for researchers after they have signed a license agreement.

In order to guarantee continued existence and accessibility of the HSN database, the HSN Foundation has linked itself by contract to the International Institute of Social History (IISH) in Amsterdam, which forms part of the Royal Netherlands Academy of Sciences (KNAW). The IISH is an internationally renowned archive and research institute in the field of social history. It is devoted to the acquirement, management and accessibility of collections in that area.
The International Institute of Social History (IISH) provides housing for the HSN’s activities and assumes the burden of the resulting costs. The IISH has guaranteed a permanent position for coordination tasks. The actual data gathering is done on the basis of projects, which are externally funded. The HSN is part of the IISH research department. Decisions regarding projects are made by the Steering Group which consists of members of the Board of the HSN and members of the management team of the IISH (for the composition of these boards, see chapter 4).

2.2 Data Collection: Starting point and sources

The Historical Sample of the Netherlands (HSN) strives to construct life histories as completely as possible for a representative portion of the nineteenth and twentieth century population in The Netherlands. The sample has been drawn from all persons born in The Netherlands between 1812 and 1922. Ultimately, the HSN database will include information on an individual level from about 85,000 persons on subjects like family structure, occupation, birthplace, literacy, social network and migration history.

These characteristics make the data set a basic resource for historical research into the areas of demography, sociology, epidemiology, genetics, economy and social geography. The importance of the HSN for the researcher is fourfold:

- The HSN provides a representative dataset with which research can be done into social developments in the 19th and 20th centuries.
- The HSN provides a control group or groups for researchers to compare with their own research population.
- The HSN is developing the expertise which individual researchers usually cannot acquire in the limited time at their disposal.
- The HSN offers the possibility for researchers to use the existing HSN dataset as a base for their own research projects.

Of course, this cuts both ways. Every researcher who wants to use the infrastructure and data of the HSN must agree that in return he or she will deliver his or her data to the central database, in accordance with the formal structure of this database. In this way the HSN has developed into a data centre that functions as a centre for quantitative research on life courses.

The sample is drawn from the birth certificates and stratified in periods of ten years. To achieve rather equally sized cohorts of persons from the age of twenty years, depending on infant and child mortality on the one hand and the number of births on the other hand, it was decided to have two sample frequencies: 0.75% for the period 1812-1872 and 0.5% for 1873-1922. This results in a sample size that is large enough to make sound statistical conclusions for subpopulations of minimal two percent of the persons born in the Netherlands during the 19th and early 20th century (in total about 14,5 million persons) at the age of 20.
The basic dataset of the HSN contains the most important data from the life courses of the sampled persons. The data about birth and death originate from the certificates of birth (see picture) and death. For the period after 1939 we use the personal cards instead of the death certificates. The certificates of death and marriage also comprise data about occupational titles and places of living of the parents and other relatives. Marriage certificates contain data about place of living, occupation, age, illiteracy (whether or not being able to write a signature) of both bride and groom, their parents and four witnesses (usually relatives like brothers or close friends).

Besides the certificates the data are drawn from the population registers. These sources are quite rich and deliver data about the occupational careers, the family structure and the migration patterns of the sample person and his or her relatives.

The Netherlands is one of the few countries in the world with a continuous population register starting as early as the mid-19th century. In the early registers each household was entered on a double page, with the head of the household first; he was followed by his wife, children, other relatives, and other members of the household. Date and place of birth, relation to the head of the household, sex, marital status, occupation, and religion were recorded for each individual. All changes occurring in the household were recorded in the register. Population registers remained in use until 1910 or 1920, after which a new form of continuous registration was introduced, consisting of single sheets, so-called family cards. From then on the registration unit was no longer the household, but the family.

In the late 1930s, the population register was replaced by the personal card; from that time on the individual person became the registration unit in all municipalities. Since then the population register in each municipality has consisted of a collection of personal cards, containing nearly the same information as the population register. All persons who were alive in July 1938 or were born after that year received a personal card. At the time of death, this card is removed from the files and sent to the Central Bureau of Statistics (CBS), where the data on the card are used for statistical purposes; and then it is sent to the Central Genealogical Bureau (CBG). Copies of the cards have been used for the data set. They contain the following information: name, municipality and date of birth of the person concerned, as well as those of his or her parents, marriage partner(s) and children. The nationality is given as ‘Dutch’ or ‘Foreign’. Successive occupations, addresses and changes therein are also indicated.

### 2.3 Content of the HSN Database

Figure 1 gives an overview of the data gathered for each RP since the start of the HSN in 1991. In the first ten years the HSN concentrated on the data entry of all birth certificates and the death certificates of children who died before the age of ten. After the year 2000 more and
more marriage certificates were entered and the HSN also started entering data from the population register.

In 2012 the HSN began to complete the sample of the birth certificates of the period 1903-1922. A main part of this sample was entered on the basis of a sample frequency of 0.25% instead of the aimed 0.5%. During 2015 we added another 2,500 birth certificates. This means that we added all remaining ones except about 200 from the province of South-Holland. After completing the period 1903-1922 the whole sample will contain 85,500 births.

The maximum number of all sources to be entered for the cases is defined by the number of birth certificates. From figure 1 it is clear that for the life courses we are nearly half way and for the combination of death certificates and personal cards at about two thirds of the number of births. During the year about 500 personal cards and 1,800 personal lists were added to the database and the data of 500 life courses were entered. During 2015 we did not work on marriage certificates.

The fact that the HSN is not yet complete poses a selection problem for each researcher. If and how the data are used depends on the research question and the selection the researcher will make from the dataset, see the following tables 1 and 2 for more detailed information.
Table 1  Number of birth and death certificates and personal cards in HSN dataset by period of birth, 31\textsuperscript{st} of December 2015

<table>
<thead>
<tr>
<th>Period</th>
<th>HSN Basic Sample (Number Birth Certificates)</th>
<th>Death Certificates and Personal Cards (PK) and Personal Lists (PL)</th>
<th>Death Certificates</th>
<th>PK’s &amp; PL ‘s</th>
<th>Death Certificates and PK’s &amp; PL’s</th>
<th>% Basic Sample**</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>% &lt; 16 year</td>
<td>N</td>
<td>N</td>
<td>% Basic Sample**</td>
<td></td>
</tr>
<tr>
<td>1812-1862</td>
<td>36,280</td>
<td>57.7</td>
<td>750</td>
<td>18,292</td>
<td>50.3</td>
<td></td>
</tr>
<tr>
<td>1863-1882</td>
<td>16,502</td>
<td>61.7</td>
<td>5,351</td>
<td>14,083</td>
<td>83.3</td>
<td></td>
</tr>
<tr>
<td>1883-1922*</td>
<td>32,572</td>
<td>79.9</td>
<td>19,228</td>
<td>25,797</td>
<td>79.2</td>
<td></td>
</tr>
<tr>
<td>Total 31-12-2015</td>
<td>85,354</td>
<td>63.2</td>
<td>25,329</td>
<td>58,172</td>
<td>68.2</td>
<td></td>
</tr>
<tr>
<td>Total 31-12-2014</td>
<td>82,704</td>
<td>63.2</td>
<td>22,856</td>
<td>55,699</td>
<td>67.3</td>
<td></td>
</tr>
</tbody>
</table>

* Sampling frequency 0.5% (except small part of South-Holland with 0.25% for the period 1913-1922).
* The percentages of deaths exclude double counting (of certificates and personal cards).

Table 1 presents the databases for three periods. We see that for the period 1863-1882 the percentage of found death records is about 83%. For the period 1883-1922 the percentage is about 4% lower, due to the ongoing data entry of birth certificates for the period 1903-1922 and because some persons are still alive. In the early years of the HSN the focus was on the data entry of death certificates of infants and children. This means that these deaths are still overrepresented in the HSN database, although for the two last periods the percentages in the table are exaggerating the situation since the personal cards are not included in this percentage (having only grown up HSN research persons).

Table 2 presents the number of life courses that we have taken in production during the period 2000-2010 (mainly by way of the NWO investment program Life Courses in Context), all in all 44,252 cases. We used schemes based on a) a distinction in birth period: 1863-1882 and 1883-1922 in which we prioritized a large part of the sample and b) region: the provinces of Utrecht, Zeeland, Friesland and the city of Rotterdam which acted as spearheads. For these areas we did not prioritize the sampled persons but completed all of them, we also included the life courses for the period 1850-1862 and we put the sample size for the period 1903-1922 on the necessary 0.5%.

The actual data release comprises a number of 37,173 life courses. Table 2 presents also a bifurcation of the life courses by region and period. Almost 2/3 of the included cases have a complete life course which means that we could follow them from the cradle to the grave or till the year 1940 when the personal card became the only form of population registration. We are still working on the incomplete cases. However, due to emigration, loss of registers (damage by water or fire), loosing track of persons, quite a lot of these cases will never have a complete recording of their life course.
Table 2  Number of Life Courses by region, date of birth and priority of data entry, HSN Release 2010.01

<table>
<thead>
<tr>
<th>Region</th>
<th>Priority</th>
<th>Period of Birth</th>
<th>Total</th>
<th>In release</th>
<th>Complete Life Course</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>N</td>
<td>%</td>
<td>N</td>
</tr>
<tr>
<td>Spearhead Regions</td>
<td>X</td>
<td>1850-1882</td>
<td>6,208</td>
<td>5,827</td>
<td>4,179</td>
</tr>
<tr>
<td>Rest of the Netherlands</td>
<td>X</td>
<td>1863-1882</td>
<td>6,795</td>
<td>5,608</td>
<td>4,009</td>
</tr>
<tr>
<td>Rest of the Netherlands</td>
<td></td>
<td>1863-1882</td>
<td>5,931</td>
<td>2,159</td>
<td>1,785</td>
</tr>
<tr>
<td>Spearhead Regions</td>
<td>X</td>
<td>1883-1922</td>
<td>6,528</td>
<td>6,309</td>
<td>4,805</td>
</tr>
<tr>
<td>Rest of the Netherlands</td>
<td>X</td>
<td>1883-1922</td>
<td>14,150</td>
<td>13,185</td>
<td>10,113</td>
</tr>
<tr>
<td>Rest of the Netherlands</td>
<td></td>
<td>1883-1922</td>
<td>4,640</td>
<td>4,085</td>
<td>3,081</td>
</tr>
</tbody>
</table>

Total                      |          |                 | 44,252  | 37,173     | 27,972               | 63.2                |

* Spearhead regions are the provinces of Friesland, Utrecht and Zeeeland and the city of Rotterdam. For these regions the life courses from the period of birth 1850-1862 are also included. And for the three provinces the life courses from the oversampling 1903-1922 (from 0.25 to 0.5%) are included as well.

Figure 2  Number HSN Research Persons with no finished life course, per production stage, 1843-1922 (n=27,599)
Most of the 7,000 persons who are not included in the release, originate from the birth period 1863-1882. From this total 2,500 persons have been collected and entered into the database of which 500 in the report year. So far, a number of 1,000 persons could not be tracked in the registers (mostly because of the incompleteness of the registers). The resulting number of 3,000 is in different stages in the process of data collection and data entry. Besides we already are working on the extension of the life course dataset, especially for Amsterdam, the Hague and the provinces of South-Holland and Noord-Brabant for persons from the birth period 1850-1862 (not included in table 2). So, for the life courses we concentrated on collecting and data entry. Figure 2 shows the different stages in the production process for all cases from the birth period 1843-1922 which are not included in the release 2010.01. After the data collecting in the region of birth the files are checked, in case the persons are not finished because they move to another area, the file moves to the ‘mail-system’ in which we produce requests for copies of registers filling the gaps that we found in the life courses.

Since October 2010 the HSN has worked on the implementation of the Intermediate Data Structure (IDS). The work started within the context of the Alfalab project for just the population registers. During this project it was decided that the work would include all types of HSN data. This implied that work on the IDS as a whole was brought under the umbrella of the HSN main activities. The IDS is built in several parts: an IDS for the population registers till 1940, an IDS for the civil certificates and an IDS for the personal cards (population register after 1940). In a second stage the three IDS-systems are integrated into one system. At the end of 2015 all software has been developed but parts of it still needed to be tested thoroughly.

2.4 Promotional activities, lectures and publications

In addition to the work directly connected to the database, activities were developed to raise our profile at home and abroad. One of the tools to achieve this is the HSN website. In 2015 this website was visited 4,904 times (2014: 6,546), the number of pages visited was 10,152, and there were 2,893 unique visitors. Apart from our website and the website of the Life Courses in Context project (LCC), launched May 2004, the HSN has also been responsible since December 2005 for the renewed website of the International Commission for Historical Demography (ICHD). This website offers, besides general information about the ICHD, information on international meetings, publications and links related to historical demography. The LCC website received 12,884 visitors (4,165 unique visitors), with 15,681 page views. The ICHD website received 12,003 visitors (5,397 unique visitors), with 20,225 page views. At the end of 2012 the portal of the European Historical Population Samples Network (EHPS-Net) was launched. Part of this portal is a collaboratory in which about eighty researchers participate. In 2015 the portal generated 4,644 visitors and 49,276 page views.

In 2015 17 HSN related articles and books were published (see appendix A for an overview), including the three working papers (Appendix C), this amounts to 20 publications. This figure includes four dissertations in which HSN or LINKS data were used.

In 2015 29 presentations and papers were written, both in the Netherlands and abroad (please refer to appendix B for an overview). As usual, with four lectures, the HSN had a strong presence at the conference of the Social Science History Association in Baltimore. At the Spring Meeting of the Research Committee on Social Stratification and Mobility (RC28) of the International Sociological Association (ISA), Tilburg, 28-30 May 2015 three HSN contributions were made.
In the publications and presentations 49 different researchers were involved (in 2014 46 researchers).

Figure 3  Development of the number of publications and presentations, 1991-2015
In addition to functioning as an important source for research and as a source for control groups for interpretation of research into specific groups, the HSN database serves as the basis for collecting new data. In practice this is realized through:

a) Designing and maintaining a data structure for use by individual researchers;

b) Taking the database as a starting point for further research, both through increasing the number of individuals included (oversampling) and by adding supplementary variables for a specific group of research subjects.

Scholars thus kill two birds with one stone. They can use both the data already recorded, and the software and expertise developed by the HSN. This expertise is an important byproduct of the data entering of the past ten years. For those wanting to use its software and already recorded data, the HSN sets the precondition that new data must be added to the data set, so that these data will eventually become available to other researchers too.

3 Projects

In this chapter all current projects and the activities during 2015 are presented. For a list of all HSN-projects we refer to Appendix D and for a complete overview to the HSN website (https://socialhistory.org/en/hsn/hsn-projects).

3.1 Genes, Germs and Resources

This infrastructure proposal involves the creation of new longitudinal databases named Long Lives and Linked Families. The project itself researches the role of the family and the disease environment in mortality and longevity in the Netherlands, 1812-2015. The project is granted by the Netherlands organization for Scientific Research (Free Competition Humanities and is supervised by Angélique Janssens (Radboud University Nijmegen) and Eline Slagboom (Leiden University Medical Center).

This research project proposes to study the phenomenon of familial influences on early death and exceptional survival in the Netherlands between 1812 and 2015 taking into account the simultaneous effects of resources, germs and genetic influences. ‘Resources’ are defined in socio-economic, and cultural terms; ‘germs’ refers to the disease environment, and ‘genetic influences’ refers to an individual’s genetically determined predisposition for exceptional survival or the lack thereof. The influence of these factors will be studied through a multi-generational approach in which families are followed over a time span of 200 years. The goal is to uncover the role of familial influences on survival and the changing interactions between social-structural and biological-genetic factors in mortality and longevity within changing disease environments from the nineteenth and twentieth centuries until today.
The project has several innovative aspects, among which the introduction of genetics into the study of historical mortality as recent advances in human genetics have shown the relevance of the genetic component for longevity and mortality.

The HSN is building two datasets for this project. The database ‘Long Lives’ starts with the HSN basic set from the birth period 1860-1875. Two groups will be distinguished: a) Persons who reached the age of 80 years or older and b) a control group with persons who died before the age of 65. From these persons data about their off-spring will be gathered from the population register (second, third and fourth generation). The database ‘Linked Families’ will contain data from the LINKS project in a GGR-project related format.

The project started in July 2014 and will continue till the end of 2016. The data collecting and entry for ‘Long Lives’ is divided in three subsets. At the end of 2015 the first ‘pilot’ one was delivered to the researchers. During 2015 over10,000 Personal Cards and Personal Lists were collected and entered. For the second project ‘Linked families’ two beta-releases were delivered: one for the province of Zeeland and one for the combination of the two provinces Groningen and Drenthe.

3.2 LINKing System for historical family reconstruction (LINKS)

For more than fifteen years volunteers have been indexing civil records at the Dutch provincial archives, insofar these records are accessible to the general public. Since 2012 the Central Agency of Genealogy (CBG) has taken over the organization collecting all the indexes from the provincial archives in one big data base (previously GENLIAS) and it maintains the website that makes the data accessible to the general public (https://www.wiewaswie.nl/en/home/). Nowadays the index contains names from more than 7,000,000 birth certificates of the period 1812-1912, names from more than 3,500,000 marriage certificates of the period 1812-1937 and names from more than 10,000,000 death certificates of the period 1812-1962. These indexed names are a multiple of the number of acts, because the acts are indexed for more than just one name; for marriage acts e.g. not only the names of the bride and groom are indexed, but also the names of both parents.

The LINKS (Linking system for historical family reconstruction) project is a cooperation of LIACS, NIDI, the Meertens Institute, the CBG and the organizations behind GENLIAS/WieWasWie (mainly Dutch regional archives) granted by the CATCH-program (Continuous Access To Cultural Heritage) of the Netherlands Organization for Scientific Research. The project started in June 2009 and was finalized at the end of 2014. During 2015 we worked further on a small follow up of the project, called LINKS valorization. In this project in which we cooperate with the CBG we provide a system in which the established links can be published on WieWasWie website as a result of search actions by the user. Secondly we built a system by which we can deliver established errors and other kind of mistakes can be channeled back to the archives that are responsible for the data entry of WieWasWie.
LINKS has generated a sophisticated, fast and general family reconstitution programme on the basis of the combination of birth, death and marriage certificates. As far as possible other sources such as church registers (baptism, funeral and marriage) are included as well. The programme was delivered at the end of 2014, although during 2015 we still worked on parts of the software that have to be improved to handle the reading and matching within acceptable time limits.

Nevertheless, scientific research based on the data of GENLIAS is flourishing. This was done on the basis of an already linked dataset by trainee Maarten Oosten who built a first version of a program linking the parents of brides and grooms in marriage acts to their own marriage acts. The work was done for five provinces where occupational titles were included in the index (Groningen, Overijssel, Gelderland, Zeeland and Limburg). Other datasets for research goals were created by Kees Mandemakers (linking birth, death and marriage certificates for the provinces of Groningen and Zeeland). Among others Frans van Poppel, Hilde Bras, Jan Kok, Christiaan Monden, Peter Ekamper, Roel Jennisen and Kees Mandemakers analyzed the relation between the ages of mother and daughter at the moment of their marriage, the development in geographical distances between spouses, the occurrence of marriages between nieces and nephews, aunts/uncles with nieces and nephews and other topics. In 2015 Antonie Knigge defended his thesis, Sources of Sibling Similarity. Status Attainment in the Netherlands during Modernization, based on linked marriage certificates of five provinces.

The LINKS project was officially closed with a conference in February 2014. The book with conference papers was published by Springer in September 2015: Gerrit Bloothooft, Peter Christen, Kees Mandemakers & Marijn Schraagen (eds.), Population Reconstruction.
Another offspring of the LINKS programme is software that combines the HSN dataset with the results of the LINKS record linkage. The HSN database is largely based on municipal population registers. A weakness of this source is that it does not provide information on the wider kin network of the sampled individuals and sometimes gives conflicting information or – especially in the early registers – simple does not contain the expected information. By combining the information from the HSN with that from another database, called LINKS, we will try to find out whether and how the integration of both databases offers a way to improve the quality and completeness of the HSN database. For an introduction to this software, see the video at the CLARIAH website (please click the bottom on the right to activate English subtitles). The creation of the final database is awaiting the completion of the LINKS-IDS database.

3.3 European Historical Population Sample Network (EHPS-Net)

The European Historical Population Samples Network (EHPS-Net) brings scholars together to create a common format for databases containing non-aggregated information on persons, families and households. This format or Intermediate Data Structure (IDS) forms an integrated and joint interface between many European databases. In June 2011 the European Historical Population Sample Network was launched in Strasbourg. Fourteen countries agreed to cooperate and fund the project. Kees Mandemakers was appointed as chair and Marja Koster as programme coordinator.

In January 2012 the first conference was held in Amsterdam and in September 2012 the second one in Budapest bringing together about 40 scholars and database administrators from all over Europe and Northern America. In September 2014 the third one was held in Alghero as a preconference of the first conference of the European Society of Historical Demography during which an equal number of scholars had gathered.
During the project period, 2011-2016, the main databases convert their material to the IDS format. In the meantime, data extraction programs for different types of studies (e.g., on migration and fertility) are being prepared in close collaboration between researchers and programmers. The intended system is open, scalable and extendable. New types of analysis can be introduced by adding new extraction modules. Anyone can contribute an extraction module, which will be peer-reviewed and published.

The work of the EHPS-Net is structured in ten working groups:

1. Development Portal (chair: Kees Mandemakers)
2. E-journal Editorial Board (chair: Koen Matthijs)
3. IDS Clearing Committee (chair: Kees Mandemakers)
4. Extraction software for IDS (chair: Anders Brändström)
5. Developing proposals for historical micro data infrastructure within European and national call structures (chair: Kees Mandemakers)
6. New Database (chair: Gunnar Thorvaldson)
7. Education (chair: Ioan Bolovan)
8. Standards for documentation about databases (chair: Nanna Floor Clausen)
9. GIS (chair: Diego Ramiro Fariñas)
10. IDS Extended (chair: Tommy Bengtsson)

Group 3, 4 and 10 act as one group in practice.

In June 2013 we had a Summer Course in Cluj, in July 2014 the second one and we had a more advanced course in Lund. In 2015 we organised a Winter Course on IDS in Amsterdam, and Summer Courses in Cluj, Umeå, Nijmegen and Lund. All working groups met each other at least once. The journal continued this year with more than five articles. Extraction software will be released in the course of 2016.
The program runs for five years till mid 2016. In the final quarter of 2014 we hired a technical writer, Tatyana Moisseenko, to improve the documentation in the website about the databases and write (extraction) software documentation. She continued her work until May 2015.

3.4 CEDA_R project

The project *Census data open linked - CEDA_R - From fragment to fabric - Dutch census data in a web of global cultural and historic information*, focuses on a better dissemination structure of the census data from the last two centuries at aggregated levels like municipalities and provinces ([www.volkstelling.nl](http://www.volkstelling.nl)).

The project is part of the Computational Humanities programme of the Royal Academy of Arts and Sciences (KNAW), under supervision of Sally Wyatt. The project is cooperation between DANS (Peter Doorn, Andrea Scharnhorst), IISH (Kees Mandemakers) and VU University Amsterdam (Frank van Harmelen, Rinke Hoekstra) and is in line with earlier launched initiatives in the realm of the semantic web such as Data2Semantics. The project offers a new impulse to the dissemination of the census data after the successful cooperation of the HSN and DANS in the NWO Large Investment project *Life Course in Context*. In this project most of the census data were entered in Excel spreadsheets. The project started on 1 December 2011 and will continue until the middle of 2016.

With the project two PhD’s (Ashkan Ashkpour and Albert Meroño-Peñuela) and a postdoc (Christophe Guéret) work together to improve the availability of the existing 2,288 spreadsheets with aggregate data from the censuses from the period 1830-1948 for researchers
by way of semantic technology. Promotors of the PhD’s are Frank van Harmelen (VU University Amsterdam) and Kees Mandemakers (IISG/Erasmus University Rotterdam).

The semantic valorization will be done in two steps: a) a structural improvement of the storage of the data (from spreadsheets into one dataset) and b) harmonizing and disseminating the data by way of RDF-techniques. RDF stands for Resource Description Framework technology. All data from the censuses will be accessible at the level of the cell by way of RDF coding. This will make all kinds of pattern recognition feasible and a much better querying of the dataset (it will be possible to collect for example the number of inhabitants of the village Besoyen for all census years between 1795 and 1947).

During 2015 several papers were written and several ones were published including one for *Historical Methods*. In the meantime two mini projects were worked on. Until now, main findings are that converting all data into RDF is not as easy as one might expect. The structures of the 19th century census tables are quite inconsistent, sometimes in a very surprising way (like having the same categories in the rows and the columns).

### 3.5 IDS meso

A researcher in the field of social, economic and demographic history needs structured data. Structured data are understood as data that include not only a value but also the meaning of that value. These structured data are usually distinguished according to the data level or data unit of the dataset. Depending on the level of the unit one speaks of micro, meso or macro data. Within social sciences micro data are usual data on the level of the individual or the household. Macro data are usually data on the level of nations (or large subdivisions such as states in the USA or economic sectors). Meso data are everything in between micro and macro. This could be organizations such as trade unions, a work environment as a ship, a disturbance like a strike or data about areas below state levels such as municipalities and provinces.

For micro- and macro-data we have already internationally established standards to structure the data in a common format. For micro-data on life courses there is the so-called Intermediate Data Structure (for IDS version 4, see *Alter/Mandemakers 2014*); for cross-survey data like census-data we have the harmonization structures of IPUMS (see https://www.ipums.org/). For macro-data we have a data standard in the CLIO-INFRA system. Both IDS and CLIO_INFRA were (partly) developed at IISH.

Meso data have characteristics of both micro and macro data. Take for example a dataset of ships: A ship has a lot of individual characteristics such as a name, period of existence, name ownership etc., journeys. However, it is not considered like real micro data because it can form a context on its own (crews of a ship are sets of individuals each individual with their own attributes). Datasets with trade unions or other organizations are comparable with ships.
Important here is context. Contexts are usually layered: A ship forms the context for a journey, the combination of ship and journey forms the context of the crew, the goods etc.

The project is financed by the KNAW within the ‘Data advisor’ programme and will built a demonstrator in which a uniform data structure will be made for five very different organizations like data about strikes and ships. The project will continue to February 2016.

4 Staff and Boards HSN

4.1 Staff HSN

The HSN is headed by Kees Mandemakers. Marja Koster functions as office manager of the HSN and coordinates the EHPS-Net program. Coordination between the steering committee and the research department of the IISH is managed by Karin Hofmeester.

Four workplaces were available for the work in the archives and the data entry in the office (one WIW workplace and three SWV workplaces). One trainee worked at HSN in 2015: Elisa Rodenburg. Until May, Tatiana Moisseenko worked to add content to the database section of the EHPS-Net website. In April Cor Munnik retired, immediately continuing as a volunteer to work on the HSN software; Fons van Laan continued his work on LINKS and other HSN software. Both are stationed at the IISH Digital Infrastructure department. Kerim Meijer and Richard Zijdeman were involved in the upgrading of HSN data entry software.

At the end of 2015 the staffing of HSN, directly and in cooperation with other organizations, numbered to 15 persons (2014: 18 persons). During the year, a total of 17 people worked for HSN, among whom 7 volunteers, who were engaged in collecting material in archives and data entry.

Staff in 2015:

<table>
<thead>
<tr>
<th>Name</th>
<th>FTE</th>
<th>Period</th>
</tr>
</thead>
<tbody>
<tr>
<td>J. Bartman</td>
<td>0.5</td>
<td>Jan. - Dec.</td>
</tr>
<tr>
<td>W. Commandeur</td>
<td>0.3</td>
<td>Jan. - Dec.</td>
</tr>
<tr>
<td>B. Gül</td>
<td>0.8</td>
<td>Jan. - Dec.</td>
</tr>
<tr>
<td>drs. J. van Hees</td>
<td>0.1</td>
<td>Jan. - Dec.</td>
</tr>
<tr>
<td>prof. dr. K. M. Hofmeester</td>
<td>0.1</td>
<td>Jan. - Dec.</td>
</tr>
<tr>
<td>drs. M. Koster</td>
<td>0.6</td>
<td>Jan. - Dec.</td>
</tr>
<tr>
<td>drs. F. Laan</td>
<td>1.0</td>
<td>Jan. - Dec.</td>
</tr>
<tr>
<td>prof. dr. C.A. Mandemakers</td>
<td>1.0</td>
<td>Jan. - Dec.</td>
</tr>
<tr>
<td>drs. B. Mouwes</td>
<td>0.3</td>
<td>Jan. - Dec.</td>
</tr>
<tr>
<td>T. Moisseeenko, MA</td>
<td>0.4</td>
<td>Jan. - May</td>
</tr>
<tr>
<td>drs. C. Munnik</td>
<td>0.3</td>
<td>Jan. - Dec.</td>
</tr>
<tr>
<td>drs. F. Nijsdaj</td>
<td>0.1</td>
<td>Jan. - Dec.</td>
</tr>
<tr>
<td>E. Rodenburg, MA</td>
<td>1.0</td>
<td>Jan. - Febr.</td>
</tr>
<tr>
<td>dr. B. Schijf</td>
<td>0.2</td>
<td>Jan. - Dec.</td>
</tr>
</tbody>
</table>
4.2 Board Foundation HSN

Two members of the board, Marco van Leeuwen and Jan Kok, have accepted a new term. At the end of 2015 the board consisted of (the year of resignation is between brackets):

- Prof. dr. F.W.A. van Poppel, Netherlands Interdisciplinary Demographic Institute (NIDI), Utrecht University, chair (2017)
- Prof. dr. H.A.J. Bras, Wageningen University, member (2019)
- Prof. dr. A. F. Heerma van Voss, Utrecht University, director Huygens ING, member (2017)
- Prof. dr. J. Kok, Radboud University Nijmegen, IISG, member (2020)
- Prof. dr. M.H.D. van Leeuwen, Utrecht University, vice chair (2020)
- Prof. dr. W.A.F. Maas, Utrecht University, secretary (2018)
- Dr. F.R.M. Portrait, VU University Amsterdam, treasurer (2016)
- Dr. P.G. Tassenaar, University of Groningen, member (2016)


The board held their meetings on 19 January, 12 March, 23 April and 14 October 2015. Main item on the agenda was the progress of the projects and the preparation of an application for the program NWO Large Investment.

4.3 Steering Committee HSN

The steering committee of the HSN is the decision-making body regarding the implementation of the work of the HSN. The steering committee was established to integrate the HSN into the structure of the IISH and to carry out the work related to the NOW investments.

The steering committee consists of the members of the HSN board (see foregoing section 4.2) and, on behalf of the IISH, prof. dr. L.A.C.J. Lucassen as head of the research department of the IISH (of which the HSN is a part). The secretary of the steering committee is prof. dr. K.M. Hofmeester. Advisor to the steering committee is prof. dr. C.A. Mandemakers, head of the HSN. The steering committee held their meetings on 19 January, 12 March, 23 April and 14 October 2015

4.4 Scientific Council of Advice

Task of the Advice Council is to provide the board with solicited and unsolicited advice. Chair is prof. dr. J. Dronkers. In the course of the year there were several informal contacts.

The Scientific Advisory Board consists of:

- Dr. P.K. Doorn, head DANS
- Prof. dr. J. Dronkers, Maastricht University
The HSN is advised by the International Advisory Board convening on an annual basis. Chair of the Board is prof. dr. A. Brändström. There were two mutation in the Board during 2015: prof. Kris Inwood left the board after the end of his term to be replaced by prof. Martin Dribe. The composition of the Board was as follows:

Prof. dr. A. Brändström, University of Umeå, director Demographic Database Umeå
Prof. dr. L. Dillon, Département de Démographie, Université de Montréal
Prof. dr. M. Dribe, Centre for Economic Demography, Lund University
Dr. D. Ramiro-Fariñas, Instituto de Economía, Geografía y Demografía, Madrid
Prof. dr. K.R. Smith, University of Utah.
Prof. dr. H. Vézina, l'Université du Québec à Chicoutimi (UQAC)

The Board met in Baltimore on 12 November 2015 during the annual conference of the Social Science History Association (SSHA). The various projects of the HSN and the future of the HSN were discussed.
Appendix A  
Publications

2015


For the publications in foregoing years see the HSN website: www.iisg.nl/hsn/products/publications.
Appendix B Lectures, symposia and other promotional activities

2015


609/ 604 40th Annual Meeting of the Social Science History Association, Baltimore, MD, USA, 12-15 November 2015, with the following contributions:
- Kees Mandemakers & Gerrit Bloothooft, ‘Record linkage with the LINKS project: From Marriage Certificates to Pedigrees’, session ‘Challenges, Methods and Solutions in Record Linkage of Micro-data’.
- Ingrid van Dijk, ‘Hundred Years of Repeated Bereavement: Childhood Mortality Clustering in The Netherlands 1812 – 1912’.


602 Antonie Knigge, ‘Sources of Sibling Similarity. Status Attainment in the Netherlands during Modernization’, dissertation presented at Department of Sociology’s colloquium, University of Groningen, Groningen, the Netherlands, 15 October 2015 (invited).

601 Jan Kok, ‘Current debates in family history’, Seminar at the Research Centre for Historical Studies, Rijksuniversiteit Groningen, 12 October 2015.


demographic technique? (PART 3), International Committee of Historical Sciences (CISH) 22nd Congress, Jinan, China, 23-29 August 2015


595 Antonie Knigge, ‘Sources of Sibling Similarity. Status Attainment in the Netherlands during Modernization’, dissertation presented at Department of Sociology’s colloquium, Utrecht University, Utrecht, 8 June 2015.

594/592 The Spring Meeting of the Research Committee on Social Stratification and Mobility (RC28) of the International Sociological Association (ISA), Tilburg, 28-30 May 2015, with the following contributions:
- Hao Dong & Ineke Maas, ‘Demographic Transition and Intergenerational Mobility: the Netherlands, 1850-1957.’.
- Marco H.D. van Leeuwen & Ineke Maas, ‘Towards Open Societies? Trends, variations and driving forces of intergenerational social mobility in Europe over the past three centuries’ (keynote).
- Antonie Knigge, ‘Competition and Sharing among Siblings: Status Differences between Brothers in the Netherlands in the Nineteenth Century’ (paper).


589/588 2015 Annual Meeting of the Population Association of America, San Diego, CA, USA, 30 April – 2 May 2015, with the following contributions:
- Julia A. Jennings & Clark L. Gray, ‘Climate, Marriage, and Fertility in the Netherlands, 1865-1937’ (paper).


586/585 Expert meeting ‘Using Genealogical Data in Scientific Research’, Utrecht University, 12 February 2015, with the following contributions:
- Kees Mandemakers, ‘LINKS project and the EHPS’.
- Gerrit Bloothooft, ‘Towards a historic person identifier’.

584 Paul Puschmann, Robyn Donrovich, Per-Olof Grönberg, Graziela Dekeyser & Koen Matthijs, ‘Revisiting the urban graveyard debate. An analysis of mortality differences between migrants and natives in North-Western European port cities: Antwerp, Rotterdam & Stockholm, 1850-


For the presentations in foregoing years see the HSN website: https://socialhistory.org/en/hsn/hsn-presentations.
Appendix C  Reports and Working papers

This list includes internal (HSN published) and external HSN related papers.

2015


35  Nigel Kragten, Long-Term Trends In Intergenerational Social Mobility In The Netherlands, Birth Cohorts 1760-1980, MSc thesis Research Master’s programme ‘Sociology And Social Research’, Utrecht University, the Netherlands

Appendix D Releases

Releases of the HSN and LINKS are only available on request and after signing a license agreement. For more information, see our website: https://socialhistory.org/en/hsn/hsn-privacy-statement

HSN


LINKS

17 LINKS Zeeland Linked Dataset (Marriages, Births and Deaths), Project Genes, Germs and Resources, Province of Zeeland, Release 2015_01, beta version

Appendix E  Project history

During the foregoing twenty five years several projects were undertaken by the HSN. The following lists these projects; most of them delivered specific datasets. For more information on these projects we refer to our website.

<table>
<thead>
<tr>
<th>Acronym</th>
<th>Project title</th>
</tr>
</thead>
<tbody>
<tr>
<td>MUT/ASG</td>
<td>Migration in the province of Utrecht</td>
</tr>
<tr>
<td>OVF</td>
<td>Reduced fecundity because of maternal high-risk conceptions</td>
</tr>
<tr>
<td>RDN</td>
<td>Regional differences in demographic behaviour, the Netherlands, 1900-1960</td>
</tr>
<tr>
<td>AKON</td>
<td>General index of death certificates in the Netherlands</td>
</tr>
<tr>
<td>TTA</td>
<td>Textile industry workers in Twente</td>
</tr>
<tr>
<td>MFZ</td>
<td>Geographic and Social Mobility of Female Domestic Servants in Zeeland, 1850-1950</td>
</tr>
<tr>
<td>DUM</td>
<td>Germans in Utrecht: a temporary minority in the 19th century</td>
</tr>
<tr>
<td>RCM</td>
<td>Religious differences in infant and childhood mortality, The Hague, 1860-1920</td>
</tr>
<tr>
<td>DVI</td>
<td>Settlement determinants for immigrants and their descendants in the Netherlands, 1853-1960</td>
</tr>
<tr>
<td>GBW</td>
<td>Family formation and living strategies in the western parts of the Netherlands 1830-1940</td>
</tr>
<tr>
<td>ESM</td>
<td>Early-life conditions, social mobility and longevity</td>
</tr>
<tr>
<td>RAM</td>
<td>Living Strategies of Born Rotterdammers</td>
</tr>
<tr>
<td>VBA</td>
<td>On the move in Amsterdam. Mobility of the Amsterdam poor 1900-1940</td>
</tr>
<tr>
<td>LCC</td>
<td>Life Courses in Context (NWO Large investment)</td>
</tr>
<tr>
<td>MNI</td>
<td>European migration to the Dutch East Indies</td>
</tr>
<tr>
<td>HVL</td>
<td>Marriage certificates Pupils of Dutch Higher Secondary Education</td>
</tr>
<tr>
<td>LINKS</td>
<td>LINKing System for historical family reconstruction</td>
</tr>
<tr>
<td>LMP</td>
<td>Long Term Mortality Effects of Potato Crisis</td>
</tr>
<tr>
<td>JDJ</td>
<td>Jewish Dutch or Dutch Jews?</td>
</tr>
<tr>
<td>LHL</td>
<td>Linking Historical Lives (Linked Lives)</td>
</tr>
<tr>
<td>MOSAIC</td>
<td>MOSAIC - the Netherlands</td>
</tr>
<tr>
<td>EHPS-Net</td>
<td>European Historical Population Samples Network</td>
</tr>
<tr>
<td>CEDAR</td>
<td>Census Data Open Linked</td>
</tr>
<tr>
<td>HLZ</td>
<td>HSN LINKS Zeeland</td>
</tr>
<tr>
<td>GGR</td>
<td>Genes, Germs and Resources</td>
</tr>
<tr>
<td>LINKS val</td>
<td>LINKS valorization project</td>
</tr>
<tr>
<td>IDS meso</td>
<td>IDS meso: Intermediate Data Structure for organisations</td>
</tr>
</tbody>
</table>