Contents

1 The HSN in 2020 (summary).................................................................................................................. 3

2 The HSN ................................................................................................................................................. 6
   2.1 Organizational Structure .................................................................................................................. 6
   2.2 Data Collection: Starting point and sources.................................................................................. 7
   2.3 Content of the HSN Database ........................................................................................................ 9
   2.4 Promotional activities, lectures and publications........................................................................... 12

3 HSN-Projects ......................................................................................................................................... 13
   3.1 LINKS and LINKS valorization ........................................................................................................ 13
   3.2 European Historical Population Sample Network (EHPS-Net)..................................................... 17
   3.3 LONGPOP ..................................................................................................................................... 17
   3.4 Against the Stream (ANDB) ........................................................................................................... 19
   3.5 ODISSEI ....................................................................................................................................... 19

4 Staff and Boards HSN ......................................................................................................................... 21
   4.1 Staff HSN ...................................................................................................................................... 21
   4.2 Board Foundation HSN .................................................................................................................. 21
   4.3 Steering Group HSN ....................................................................................................................... 22
   4.4 Scientific Council of Advice .......................................................................................................... 22
   4.5 International Advisory Board ...................................................................................................... 23

Appendix A Publications.......................................................................................................................... 24
Appendix B Lectures, presentations, symposia and other promotional activities............................... 26
Appendix C Reports and Working papers.............................................................................................. 27
Appendix D Releases................................................................................................................................ 28
Appendix E Project history ..................................................................................................................... 29
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 its website and journal is published by the IISH/HSN.

1 The HSN in 2020 (summary)

The year 2020 will also for the HSN be a year to be remembered. The spread of the coronavirus and the subsequential lockdown in the Netherlands from the middle of March onwards, interrupted the usual workflows in several ways. First of all, the Institute was closed, so all staff members had to work at home and all staff meetings took place online. Secondly all archives were closed as well, so data gathering was limited to what the archives presented on their websites. Thirdly the HSN became part of a new department within the IISG (‘Data and Collection Management’), which ensued several reorganisation issues including the succession of the head of the HSN, Kees Mandemakers, who officially retired from work just two days before the start of the lockdown. Luckily, he was prepared to continue leading the HSN on a part time base and teaching the ropes to his successor Richard Zijdeman. Within these limits we continued our work as much as possible on the different projects.

Besides the continuous work on the extension of the HSN and LINKS database and the EHPS Network, the HSN staff was engaged in three projects during 2020: LONGPOP, Against the Stream (ANDB, the Dutch diamond workers union) and HSN-ODISSEI. The ANDB project concerns the gathering of about 1,000 life courses of Jewish persons, most of them sampled from the ANDB membership archives. ODISSEI creates a new national data infrastructure for the social sciences. The HSN part concentrates on connecting the HSN database with contemporary data kept at Statistics Netherlands (SSD, System of Social statistical Datasets).

Definitive 2020-highlights were the three PhDs that defended a thesis which was (fully or partly) based on HSN material or related projects: Niels van den Berg, Matthias Rosenbaum-Feldbrügge and Tim Riswick.

Niels van den Berg  Matthias Rosenbaum-Feldbrügge  Tim Riswick
On the 22nd of January Niels van den Berg defended his thesis on familial clustering and transmission of human longevity. In this study he made use of data from several historical genealogical databases among which the LINKS database and the Historical Sample of the Netherlands. One week later, on the 30th of January, Matthias Rosenbaum-Feldbrügge defended his thesis about the effects of parental death during childhood and the transition to adulthood in the Netherlands. The 13th of November saw the defense of Tim Riswick concerning the impact of brothers and sisters on infant and child mortality in the Netherlands (1863-1910) and Taiwan (1906-1945).

Besides these three PhDs there was more published research. During 2020, 35 different researchers were involved in publishing and presenting studies based or partly based on the HSN and LINKS databases. In total 19 publications in relation to the HSN or using data from the HSN and LINKS databases were published. Quite a lot of them in top journals such as Aging Cell, Population studies and Economic History Review. Another highlight was the overview article of Kees Mandemakers and Jan Kok, ‘Dutch Lives. The Historical Sample of the Netherlands (1987–): Development and research’, presenting and discussing about 300 articles, dissertations and (parts of) edited volumes which were based on HSN and LINKS data.

Due to corona lockdowns the number of lectures, presentations, interviews and other promotional activities dropped to 6 after the all-time record of 63 activities in 2019. This was mainly a consequence of the cancellation of conferences of which most of them had to wait till 2021 to reappear in a virtual environment.

Work on the HSN database itself continued throughout the year. About 1,500 marriage certificates and about 500 files with life courses were entered into the database. We are still collecting data to complement cases and extra data were gathered of about 1,000 life courses.

During the year the staff of the HSN did not change, which implies that the number of HSN employees including volunteers still counted 16 persons at the end of the year. Part of the employees work part-time and some of them work outside the institute, to collect data in various archives.

The LINKS project finished a project granted by Edinburgh University to enlarge the existing release with standards of occupational titles. A new release with 281,355 different occupational titles found in the sources used by the HSN was published, including the coding of the standardized occupations in the HISCO, HISCLASS, OCC1950, HISCAM and SOCPO classifications. The work on the linking of the certificates itself continued and a new release contained about 4,2 million marriage certificates with links between the parents of a bride or a groom to their own certificate. In a second release LINKS presented 7,7 million births linked with the marriage records of the parents. From this release also a web version was produced to present these links on the WieWasWie website as well.
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 an EU-project within the framework of the Marie Skłodowska-Curie Innovative Training Network. LONGPOP is a consortium of high profile universities, research institutions and companies located in Spain, Sweden, the Netherlands, Italy, the United Kingdom, Belgium and Switzerland. LONGPOP focuses on transformations in European societies, covering family structures, fertility and the decline of mortality. It has created 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. The project officially came to an end in July 2020. In this report year we published the results of the linking of the HSN research persons with the American censuses.

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 not only concentrates 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

HSN staff (25 April 2019)

*From left to right and top to bottom: Kees Mandemakers, Rolf Wasser, Reinier Sijbrands, Kirsten Schimmel, Jan Bartman, Marja Koster, Theo Dibbets, Behice Gül, Bert Schijff, Ben Mounwes, Walfried Commandeur*  
*Not on this picture: Jos van Hees and Frits Nijstad* (Picture: Francisco Anguita).

Due to COVID-19 no general staff meeting took place in 2020.
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 published eleven articles, mainly in two developing volumes about the outreach and content of historical population databases.

Chapter 2 of this report gives an overview of the HSN organization, of the development of the database during 2020 and of the outreaching activities. Chapter 3 contains a more detailed account of the projects that we 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 2020 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 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 19th and 20th century population in the Netherlands. The sample was 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 on 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), nowadays the Dutch Family center. 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 (except about 200 from the province of South-Holland). The whole sample contains 85,408 births at the end of 2020.

![Figure 1: Development of the HSN-database, 1991-2020](image)

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 half way and the combination of death certificates and personal cards stands at almost 80 percent of the number of births. During the year about 200 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 7,000 certificates were entered of which during 1920 a number of 1,485. In total 37,279 marriage certificates are entered; in the end we expect to collect 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 88%, For the period 1812-1862 the percentage is about 71%. The reason for this lower percentage is mainly because part of the civil certificates of death have not yet been indexed, so we do not know the date and place of death (especially for the provinces of North- and South-Holland). 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 we only enter death certificates for the period before 1940 focusing on ‘young deaths’.

Table 1 Number of birth and death certificates and personal cards in HSN dataset by period of birth, 31st of December 2020

<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>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>N</td>
<td>% &lt; 16 year</td>
<td>N</td>
</tr>
<tr>
<td>1812-1862*</td>
<td>36,280</td>
<td></td>
<td>25,136</td>
<td>42.5</td>
<td>865</td>
</tr>
<tr>
<td>1863-1882*</td>
<td>16,502</td>
<td></td>
<td>9,585</td>
<td>57.5</td>
<td>5,570</td>
</tr>
<tr>
<td>1883-1922*</td>
<td>32,626</td>
<td></td>
<td>7,899</td>
<td>78.8</td>
<td>20,903</td>
</tr>
<tr>
<td>Total 31-12- 2020</td>
<td>85,408</td>
<td></td>
<td>42,620</td>
<td>52.6</td>
<td>27,338</td>
</tr>
<tr>
<td>Total 31-12- 2019</td>
<td>85,383</td>
<td></td>
<td>42,576</td>
<td>54.4</td>
<td>27,285</td>
</tr>
</tbody>
</table>

* Sampling frequency 0.75% for the birth period 1812-1872 and 0.5% for the birth period 1873-1922.
** The percentages of deaths exclude double counting (of certificates and personal cards).

Table 2 presents the number of life courses that we put into 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 brought to 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>5,827</td>
<td>93.9</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>82.5</td>
</tr>
<tr>
<td>Rest of the Netherlands</td>
<td></td>
<td>1863-1882</td>
<td>5,931</td>
<td>2,159</td>
<td>36.4</td>
</tr>
<tr>
<td>Spearhead Regions</td>
<td>X</td>
<td>1883-1922</td>
<td>6,528</td>
<td>6,309</td>
<td>96.6</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>93.2</td>
</tr>
<tr>
<td>Rest of the Netherlands</td>
<td></td>
<td>1883-1922</td>
<td>4,640</td>
<td>4,085</td>
<td>88.0</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
<td>44,252</td>
<td>37,173</td>
<td>84.0</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 and the full sample from 1903-1922 are included (other regions only 0.25% for 1903-1922).

The actual data release comprises a number of 37,173 life courses. Table 2 also presents 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 3,000 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 2,500 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 2020 all software had been developed but parts of it still needed to be tested thoroughly, especially the software to handle 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 2020 this website was visited 2,025 times, and the number of page views was 7,191.

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 2020 the EHPS-Net portal generated 4,065 visitors and 15,273 page views.

In 2020 eighteen HSN related articles and books were published (see appendix A for an overview), together with one working paper (Appendix C), this amounts to nineteen publications. Highlights were the three dissertations of Niels van den Berg, Matthias Rosenbaum-Feldbrügge and Tim Riswick and the overview of all work within the context of the HSN and LINK over the last 30 years by Kees Mandemakers and Jan Kok.

Due to the coronavirus pandemic and the following lockdowns worldwide only six conference contributions, lectures and other public activities were counted, presented in the Netherlands and abroad. This number of six contributions is a sharp contrast with the number of 63 in 2019 which was an all-time record.

In total five data sets for research were released. Three by the LINKS project: a release with marriage links and one with links between births and the marriages of the parents. From the last one also a version was created to publish the links on the website of WieWasWie. The HSN released a new version of the standardized and coded values, now with 281,355 different occupational titles found in the sources used by the HSN until the 1st of January 2020 and a new version with the data from the civil certificates. Secondly, we published a new improved release with the civil certificates of the HSN Research Persons.

In the publications and presentations 35 different researchers were involved (because the drop in the number of presentations this number was 17 less than that in 2019).

Figure 2 Development of the number of publications and presentations, 1991-2020
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:

- designing and maintaining a data structure for use by individual researchers;
- 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 2020 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](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. During the year we worked on new releases, the implementation of the software on the servers of HUC Digital Infrastructure and the development of a new and faster system to link the data.

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/](https://www.wiewaswie.nl/en/)). Nowadays the index contains names from about 10.5 million birth certificates of the period 1812-1920, names from about 4.5 million marriage certificates of the period 1812-1945 and names from about 12 million death certificates of the period 1812-1970. 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 2020 we worked on improving the system, especially software to fasten the handling, reading and matching of the data. Within the context of the CLARIAH project Joe Raad (VU Computer Science) worked on the development of a new and especially faster linkage system than the present one.

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 found 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 and Fons Laan (linking birth, death and marriage certificates for the provinces of Groningen and Zeeland). During 2020 two datasets were released, one with the links between marriages over two generations and one with the births linked with the marriages of the parents. From the last one we also constructed a ‘WieWasWie-version’.

During 2020 several studies were published which were fully or partly based on LINKS data among which the dissertation of Niels van den Berg. This dissertation was written within the context of the project Genes, Germs & Resources (P.I. Angélique Janssens and Eline
Slagboom). Besides these dissertations LINKS data formed the basis for several articles under which one in *Aging Cells* (Niels van den Berg et al) on the subject on longevity and genetics.

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 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).

Niels van den Berg, Rick Mourits and others published an article in *Population Studies* in which for the province of Zeeland for the birth cohort 1863-1872 a systematic comparison was made between the outcomes of the HSN and the LINKS database.

In September 2018 the first release of the LINKS valorization project was forwarded to the Dutch Family Center which is the organization behind the WieWasWie data collection. On 23 April 2019 the already existing cooperation between the Dutch Family Center and the IISG was formalized in an agreement to continue and intensify the data sharing. In December 2019 an updated revised version saw the light which contained about 5.4 million links between the parents of a bride or a groom on a marriage certificate and their own marriage certificate. In 2020 we constructed a new type of release with 8.5 million births linked with the marriage certificates of their parents. See figure 4 for the way these links are presented at the website of WieWasWie. The orange blocks present 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.
In cooperation with the department of economics of Edinburgh University the HSN worked on a new release with standardization and classified occupational titles. Thanks to a grant of Edinburgh University it was possible to offer Jan Hornix a contract to work on these titles. The release was finished in August 2020 and included 281,355 different occupational titles (about 150,00 more than in the foregoing release). New in this release was also a systematic classification of so-called ‘double’ titles and ambiguous titles (see https://iisg.amsterdam/en/hsn/data/occupations).

Figure 4 Still from the website ‘WieWasWie’
3.2 European Historical Population Sample Network (EHPS-Net)

The European Historical Population Sample 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. Key activities of the network are the website including the collaboratory and the IDS repository and the e-journal.

During 2020 several activities like the development of extraction software were continued under the flag of the LONGPOP project (see section 3.3); several data extraction programs were published on the EHPS website. The e-journal *Historical Life Course Studies* published eleven articles. Most of them within two special volumes designated to give overviews of the use and content of large databases with demographic historical micro-data.

![Participants of the last general meeting of EHPS-Net (Picture: Marja Koster)](image)

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 it 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. Although most of the practical work was finished during the foregoing year, the project was officially closed at the end of July 2020. In this year the IISG Early Stage Researchers, Diogo Paiva and Francisco Anguita, published their results of the linking of the HSN Research Persons with the persons listed in the USA censuses over the period 1850-1940.
Part of the project ‘Against the stream. Social Mobility of Dutch Jews, 1880-1940’ is the analysis of the social mobility and integration of Dutch Jews in the late 19th and early 20th century. The project is granted by the Royal Academy of Science and is a cooperation project with Huygens ING (https://iisg.amsterdam/en/research/projects/tegen-de-stroom) and supervised by Karin Hofmeester, Leo Lucassen and Lex Heerma van Voss.

The mobility part of the project is realized by Joris Kok. He will focus on individual life strategies employed by Dutch Jews and the relation between these strategies and outcomes in integration, social mobility, and (de)segregation. Among the strategies studied will be residential mobility, inter-/intragenerational occupational mobility, intermarriage, and other demographic patterns.

For this research a special sample (n=800) was drawn from the database of the Diamond Workers Union the ANDB (Algemene Nederlandse Diamantbewerkersbond, https://iisg.amsterdam/nl/onderzoek/projecten/andb). Microdata on diamond workers will allow for comparing trajectories of Jewish and non-Jewish diamond workers. Life courses from the Historical Sample of the Netherlands will be used to extend this analysis to Jewish workers outside the diamond industry and the general Dutch population.

The Amsterdam diamond industry employed a large share of Dutch Jews. Besides the mentioned ANDB sample and the basis dataset the HSN will also supply life course data from an earlier project ‘Dutch Jews or Jewish Dutch’ (JDJ, see https://iisg.amsterdam/en/hsn/projects/jdj). This dataset was enlarged with the birth cohort 1873-1882 (n=200).

During the report year most of the collecting, coding and entering of the data was done. The dataset is aimed to be finished early 2022.

3.5 ODISSEI

ODISSEI (Open Data Infrastructure for Social Science and Economic Innovations) is the national research infrastructure for the social sciences in the Netherlands. ODISSEI brings together researchers with the necessary data, expertise and resources to conduct groundbreaking research and embrace the computational turn in social enquiry (see https://odissei-data.nl/). One of the aims of the ODISSEI Observatory is to broaden the types of data that are integrated into ODISSEI and the Observatory seeks to integrate data from new sources and open up new lines of research. One new source are data from the Historical Sample of the Netherlands (HSN) which will be linked with micro-data from the current Statistics Netherlands SSD (System of Social Statistical Datasets). This will enable links between the historical research conducted using the HSN and contemporary society and the outcomes as studied by social scientists, also enabling stronger links with the CLARIAH infrastructure.

In the long run linkage between the SSD and the HSN needs to enlarge the HSN database with the life courses of the second generation of the original HSN Research Persons born between
1883 and 1922 (see figure 5). This will be an enlargement of the sample with an estimation of 70,000 life courses. However, for the ongoing project the goals are limited to the following:

1. Proof of concept of the linkage of HSN Research Persons (and their children) with the SSD
2. Continuation of the work on the life courses of the HSN Research Persons first generation, especially from the birth period 1903-1922
3. Data entry of a limited number of Personal Cards of the second generation

During the report year efforts were concentrated on the second goal which progress is reported in chapter 2.3 (development of the HSN database).

Figure 5 Lexis diagram showing the main sources of the HSN’s life course data and the integration with the SSD dataset
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 group and the research department of the IISH is managed by Karin Hofmeester.

Seven workplaces were available for the work in the archives and the data entry in the office (among two SWV workplaces). Cor Munnik continued as a volunteer to work on the HSN software; Fons Laan continued his work on LINKS by way of the LINKS valorization project. Jan Hornix continued his work on occupational coding till the summer after which he switched to the ANDB project. Jos van Hees ended his volunteership at the end of the year.

At the end of 2020 the total number of HSN-staff, directly and in cooperation with other organizations, was 16 persons (2019: 16 persons). During the year, a total of 16 persons worked for the HSN, among whom eight volunteers, who were engaged in collecting material in archives and in data entry.

Staff in 2020:

<table>
<thead>
<tr>
<th>Name</th>
<th>FTE</th>
<th>Period</th>
</tr>
</thead>
<tbody>
<tr>
<td>J. Bartman</td>
<td>0.4</td>
<td>Jan. - Dec.</td>
</tr>
<tr>
<td>W. Commandeur</td>
<td>0.3</td>
<td>Jan. - Dec.</td>
</tr>
<tr>
<td>Th. Dibbets</td>
<td>0.4</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>J. Hornix</td>
<td>0.6</td>
<td>July - 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>0.8</td>
<td>Jan. - Dec.</td>
</tr>
<tr>
<td>dr. F. Laan</td>
<td>0.2</td>
<td>Jan. - Dec.</td>
</tr>
<tr>
<td>prof. dr. C.A. Mandemakers</td>
<td>0.8</td>
<td>Jan. - Dec.</td>
</tr>
<tr>
<td>drs. B. Mouwes</td>
<td>0.3</td>
<td>Jan. - Dec.</td>
</tr>
<tr>
<td>drs. C. Munnik</td>
<td>0.2</td>
<td>Jan. - Dec.</td>
</tr>
<tr>
<td>dr. B. Schijf</td>
<td>0.2</td>
<td>Jan. - Dec.</td>
</tr>
<tr>
<td>drs. K. Schimmel</td>
<td>0.4</td>
<td>Jan. - Dec.</td>
</tr>
<tr>
<td>R. Sijbrands</td>
<td>0.8</td>
<td>Jan. - Dec.</td>
</tr>
<tr>
<td>drs. R. Wasser</td>
<td>1.0</td>
<td>Jan. - Dec.</td>
</tr>
</tbody>
</table>

4.2 Board Foundation HSN

During the year three new members entered the board: Marian Beekman, Hilde Bras and Karin Hofmeester. Karin Hofmeester replaced Leo Lucassen as representative of the management team of the IISG in April 2020. Another person who left the board was Frans van Poppel who retired after a period of duty of over thirty years. We own him a lot, not only as chair, but also as a very active fund raiser and researcher. Jan Kok replaced Frans as chair of the steering
group. At the end of 2020 the board consisted of the following persons (the year of resignation is between brackets):

Dr. M. (Marian) Beekman, Leiden University Medical Center, member (2024)
Prof. dr. C.J.H. (Catrien) Bijleveld, VU University Amsterdam, director Netherlands Institute for the Study of Crime and Law Enforcement (NSCR), member (2022)
Dr. H. (Hilde) Bras, University of Groningen, member (2024)
Prof. dr. 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. K. (Karin) Hofmeester, head research department IISG, University of Antwerp (2025)
Prof. dr. J. (Jan) Kok, Radboud University Nijmegen, IISG, chair (2025)
Prof. dr. M.H.D. (Marco) van Leeuwen, Utrecht University, vice chair (2025)
Dr. F.R.M. (France) Portrait, VU University Amsterdam, treasurer (2021)
Dr. P.G. (Vincent) Tassenaar, University of Groningen, member (2021)


The board met on 2 February, 28 May and 5 October 2020. Main items on the agenda were the progress of the several projects and ongoing applications.

4.3 Steering Group HSN

The board of the HSN acts as a steering group and is the decision-making body regarding HSN activities. The steering group may include also other staff from the IISH. Advisor of the steering group is prof. dr. C.A. Mandemakers, head of the HSN.

The steering group had three meetings on 2 February, 28 May and 5 October 2020.

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 several mutations in the Board during 2020: dr. Diego Ramiro-Fariñas, Instituto de Economía, Geografía y Demografía, Madrid left the board after the end of his term in 2019 and the board was extended with two new members: prof. dr. Chris Dibben, University of Edinburgh and prof. dr. Hilde Sommerseth, Arctic University, Tromsø. Term of each board member is seven years. The composition of the Board is as follows:

Prof. dr. C. (Cameron) Campbell, University of Science and Technology, Hong Kong
Prof. dr. C. (Chris) Dibben, Geography (Health and Environment), University of Edinburgh, Director of the Longitudinal Studies Centre Scotland.
Prof. dr. L. (Lisa) Dillon, Département de Démographie, Université de Montréal
Prof. dr. M. (Martin) Dribe, Centre for Economic Demography, Lund University
Prof. dr. H.L. (Hilde) Sommerseth, Norwegian Historical Data Centre, Arctic Univ. Tromsø
Prof. dr. E. (Evan) Roberts, University of Minnesota, Minneapolis
Prof. dr. H. (Hélène) Vézina, l'Université du Québec à Chicoutimi (UQAC)

The Board met virtual in Amsterdam on 4 December 2020. The various projects of the HSN and the future of the HSN were discussed.
Appendix A  Publications

2020


For the publications in foregoing years see the HSN website: https://iisg.amsterdam/en/hsn/products/publications. For citations h-index and i10-index of the HSNDB see: https://scholar.google.com/citations?hl=en&user=9xP8wkQAAAAJ.
Appendix B  Lectures, presentations, symposia and other promotional activities

2020

807 Quanjer, B. (2020, October 16). *Height and the disease environment, the impact of infant mortality on height in the Netherlands 1850-1940*. Annual Posthumus Conference (online).


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

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

2020


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

2020  HSN


73 Historical Sample of the Netherlands (HSN) (2020). Civil Certificates (release 2020.01 (n=85,383)) [Data set].

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

2020  LINKS

39 Mandemakers, K., & Laan, L. (2020). LINKS Dataset Linked Marriages, Netherlands, 1796-1943 (Release 2020.03 (n=4,158,388); also a version including first names of bride/groom and parents Release 2020.03_f).

38 Mandemakers, K., & Laan, L. (2020). LINKS dataset Linked Births and Marriage Certificates Parents, Netherlands, WieWasWie website (Release 2020.02 (n=8,087,449)).

37 Mandemakers, K., & Laan, L. (2020). LINKS Dataset Linked Births and Marriage Certificates Parents, Netherlands (Release 2020.01 (n= 7,695,006)).

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 thirty 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>HLN</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>LINKS oce</td>
<td>LINKing System for historical family reconstruction standardization occupational titles</td>
</tr>
<tr>
<td>ANDB</td>
<td>Against the Stream, Life Courses of members of the diamond workers union (ANDB)</td>
</tr>
<tr>
<td>ODISSEI</td>
<td>Linking the HSN with the SSD, system of social statistical datasets</td>
</tr>
</tbody>
</table>