Historical Sample of the Netherlands

HSN

Annual Report 2018

International Institute of Social History, Amsterdam
Historical Sample of the Netherlands

Annual Report 2018

Historical Sample of the Netherlands (HSN)
International Institute for Social History, Amsterdam
Contents

1 The HSN in 2018 (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......................................................................................................... 10
   2.4 Promotional activities, lectures and publications................................................................................. 13

3 HSN-Projects ......................................................................................................................................... 15
   3.1 LINKS and LINKS valorization........................................................................................................ 15
   3.2 European Historical Population Sample Network (EHPS-Net)........................................................ 18
   3.3 LONGPOP ....................................................................................................................................... 19
   3.4 GIANTS ............................................................................................................................................ 20
   3.5 CLARIAH Cure ............................................................................................................................. 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........................................................................................................... 24
   4.5 International Advisory Board ...................................................................................................... 24

Appendix A Publications........................................................................................................................... 25
Appendix B Lectures, presentations, symposia and other promotional activities....................................... 27
Appendix C Reports and Working papers................................................................................................. 31
Appendix D Releases................................................................................................................................ 32
Appendix E Project history........................................................................................................................ 33
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.

Within the environment of the HSN two other initiatives developed. One is the so-called LINKS database which is a software project to link the indices of all civil certificates of the Netherlands into pedigrees and families. The second one is the European Historical Population Sample Network (EHPS-Net) which is chaired by the HSN and which website and journal is published by the IISH/HSN.

1 The HSN in 2018 (summary)

Besides the continuous work on the extension of the HSN and LINKS database and the EHPS Network, the HSN staff was engaged in four projects during 2018: GIANTS, collecting heights from militia registers, LONGPOP, employing two so-called Early Stage Researchers (ESR), CLARIAH Cure improving the existing HSN database and LINKS valorization, disseminating LINKS matching results to the website of WieWasWie.

At the end of the report year it became known that a project was granted about the social mobility of Jewish Dutch before the second world war. This project will start in 2019 and the HSN part consists of the gathering and entering of 1,000 life courses of Jewish persons, sampled from the ANDB the Dutch diamond workers union.

During 2018 53 different researchers were involved in publishing and presenting studies based or partly based on the HSN and LINKS databases. In total fourteen publications in relation to the HSN or using data from the HSN and LINKS databases were published.

The number of lectures, presentations, interviews and other promotional activities amounted to 52, nearing the record of 2010. With 10 papers, the HSN had a strong presence at the European Social Science History Conference which took place in Belfast, Northern Ireland. With three or four presentations the HSN was also very visible at the 43rd Annual Meeting of the Social Science History Association, The Dutch Demographic Day and the European Population Conference (EPC).

Work on the HSN database itself has continued throughout the year. The CLARIAH grant gave new impetus on the curing and completing of the HSN database. Specific goal of CLARIAH is to bring the death certificates (including personal cards) till a level of 85% of all research persons, to add 3,000 marriage certificates to the HSN-database and to cure and collect about 5,000 life courses. This year about 2,500 death certificates/personal cards were entered into the database and extra data were gathered of about 2,000 life courses.

The number of HSN employees including volunteers decreased from 24 to 20 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.
The LINKS project received a grant from the Netherlands Organization for Scientific Research (NWO) to valorize the results of the matching of the LINKS algorithms. Work started in January and in September the first links were published by the website WieWasWie (WhoWasWho) of the Dutch Family Center. The release contained about 4.7 million links of marriage certificates linking the parents of a bride or a groom to their own certificate.

The GIANTS project aims at recovering the heights from HSN Research persons and their relatives from the national militia registers of the Netherlands. During the second half of the nineteenth and the first half of the twentieth centuries, the Netherlands experienced a remarkable growth in stature, both in absolute and relative terms, which resulted in making the Dutch the tallest people on earth. GIANTS aims to understand this development, by zooming in on processes at both micro and macro levels. The existing HSN database will be enriched with information on heights of a representative sample of about 20,000 individuals, namely of the HSN male persons themselves, as well as of their fathers, brothers and sons. Thus, the researchers will gain a unique longitudinal and intergenerational perspective on the remarkable history of heights and health in The Netherlands. At the end of the year the definitive dataset of the height data prior to 1940 was released and the work on heights from post 1945 archives of the Ministry of Defense was in full progress. Principal Investigator of this project is Jan Kok from Radboud University Nijmegen. The project is financed by the Dutch National Scientific Organization.

LONGPOPO stands for the project Methodologies and Data mining techniques for the analysis of Big Data based on LONGitudinal POPulation and Epidemiological Registers. LONGPOP is a EU-project within the framework of the Marie Skłodowska-Curie Innovative Training
LONGPOP is part of a broader movement of cooperation between databases with population data. For quite some time several important databases with historical life course data have been working together to develop comparable datasets and joint software. A grant of the European Science Foundation (ESF) gave this cooperation a strong impulse founding the European Historical Population Sample Network (EHPS-Net). The network concentrates 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 network in which over ten countries and twenty databases are cooperating. The e-journal of the network, Historical Life Course Studies, continued and overreached the goal of minimal five articles per year.

Chapter 2 of this report gives an overview of the HSN organization, of the development of the database during 2018 and of the outreaching activities. Chapter 3 contains a more detailed account of the 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 2018 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.

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 European Union as laid down in the General Data Protection Regulation and as implemented by Dutch law. 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://iisg.amsterdam/en/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 activities and incorporates 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) aims 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 85,500 persons on subjects like family structure, occupation, birth place, 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 14.5 million persons born in the Netherlands during the 19th and early 20th century.

The basic dataset of the HSN contains the most important data from the life courses of the sampled persons. Data about birth and death originate from the certificates of birth (see picture) and till 1940 the certificates of death. For the period after 1939 information about death is extracted from the personal cards. Civil certificates also comprise data about occupational titles and places of living of the parents and other relatives. Especially the marriage certificates are quite rich, containing data about location of residence, 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).

The Netherlands is one of the few countries in the world with a continuous population register starting as early as the mid-19th century. These sources deliver data about the occupational careers, the family structure and the migration patterns of the sample person and his or her relatives.

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 at 1 January 1940 or were born after that year received a personal card. At the time of death, this card was removed from the files and sent to the Central Bureau of Statistics (CBS or Dutch Statistics), where the data on the card were used for statistical purposes; and then it was sent to the Central Genealogical Bureau (CBG). Copies of the cards could be used for the HSN database. 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. From 1 October 1994 onwards this system has been replaced by a centralized electronic system (Basic Registration Persons). After a person had been recorded as deceased a list with personal information is sent to the CBG (or Dutch Family Center). This archive is used by the HSN to get data for sampled persons who died after the 1st of October 1994.

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 registers. During 2012-2015 the HSN completed the sample of the birth certificates of the period 1903-1922, all but remaining about 200 from the province of South-Holland. The whole sample contains 85,383 births at the end of 2018.
The maximum number of all sources to be entered is defined by the number of births. From figure 1 it is clear that complete life courses are nearly halfway and the combination of death certificates and personal cards stands at almost 80 percent of the number of births. During the year about 2,500 death certificates and personal cards were added to the database and data of 500 life courses were completed. During 2016 we restarted working on the marriage certificates, since then about 5,000 certificates were entered. In total, we expect to enter about 45,000 to 50,000 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 presents the databases for three periods. We see that for the period 1863-1922 the percentage of found death records is 87 to 90%, for the period 1812-1861 the percentage is about 25% less, partly because the date and place of death is not known yet (especially for the provinces of North- and South-Holland) and about 1,000 known certificates of death are still to be entered. 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 certificates of death are still overrepresented in the HSN database, although for the two last periods the percentages in the table exaggerate the situation since data entry of death certificates ends in 1940.

Table 1 | Number of birth and death certificates and personal cards in HSN dataset by period of birth, 31st of December 2018

<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>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Death Certificates</td>
</tr>
<tr>
<td></td>
<td></td>
<td>N</td>
</tr>
<tr>
<td>1812-1862</td>
<td>36,280</td>
<td>23,543</td>
</tr>
<tr>
<td>1863-1882</td>
<td>16,502</td>
<td>9,299</td>
</tr>
<tr>
<td>1883-1922*</td>
<td>32,601</td>
<td>7,683</td>
</tr>
<tr>
<td>Total 31-12-2018</td>
<td>85,383</td>
<td>40,525</td>
</tr>
<tr>
<td>Total 31-12-2017</td>
<td>85,403</td>
<td>38,302</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 2 presents the number of life courses that we have taken in production since 2002 (mainly by way of the NWO investment program Life Courses in Context), all in all 44,252 cases. We prioritized parts of the sample using schemes based on distinctions in the birth period: 1863-1882 and 1883-1922 and region: the provinces of Utrecht, Zeeland, Friesland and the city of Rotterdam acted as spearheads. For these areas sampled persons were not prioritized but all of
them were completed, the life courses for the period 1850-1862 were also included and the sample size for the period 1903-1922 was already put on the necessary 0.5%.

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></td>
<td>5,827</td>
</tr>
<tr>
<td>Rest of the Netherlands</td>
<td>X</td>
<td>1863-1882</td>
<td>6,795</td>
<td></td>
<td>5,608</td>
</tr>
<tr>
<td>Rest of the Netherlands</td>
<td>X</td>
<td>1863-1882</td>
<td>5,931</td>
<td></td>
<td>2,159</td>
</tr>
<tr>
<td>Spearhead Regions</td>
<td>X</td>
<td>1883-1922</td>
<td>6,528</td>
<td></td>
<td>6,309</td>
</tr>
<tr>
<td>Rest of the Netherlands</td>
<td>X</td>
<td>1883-1922</td>
<td>14,150</td>
<td></td>
<td>13,185</td>
</tr>
<tr>
<td>Rest of the Netherlands</td>
<td></td>
<td>1883-1922</td>
<td>4,640</td>
<td></td>
<td>4,085</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
<td>44,252</td>
<td></td>
<td>37,173</td>
</tr>
</tbody>
</table>

* Spearhead regions are the provinces of Friesland, Utrecht and Zeeland 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.

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.

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, 500 of whom 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 are in different stages in the process of data collection and data entry. Besides this we have already started 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 and 1903-1922 (not included in table 2).

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 parts are integrated into one IDS-system. At the end of 2018 all software had been developed but parts of it still needed to be tested thoroughly, especially the data from the population registers till 1940.
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 2018 this website was visited 6,751 times, the number of pages visited was 14,615 and there were 4,252 unique visitors.

At the end of 2012 the portal of the European Historical Population Samples Network (EHPS-Net) was launched. Part of this portal are, among others, a collaboratory in which about eighty researchers participate and an e-journal, Historical Life Course Studies. In 2018 the EHPS-Net portal generated 3,659 visitors and 19,935 page views.

In 2018 thirteen HSN related articles and books were published (see appendix A for an overview), including one working paper (Appendix C), this amounts to fourteen publications. During the year 52 conference contributions were counted, both presented in the Netherlands and abroad, nearing the record of 2010 (please refer to appendix B for an overview). With ten presentations/papers, the HSN had a strong presence at the European Social Science History Conference, Belfast, Northern Ireland, 4-7 April 2018. Three to four presentations were given at the 43rd Annual Meeting of the Social Science History Association, Phoenix, USA, 8-11 November 2018, The Dutch Demographic Day, Utrecht, 28 November 2018, the European Population Conference (EPC), Brussels, Belgium, 6-9 June 2018, the CLARIAH Person Observation Workshop, IISG, Amsterdam, 31 May-1 June 2018 as well as the Asschauer Soiree, Altenhof, Germany, 26 May 2018.
In the publications and presentations 53 different researchers were involved (in 2017 55 researchers).

Figure 2  Development of the number of publications and presentations, 1991-2018

Figure 2  Development of the number of publications and presentations, 1991-2018
The HSN-database is not only an important source for research and a source for control groups, the HSN also serves as the basis for collecting new data. In practice this is realized by:

a) designing and maintaining a data structure for use by individual researchers;
b) taking the database as a starting point for further research, but by 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 researchers 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 HSN-Projects

In this chapter all current projects and the activities during 2018 are presented. For a list of all HSN-projects we refer to Appendix E and for a complete overview of all activities to the HSN website (https://iisg.amsterdam/en/hsn).

3.1 LINKS and LINKS valorization

LINKS entails the development of software for a LINKing System for historical family reconstruction and was financed by the CATCH program of the Netherlands Organization for Scientific Research (NWO). It aims to link all Dutch civil certificates into families and pedigrees from the 19th and early 20th century. Through the year we worked a lot on debugging the system, several releases were made. During the year we worked on the so-called valorization project publishing the results of the matching at the website of WieWasWie.

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 Centre for Family History (CBG) has taken over the organization collecting all the indexes from the provincial archives in one big database (called WieWasWie previously GENLIAS) and it maintains the website that makes the data accessible to the general public (https://www.wiewaswie.nl/en/). Nowadays the index contains names from about 10 million birth certificates of the period 1812-1918, names from about 4.5 million marriage certificates of the period 1812-1943 and names from about 12 million death certificates of the period 1812-1968. These indexed names are a multiple of the number of certificates, because the certificates are indexed for more than just one name; for marriage certificates e.g. not only the names of the bride and groom are indexed, but also the names of both parents.

Initially, LINKS was 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. At the end of 2017 by way of the CATCH project a grant was received to valorize the results from the matching system (LINKS Valorisation) by presenting the links on the website 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 first version of the programme was delivered at the end of 2014. During 2018 we worked on improving the system especially software to fasten the handling, reading and matching of the data.

Scientific research based on LINKS datasets is flourishing. Research already started with a dataset linked 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.
Another off spring 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 LINKS, we will offer 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 (the bottom on the right activates English subtitles). At the European Population Conference (EPC), Brussels, Belgium, 6-9 June 2018, Niels van den Berg, Ingrid van Dijk, Rick Mourits and others presented a comparison between the content of the HSN- and LINKS database for a sample form HSN Zeeland 1863-1872.

Kees Mandemakers gave a key lecture about the way the LINKS program is organized at the Centre d'Estudis Demogràfic, Universitat Autònoma de Barcelona. Several researchers presented outcomes of their research on the basis of LINKS data releases at the Social Science History Conference in Phoenix (Ingrid van Dijk), The Dutch Society for Research and Ageing (DUSRA (Rick Mourits, Niels van den Berg et al), Dutch Demography Day 2018 (Rick Mourits, Niels van den Berg and Angélique Janssens), Clariah toogdag (Gerrit Bloothooft et al), CLARIAH Person Observation Workshop (Gerrit Bloothooft, Kees Mandemakers).

In September 2018 the first release of the LINKS valorization project was forwarded to the Dutch Family Center who is the organization behind the WieWasWie data collection. It concerned about 4,7 million links between the parents of a bride or a groom on a marriage certificates and their own marriage certificate. See figure 4 for the way these links are presented at the website of WieWasWie. The orange blocks presents the links with other marriage certificates; not only the pedigrees but also the reverse, linking parents with their children. The first block refers to documentation by which the linkage process is explained.
3.2 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.

On 20 September 2016, the fourth general meeting of the EHPS-Net and the closing meeting of the project, took place at KU Leuven, Belgium. Most important decision in Leuven was how to continue the network in the future. This will be done on a budget that will guarantee the maintenance of the key activities of the network. These are the website including the collaboratory and the IDS repository, the e-journal, the IDS framework and the summer school system. The budget will be covered by contributions of partners and by small applications. Another change will be the structural involvement of partners that did not fully participate in the network. These are not only the databases in the UK, Spain, Italy, etc. of which the national scientific organization didn’t commit themselves to EHPS-Net, but also partners from outside of Europe, such as databases from China, USA, Australia, Canada, etc.

During 2018 several activities as the development of extraction software were continued under the flag of the LONGPOP project (see section 3.3); the e-journal Historical Life Course Studies published 13 articles among which several around the theme Infant Mortality Family Clustering. Special about this theme was that all studies were analyzed with the same software scripts on the basis of a uniform data structure (IDS).
3.3 LONGPOP

On 4 February 2016 the kick-off meeting of LONGPOP took place at CCHS-CSIC, in Madrid, Spain. LONGPOP stands for the project Methodologies and Data mining techniques for the analysis of Big Data based on Longitudinal Population and Epidemiological Registers.

LONGPOP is a project within the framework of the Marie Skłodowska-Curie Innovative Training Network within the Horizon 2020 Programme of the European Commission. LONGPOP is a consortium of high profile universities, research institutions and companies located in Spain, Netherlands, Sweden, Italy, United Kingdom, Belgium and Switzerland. Principal investigator is Diego Ramiro Fariñas based at the Agencia Estatal Consejo Superior de Investigaciones Científicas (CSIC, Madrid).

LONGPOP focuses on the rapidly changing European societies. These transformations cover changes in family structures, fertility, the decline of mortality and increase of longevity, and periods of economic and social instability. Owing to population ageing across Europe, countries are now experiencing the impact of these rapid changes on the sustainability of their welfare systems. At the same time, the use of the space and residential mobility has become a key topic, with migration within the EU countries and from outside Europe being at the center of the political agenda. Over the past decade research teams across Europe have been involved in the development and construction of longitudinal population registers and large research databases, while opening up avenues for new linkages between different data sources (i.e. administrative and health data) making possible to gain an understanding of these fast societal transformations.
However, in order to work with these types of datasets one requires advanced skills in both data management and statistical techniques. LONGPOP aims at creating a network in which the different research teams share experiences, construct joint research, create a training track for specialists in the field and increase the number of users of these large - possibly underused - databases, making more scientists and stakeholders aware of the richness in the databases. In total 15 so-called ‘Early Stage Researchers’ are positioned at the mentioned institutions, of whom two at the IISH. Both started at the first of September 2016: Francisco Anguita works on the documentation and extension of IDS related extraction software and Diogo Paiva works on methods of standardizing addresses. And both worked on the linking of HSN Research Persons with the American censuses. During 2018 they presented their results of the linkage at the European Social Science Conference in Belfast, a workshop on record linkage in Guelph. Diogo Paiva organised a workshop on The use of GIS on Historical Databases in Lisbon at the Centre for the Humanities, NOVA-FCSH, 25 June 2018. Francisco Anguita presented his work on the IDS at the Social Science Conference in Phoenix, November 2018.

LONGPOP is oriented on the rapidly changing European societies. These transformations covers changes in family forms, fertility, the decline of mortality and increase of longevity, and periods of economic and social instability.

3.4 GIANTS

The GIANTS project aims at recovering the heights of HSN Research persons and their relatives from the national militia registers of the Netherlands. During the second half of the nineteenth and the first half of the twentieth centuries, the Netherlands experienced a remarkable growth in stature, both in absolute and relative terms, which resulted in making the Dutch the tallest people on earth. Given the known impact of early life diseases and nutrition on stature, this trend indicates a remarkable improvement in health. In the proposed project we aim to understand this development, by zooming in on processes at both micro and macro levels. We study the impact on young adult stature of heritability and early life conditions such as family
size, parental socioeconomic status, the availability of nutrition and the local disease environment. Moreover, we look at the consequences of adult height and health on people’s later lives. Were taller people more successful on the marriage market, in their careers, and in reproduction? Can we discern ‘virtuous cycles’ or selection processes which allowed each successive generation to be taller?

At the macro level the role of (changing national and regional socioeconomic) inequality in explaining the Dutch gains in heights and health will be studied. For this study the HSN database will be enriched with information on heights of a representative sample of about 12,000 individuals, namely of the HSN male persons themselves, as well as of their fathers, brothers and sons (total n = about 8,000 relatives). During the year it was decided to include also the archives of the Ministry of Defense. This archive was relevant for the data of the sons of HSN Research Persons of which many of them were examined after 1945. In September permission was granted from the Ministry to collect these data for the birth years 1924 till 1965.

Principal Investigator of this project is Jan Kok from Radboud University Nijmegen. The project is financed by the Dutch National Scientific Organisation (the free competition program of the Humanities), other involved researchers are France Portrait (VU University Amsterdam), Vincent Tassenaar (University Groningen), Kristina Thompson and Björn Quanjer. The researchers will gain a unique longitudinal and intergenerational perspective on the remarkable history of heights and health in The Netherlands.

The database is aimed to be finished in the spring of 2019. The end of 2018 saw already the release of the data of the HSN Research Persons (n=11,384), fathers (n=2,154), brothers (n=5,800) and part of the sons (n=1,136).

3.5 CLARIAH Cure

CLARIAH Cure is part of the Common Lab Research Infrastructure for the Arts and Humanities (CLARIAH) which is a distributed research infrastructure for the humanities and social sciences. The CLARIAH infrastructure aims to provide researchers access to large collections of digital data and to innovative and user-friendly applications for the processing of these data. The HSN was funded with 200,000 euro to improve and cure the existing database.

Specific goals of the HSN part of CLARIAH are a) to complete the death certificates and personal cards till a level of 80% of all research persons (additional n=13,000), b) to add 3,000 marriage certificates and c) to publish a new improved release of the life course database including 4,400 not earlier published life courses, mainly from the period 1863-1902. Important goal of the project is the updating of the province of Groningen which is seriously lagging behind in the availability of life courses and finishing life courses from part of the focus area: the cities of The Hague and Rotterdam. ‘Focus area’ means that the sample period 1850 -1862 and 1903-1922 will be fully implemented as well.
During 2018 the focus was on collecting population registers and data entry. All files of the provinces of Drenthe, Noord-Brabant, Limburg and Overijssel and the city of Amsterdam were checked on completeness of data collection. After the middle of the year the project was brought back to smaller proportions and some of the staff left.

During 2018 no marriage certificates were entered since the entered number already overreached with a total of 5,378 the projected amount of 3,000. The number of death certificates and personal cards was upgraded with about 2,500 entries till 11,000.

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 (under which two SWV workplaces). Both the GIANTS project as the funding from the CLARIAH Cure project resulted in the appointment of in total six persons for 0,4 to 0,6 fte. Francisco Anguita and Diogo Paiva worked as ‘Early Stage Researchers’ within the framework of the LONGPOP project. Cor Munnik continued as a volunteer to work on the HSN software; Fons van Laan continued his work on LINKS by way of the LINKS valorization project. Both are stationed at the IISH Digital Infrastructure department.

At the end of 2018 the total number of HSN-staff, directly and in cooperation with other organizations, was 20 persons (2017: 24 persons). During the year, a total of 25 persons worked for the HSN, among whom seven volunteers, who were engaged in collecting material in archives and in data entry.

Staff in 2018:

- F. Anguita, MSc 1,0 fte Jan. - Dec.
- J. Bartman 0,4 fte Jan. - Dec.
- D. Basse 0,8 fte May - June
- R. Boemen 0,4 fte Jan. - Aug.
- T. Breugelmans 0,4 fte Jan. - Dec.
- W. Commandeur 0,3 fte Jan. - Dec.
- Th. Dibbets 0,4 fte Jan. - Dec.
- B. Gül 0,8 fte Jan. - Dec.
- drs. J. van Hees 0,1 fte Jan. - Dec.
- prof. dr. K. M. Hofmeester 0,1 fte Jan. - Dec.
- drs. M. Koster 0,8 fte Jan. - Dec.
- drs. F. Laan 0,3 fte Jan. - Dec.
- prof. dr. C.A. Mandemakers 0,8 fte Jan. - Dec.
4.2 Board Foundation HSN

In September Ineke Maas resigned from the Board after almost nineteen years of service. At the end of 2018 the board consisted of the following persons (the year of resignation is between brackets):

Prof. dr. F.W.A. (Frans) van Poppel, Netherlands Interdisciplinary Demographic Institute (NIDI), Utrecht University, chair (2022)
Prof. C.J.H. (Catrien) Bijleveld, VU University Amsterdam, director Netherlands Institute for the Study of Crime and Law Enforcement (NSCR), member (2022)
Prof. R.I.A. (Ruben) van Gaalen, University of Amsterdam, Statistics Netherlands, member (2022).
Prof. dr. A. F. (Lex) Heerma van Voss, Utrecht University, director Huygens ING, member (2022)
Prof. dr. J. (Jan) Kok, Radboud University Nijmegen, IISG, member (2020)
Prof. dr. M.H.D. (Marco) van Leeuwen, Utrecht University, vice chair (2020)
Dr. F.R.M. (France) Portrait, VU University Amsterdam, treasurer (2021)
Dr. P.G. (Vincent) Tassenaar, University of Groningen, member (2021)


The board held their meetings on 12 February and 19 September 2018. Main item on the agenda was the progress of the projects and ongoing applications.

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. (Leo) 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. (Karin) Hofmeester. Advisor to the steering committee is prof. dr. C.A. (Kees) Mandemakers, head of the HSN. The steering committee held their meetings on 12 February and 19 September 2018.

### 4.4 Scientific Council of Advice

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

The Scientific Advisory Board consists of:

Dr. P.K. Doorn, director DANS  
Prof. dr. M.G.J. Duijvendak, University of Groningen  
Prof. dr. H. van Dijk, em. university lecturer Erasmus University Rotterdam  
Prof. dr. W.Th.M. Frijhoff, em. university lecturer VU University Amsterdam  
Prof. dr. H. Knippenberg, em. university lecturer University of Amsterdam  
Prof. dr. P.Th. van de Laar, Erasmus University Rotterdam  
Prof. dr. C.H. Mulder, University of Groningen  
Prof. dr. J. Plantenga, Utrecht University  
Prof. dr. F.N. Stokman, em. university lecturer, University of Groningen  
Prof. dr. W.C. Ultee, em. university lecturer Radboud University Nijmegen  
Prof. dr. J.L. van Zanden, Utrecht University

### 4.5 International Advisory Board

The HSN is advised by the International Advisory Board convening on an annual basis. Chair of the Board is prof. Hélène Vézina. There were no mutations in the Board during 2018. All terms have been extended from 5 to 7 years. The composition of the Board is as follows:

Prof. dr. C (Cameron) Campbell, University of Science and Technology, Hong Kong  
Prof. dr. L. (Lisa) Dillon, Département de Démographie, Université de Montréal  
Prof. dr. M. (Martin) Dribe, Centre for Economic Demography, Lund University  
Dr. D. (Diego) Ramiro-Fariñas, Instituto de Economía, Geografía y Demografía, Madrid  
Prof. dr. H. (Hélène) Vézina, l'Université du Québec à Chicoutimi (UQAC)

The Board met in Phoenix on 8 November 2018 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

2018


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

2018

739 Vincent Tassenaar, ‘De bewoners van de Maatschappij van Weldadigheid. Onderzoek naar de biologische levensstandaard van de deelnemers aan een sociaal experiment (1820-1860)’, Frederiksoord, 10 December 2018.

738 Vincent Tassenaar, ‘Anthropometric history in and of the Netherlands’, Radboud University Nijmegen, 6 December 2018.


736/733 Dutch Demography Day 2018, of the Netherlands Demographic Society (NVD), Utrecht, 28 November 2018, with the following contributions:
- Jan Kok, ‘Something Old, Something New: Life Courses in Historical Demography using the Historical Sample of the Netherlands (HSN)’ (keynote lecture).
- Rick Mourits, Niels van den Berg & Angélique Janssens, Intergenerational transmission of longevity is not affected by other familial factors: Evidence from 16,905 Dutch families from Zeeland, 1812-1962.
- Matthias Rosenbaum-Feldbrügge, ‘Evolutionary Explanations do not Qualify to Predict the Relationship between Parental Death and Children’s Transition to First Marriage in Historical Populations: Findings from the Netherlands, 1850-1940’.


731/729 43rd Annual Meeting of the Social Science History Association, Phoenix, USA, 8-11 November 2018, with the following contributions:
- Kees Mandemakers & Ruben van Gaalen, ‘Linking the Historical Sample of The Netherlands (HSN) with the System of Social statistical Datasets (SSD)’.
- Ingrid van Dijk, ‘Exposure to Early-Life Mortality and Offspring Mortality’.


World Economic History Conference (WEHC), Boston, USA, 29 July-3 August 2018, with the following contributions:
- Matthias Rosenbaum-Feldbrügge, ‘The impact of parental and sibling death during childhood on women’s ages at marriage in the Netherlands, 1850-1940’.
- Kristina Thompson, ‘Adult body height as a mediator between early-life conditions and socio-economic status: The case of the Dutch Potato Famine, 1846-1847’.

Kees Mandemakers, ‘LINKS, the HSN Linking System for Historical Family Reconstruction’, invited lecture Centre d’Estudis Demogràfic, Universitat Autònoma de Barcelona, 3 July 2018.


European Population Conference (EPC), Brussels, Belgium, 6-9 June 2018, with the following contributions:
- Niels van den Berg, Ingrid van Dijk, Rick Mourits, Eline Slagboom, Angélique Janssens & Kees Mandemakers, ‘Families in comparison: An exploration of results of family reconstructions using population and vital event registrations’.
- Ingrid van Dijk, ‘Exposure to Early-Life Mortality and Offspring Mortality’.

CLARIAH Person Observation Workshop, IISG, Amsterdam, 31 May-1 June 2018, with the following contributions:
- Kees Mandemakers & Gerrit Bloothooft, ‘LINKS, the IISH Linking System for Historical Family Reconstruction’.
- Kees Mandemakers, ‘The Historical Person Identifier (HPI)’.
- Kees Mandemakers, ‘The Intermediate Data Structure (IDS) for databases with longitudinal historical data’.
- Gerrit Bloothooft, David Onland, Martin Reynaert, Katrien Depuydt & Tanneke Schoonheim, ‘Name standardization and nominal record linkage’.

Asschauer soiree, Altenhof, Germany, 26 May 2018, with the following contributions:
- Vincent Tassenaar, ‘Development of regional variety in the Netherlands’.
- Kristina Thompson, ‘Adult body height as a mediator between early-life conditions and socio-economic status: The case of the Dutch Potato Famine, 1846-1847’.


702/ 693 European Social Science History Conference, Belfast, Northern Ireland, 4-7 April 2018, with the following contributions:
- Dolores Sesma Carlos, Paul Puschmann & Jan Kok, ‘Internal migrations and later-life mortality. The Netherlands, birth cohorts 1850-1922’.
- Kees Mandemakers & Gerrit Bloothooft, ‘Consequences of Linking Decisions Based on Geographical Distance for Research on Social Mobility’.
- Thijs Hermansen, Dolores Sesma Carlos, Jan Kok & Paul Puschmann, ‘Visualizing migration trajectories using the Intermediate Data Structure’.
- Tim Riswick, ‘Between Rivalry and Support: Differences in the Mortality Chances of Brothers and Sisters in Taiwan (1906-1946) and the Netherlands (1863-1910)’.
- Vincent Tassenaar, ‘Development of regional variety in the Netherlands’.
- Kristina Thompson, ‘Adult body height as a mediator between early-life conditions and socio-economic status: The case of the Dutch Potato Famine, 1846-1847’.


For the presentations in foregoing years see the HSN website: https://iisg.amsterdam/en/hsn/products/presentations
Appendix C  Reports and Working papers

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

2018

45. Stefan Haarman, *De ontbrekende schakels. Onderzoek naar zoekstrategieën voor cold cases van de Historische Steekproef Nederland (HSN)*, Radboud Universiteit Nijmegen.

For the papers in foregoing years see the HSN website: [https://iisg.amsterdam/en/hsn/products/papers](https://iisg.amsterdam/en/hsn/products/papers)
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://iisg.amsterdam/en/hsn/privacy-statement

2018 - HSN

67  Historical Sample of the Netherlands (HSN). Data Set Civil Certificates Release 2018.01_Beta

66  Kees Mandemakers, HSN dataset Heights and Life Courses, Release 2018_02

65  Kees Mandemakers, HSN dataset Heights and Life Courses, Release 2018_01


For a complete overview of releases see the HSN website: https://iisg.amsterdam/en/hsn/products/releases

2018 - LINKS

33  Kees Mandemakers and Fons Laan, LINKS dataset Linked Marriage Certificates Netherlands, WieWasWie website, Release_2018_01 (n=4,7 million).

For the LINKS releases in foregoing years see the HSN website: https://iisg.amsterdam/en/hsn/projects/links/links-releases
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>Alfalab</td>
<td>KNAW software integration project (HSN focusing on IDS structure)</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>HZZ</td>
<td>HSN LINKS Zeeland (CLARIAH seed)</td>
</tr>
<tr>
<td>GGR</td>
<td>Genes, Germs and Resources</td>
</tr>
<tr>
<td>LINKS val</td>
<td>LINKing System for historical family reconstruction valorization</td>
</tr>
<tr>
<td>IDS meso</td>
<td>IDS meso: Intermediate Data Structure for organizations</td>
</tr>
<tr>
<td>LONGPOP</td>
<td>Methodologies and Data Mining Big Data based on Longitudinal Population and Epidemiological Registers</td>
</tr>
<tr>
<td>GNT</td>
<td>GIANTs of the modern world. A new history of heights and health in The Netherlands, 1811-1940’</td>
</tr>
<tr>
<td>CCH</td>
<td>CLARIAH Curing the Historical Sample of the Netherlands</td>
</tr>
<tr>
<td>ANDB</td>
<td>Life Courses of members of the diamond workers union (ANDB)</td>
</tr>
</tbody>
</table>