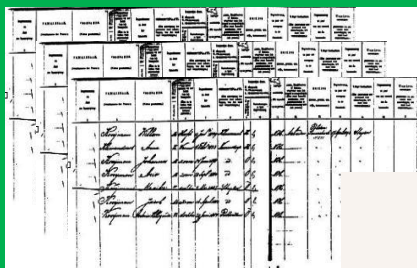


Historical Sample of the Netherlands

HSN

Annual Report 2016



DATE	DESCRIPTION	AMOUNT	CURRENCY
1798
1799
1800



International Institute of Social History, Amsterdam

Historical Sample of the Netherlands

Annual Report 2016

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

Contents

1	The HSN in 2016 (summary).....	5
2	The HSN.....	7
2.1	Organizational Structure.....	7
2.2	Data Collection: Starting point and sources.....	9
2.3	Content of the HSN Database.....	11
2.4	Promotional activities, lectures and publications.....	14
3	HSN-Projects.....	16
3.1	Genes, Germs and Resources.....	16
3.2	LINKing System for historical family reconstruction (LINKS).....	17
3.3	European Historical Population Sample Network (EHPS-Net).....	19
3.4	CEDA_R project.....	21
3.5	IDS meso.....	22
3.6	LONGPOP.....	23
3.7	GIANTS.....	24
4	Staff and Boards HSN.....	25
4.1	Staff HSN.....	25
4.2	Board Foundation HSN.....	26
4.3	Steering Committee HSN.....	26
4.4	Scientific Council of Advice.....	26
4.5	International Advisory Board.....	27
Appendix A	Publications.....	28
Appendix B	Lectures, symposia and other promotional activities.....	30
Appendix C	Reports and Working papers.....	33
Appendix D	Releases.....	34
Appendix E	Project history.....	35

The Historical Sample of the Netherlands (HSN) is an initiative of researchers from different disciplines within the social and historical sciences. The goal of the HSN is to create a representative database of nineteenth- and twentieth-century life courses. A sample of the birth certificates from the period 1812-1922 serves as the basis for the HSN database.

1 The HSN in 2016 (summary)

In the beginning of the year we were shocked by the news of the unexpected death of Jaap Dronkers at the age of 71. Jaap, one of the founding fathers of the HSN, had not only a sharp eye for social questions of the day especially in the field of education and social inequality, but as a historically interested researcher he was also very important for the development of quantitative historical research in the Netherlands. Jaap Dronkers not only gave the study of the role of the nobility and other elite networks a strong impetus, he also stood at the basis of the realization of the Historical Sample of the Netherlands.



Picture: NSV-sociology

Meetings that took place from 1986 onwards, some of them in his house in Haarlem, in which sociologists, demographers and geographers participated, resulted in 1989 in the creation of the Historical Sample of the Netherlands (HSN). Jaap remained involved in the HSN as chair of the Advisory Board and kept close track of the development of the HSN. The dozens of researchers who used the data of the HSN over the last 25 years and those who will do so in future, owe a lot to the work and enthusiasm of Jaap. We will keep in memory Jaap's passionate commitment, his sharpness in discussions, his critical attitude, his humour and his typical ironic smile.

Besides the continuous work on the extension of the HSN and LINKS database, the HSN staff was engaged in six projects during 2016. Four projects were finished this year: the demonstrator for an Intermediate Data Structure of organizations (IDS Meso), *Genes, Germs and Resources*, EHPS-Net and CEDA-R. Two new projects: GIANTS, collecting heights from militia registers, and LONGPOP, employing two so-called Early Stage Researchers (ESR), have started this year. The work on the HSN database got a fresh impetus by the grant of 200,000 euro by way of the CLARIAH infrastructure project.

Genes, Germs and Resources studies the phenomenon of familial influences on early death and exceptional survival in the Netherlands between 1812 and 2015. It takes into account the simultaneous effects of *resources, germs and genetic influences*. The HSN has built two datasets for this project. One starts with the HSN basic set from the birth period of 1860-1875, selecting all children of these research persons and collecting their personal cards. In a second step all personal cards of the third generation are collected and entered. The second database uses the LINKS dataset as a starting-point to reconstruct the families of provinces with relatively rich indices of the civil registers. At the end of 2016 the project was closed with the release of a database of about 40,000 persons originating from the modern Dutch population register.



For quite some time several important databases with historical life course data have been working together to develop comparable datasets and joint software. In 2011, a grant of the European Science Foundation (ESF) gave this cooperation a strong impulse when the *European Historical Population Sample Network* (EHPS-Net) was founded. Within this network, we concentrate not only on the creation of common data structures and software, but also on education by way of summer schools, on developing new databases and on the publication of results in an E-journal. The HSN is chair of this project in which over ten countries and twenty databases are cooperating. In 2016 several working groups were engaged on all kind of topics relevant for our kind of research and on data gathering and data mining. We had summer courses in Cluj, Lund and Nijmegen, the E-journal, *Historical Life Course Studies* continued and overreached the goal of minimal five articles per year. A winter school was held on the ‘Intermediate Data Structure and Extraction Software’ at the IISH in Amsterdam at the end of November.

The CEDA_R project aims at the conversion of about 2,288 spreadsheets in which the Dutch censuses 1830-1947 were transcribed. For this period the results are available at the aggregative level of municipality or province. All spreadsheets were converted into one database in order to convert the data into the RDF structure. The team concentrated on testing and developing the RDF system and the harmonization of those spreadsheets that contain population and dwellings for each location in the Netherlands. In the spring Albert Meroño-Peñuela successfully defended his thesis at Free University in Amsterdam. The IISH based PhD, Ashkan Ashkpour, is expected to defend his thesis in 2017.

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, Sweden, the Netherlands, Italy, the United Kingdom, Belgium and Switzerland. 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. 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. One works on the documentation and extension of IDS related extraction software and one works on methods of standardizing addresses.

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 30,000 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. Principal Investigator of this project is Jan Kok from Radboud University Nijmegen. The project is financed by the Dutch National Scientific Organization.

In 2016 about 15 publications and 4 working papers in relation to the HSN database or using data from the HSN database were published. This figure includes the dissertation of Julia Efremova, *Mining Social Structures from Genealogical Data* in which HSN and LINKS data were used. The number of lectures, presentations, interviews and other promotional activities amounted to 37. In total about 49 different researchers were involved in these activities.



Julia Efremova

At the conference of the 'European Consortium for Sociological Research', held 22-24 September 2016 in Oxford, UK, Antonie Knigge was awarded with the ECSR Prize for best PdD thesis written in 2015. His thesis is entitled *Sources of Sibling Similarity. Status Attainment in the Netherlands during Modernization*. The analyses in his research are based on data from the LINKS project.

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 complete the death certificates (including personal cards) till a level of 85% of all research persons and to cure about 5,000 life courses to be added to the database at the end of 2018. This year about 7,000 death certificates were entered into the database and extra data were gathered of about 500 life courses.

The number of HSN employees including volunteers increased from 15 to 19 at the end of the year. Part of the employees work part-time and some of them work away from the institute, to collect data in various archives.

Chapter 2 of this report gives an overview of the HSN organization, of the development of the database during 2016 and of the outreaching activities. Chapter 3 contains a more detailed account of the six projects that we have worked on. Chapter 4 presents the composition of the staff and the several boards of the HSN.

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

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.



HSN staff (14 June 2016)

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

Picture: Richard Zijdeman

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

In order to guarantee continued existence and accessibility of the HSN database, the HSN Foundation has linked itself by contract to the International Institute of Social History (IISH) in Amsterdam, which forms part of the Royal Netherlands Academy of Sciences (KNAW). The IISH is an internationally renowned archive and research institute in the field of social history. It is devoted to the acquirement, management and accessibility of collections in that area.

The International Institute of Social History (IISH) provides housing for the HSN's activities and assumes the burden of the resulting costs. The IISH has guaranteed a permanent position for coordination tasks. The actual data gathering is done on the basis of projects, which are externally funded. The HSN is part of the IISH research department. Decisions regarding projects are made by the Steering Group which consists of members of the Board of the HSN and members of the management team of the IISH (for the composition of these boards, see chapter 4).



IISH building

2.2 Data Collection: Starting point and sources

The Historical Sample of the Netherlands (HSN) strives to construct life histories as completely as possible for a representative portion of the nineteenth and twentieth century population in The Netherlands. The sample has been drawn from all persons born in The Netherlands between 1812 and 1922. Ultimately, the HSN database will include information on an individual level from about 85,000 persons on subjects like family structure, occupation, 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 persons born in the Netherlands during the 19th and early 20th century (in total about 14,5 million persons) at the age of 20.

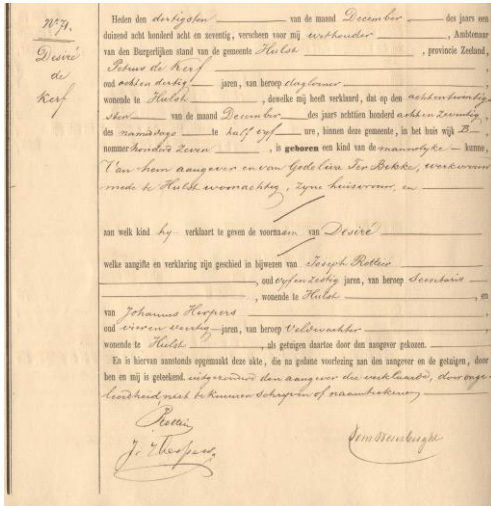
The Historical Sample of the Netherlands (HSN) strives to construct life histories as completely as possible for a representative portion of the nineteenth and twentieth century population in The Netherlands. The sample has been drawn from all persons born in The Netherlands between 1812 and 1922. Ultimately,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



(Birth certificate of Desiré de Kerf, Hulst, 28 December 1878 Zeeuws Archief).

The sample is drawn from the birth certificates and stratified in periods of ten years. To achieve rather equally sized cohorts of persons from the age of twenty years, depending on infant and child mortality on the one hand and the number of births on the other hand, it was decided to have two sample frequencies: 0.75% for the period 1812-1872 and 0.5% for 1873-1922. This results in a sample size that is large enough to make sound statistical conclusions for subpopulations of minimal two percent of the persons born in the Netherlands during the 19th and early 20th century (in total about 14,5 million persons) at the age of 20.

The basic dataset of the HSN contains the most important data from the life courses of the sampled persons. The data about birth and death originate from the certificates of birth (see picture) and death. For the period after 1939 we use the personal cards instead of the death certificates. The certificates of death and marriage also comprise data about occupational titles and places of living of the parents and other relatives. Marriage certificates contain data about place of living, occupation, age, illiteracy (whether or not being able to write a signature) of both bride and groom, their parents and four witnesses (usually relatives like brothers or close friends).

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

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

In the late 1930s, the population register was replaced by the personal card; from that time on the individual person became the registration unit in all municipalities. Since then the population register in each municipality has consisted of a collection of personal cards, containing nearly the same information as the population register. All persons who were alive at 1 January 1940 or were born after that year received a personal card. At the time of death, this card is removed from the files and sent to the Central Bureau of Statistics ([CBS](#)), where the data on the card are used for statistical purposes; and then it is sent to the Central Genealogical Bureau ([CBG](#)). Copies of the cards have been used for the data set. They contain the following information: name, municipality and date of birth of the person concerned, as well as those of his or her parents, marriage partner(s) and children. The nationality is given as 'Dutch' or 'Foreign'. Successive occupations, addresses and changes therein are also indicated. 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. This archive is used by the HSN to get data for sampled persons who died after 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 register.

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

The maximum number of all sources to be entered for the cases is defined by the number of birth certificates. From figure 1 it is clear that for the life courses we are nearly half way and for the combination of death certificates and personal cards at about three quarters of the number of births. During the year about 5.000 death certificates were added to the database and extra data of 500 life courses were collected. During 2016 we restarted working on the marriage certificates and entered about 1,200 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 1863-1882 the percentage of found death records is about 88%, For the period 1883-1922 the percentage is about 6% lower, due to the ongoing data entry of birth certificates for the period 1903-1922 and because some persons are still alive. In the early years of the HSN the focus was on the data entry of death certificates of infants and children. This means that these deaths are still overrepresented in the HSN database, although for the two last periods the percentages in the table exaggerate the situation since date entry of death certificates ends in 1940 (personal cards

are not included in this percentage including most of the grown up HSN research persons from these periods of birth).

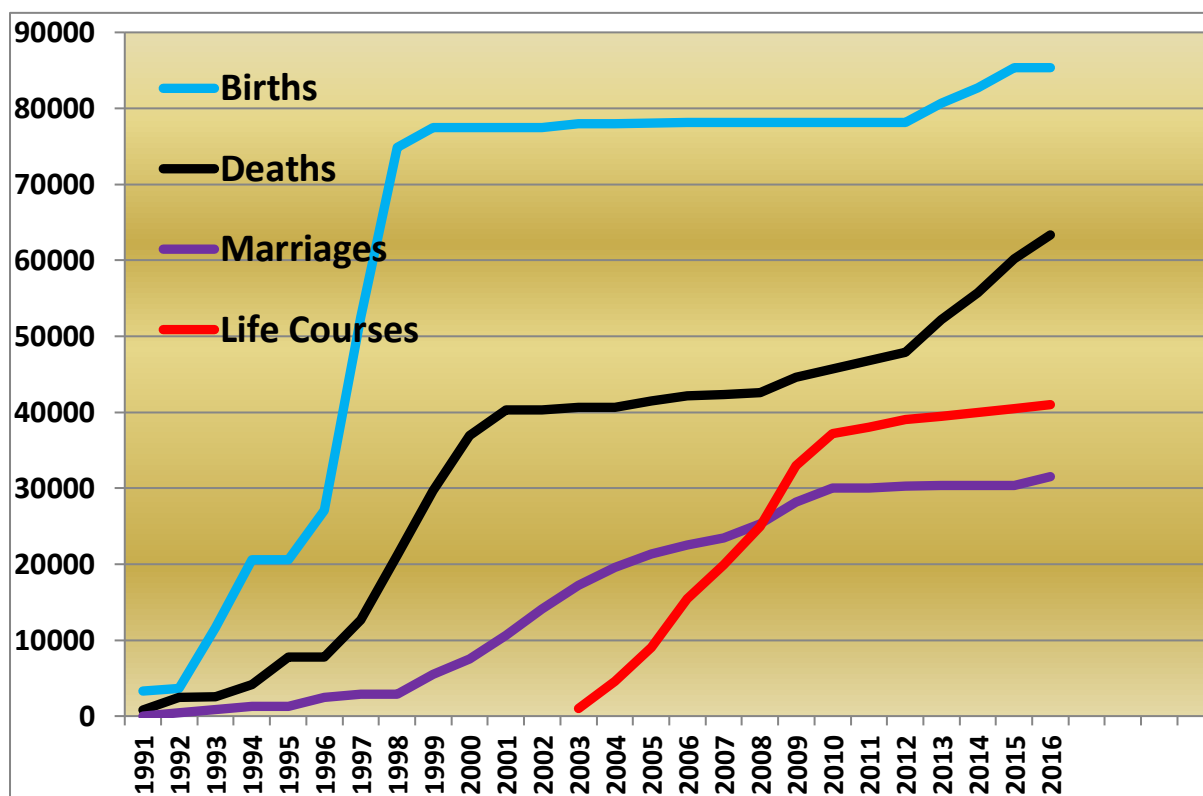


Figure 1 Development of the HSN-database, 1991-2016

Table 1 Number of birth and death certificates and personal cards in HSN dataset by period of birth, 31st of December 2016

Period	HSN Basic Sample (Number Birth Certificates)	Death Certificates and Personal Cards (PK) and Personal Lists (PL)				
		Death Certificates		PK's & PL's	Death Certificates and PK's & PL's	
		N	% < 16 year	N	N	% Basic Sample**
1812-1862	36,280	21,587	48.4	750	22,337	61.6
1863-1882	16,502	9,141	59.4	5,351	14,492	87.8
1883-1922*	32,572	7,185	78.9	19,325	26,510	81.4
Total 31-12- 2016	85,354	37,913	56.8	25,426	63,339	74.2
Total 31-12- 2015	85,354	32,843	63.2	25,329	58,172	68.2

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

Table 2 Number of Life Courses by region, date of birth and priority of data entry, HSN Release 2010.01

Region	Priority	Period of Birth	Total	In release		Complete Life Course	
				N	%	N	%
Spearhead Regions	X	1850-1882	6,208	5,827	93.9	4,179	67.3
Rest of the Netherlands	X	1863-1882	6,795	5,608	82.5	4,009	59.0
Rest of the Netherlands		1863-1882	5,931	2,159	36.4	1,785	30.1
Spearhead Regions	X	1883-1922	6,528	6,309	96.6	4,805	73.6
Rest of the Netherlands	X	1883-1922	14,150	13,185	93.2	10,113	71.5
Rest of the Netherlands		1883-1922	4,640	4,085	88.0	3,081	66.4
Total			44,252	37,173	84.0	27,972	63.2

* 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). Figure 2 shows the different stages in the production process for all cases from the birth period 1843-1922 which are not included in the release 2010.01. After the data collecting in the region of birth the files are checked on completeness. In case a person moved to another area, the file moves to the 'mail-system' in which we produce requests for copies of registers filling the gaps that we found in the life courses. So, for the life courses we concentrated on collecting and data entry during 2016.

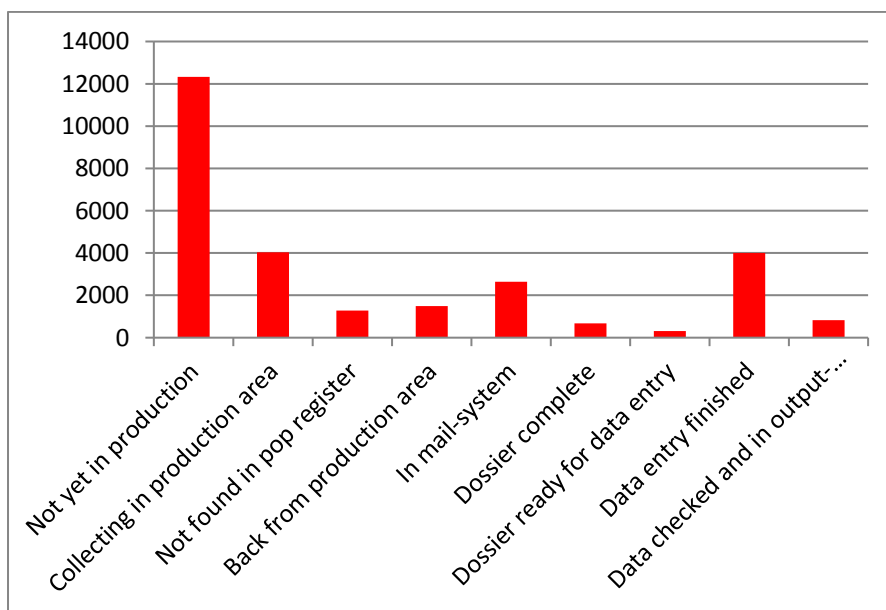


Figure 2 Number HSN Research Persons with no finished life course, per production stage, 1843-1922 (n=27,599)

Within the CLARIAH cure project we aim a) to bring the number of death certificates and personal cards on 85% of the sample, b) to enter 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 (The Hague and Rotterdam also from 1903-1922).

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 2016 all software had been developed but parts of it still needed to be tested thoroughly.

2.4 Promotional activities, lectures and publications

In addition to the work directly connected to the database, activities were developed to raise our profile at home and abroad. One of the tools to achieve this is the HSN website. In 2016 this website was visited 4,843 times, the number of pages visited was 10,876 and there were 2,868 unique visitors. Apart from our website and the website of the [Life Courses in Context](#) project (LCC), launched May 2004, the HSN has also been responsible since December 2005 for the renewed website of the [International Commission for Historical Demography \(ICHD\)](#). This website offers, besides general information about the ICHD, information on international meetings, publications and links related to historical demography. In September 2016 the website moved to the Radboud University of Nijmegen.

The LCC website received 12,024 visitors (3,542 unique visitors), with 18,411 page views. The ICHD website received 15,324 visitors (6,418 unique visitors), with 29,262 page views.

At the end of 2012 the portal of the [European Historical Population Samples Network \(EHPS-Net\)](#) was launched. Part of this portal is a [collaboratory](#) in which about eighty researchers participate. In 2016 the portal generated 5303 visitors and 40.003 page views.

In 2016 15 HSN related articles and books were published (see appendix A for an overview), including the four working papers (Appendix C), this amounts to 19 publications. In 2016 37 presentations and papers were written, both in the Netherlands and abroad (please refer to appendix B for an overview). As usual, with seven presentations/papers, the HSN had a strong presence at the *11th European Social Science History Conference*, Valencia, Spain, 30 March- 2 April 2016. At the second *Biennial Meeting of the European Society of Historical Demography (ESHD)*, KU Leuven the HSN was presented with three presentations. In the publications and presentations 49 different researchers were involved (in 2015 52 researchers).

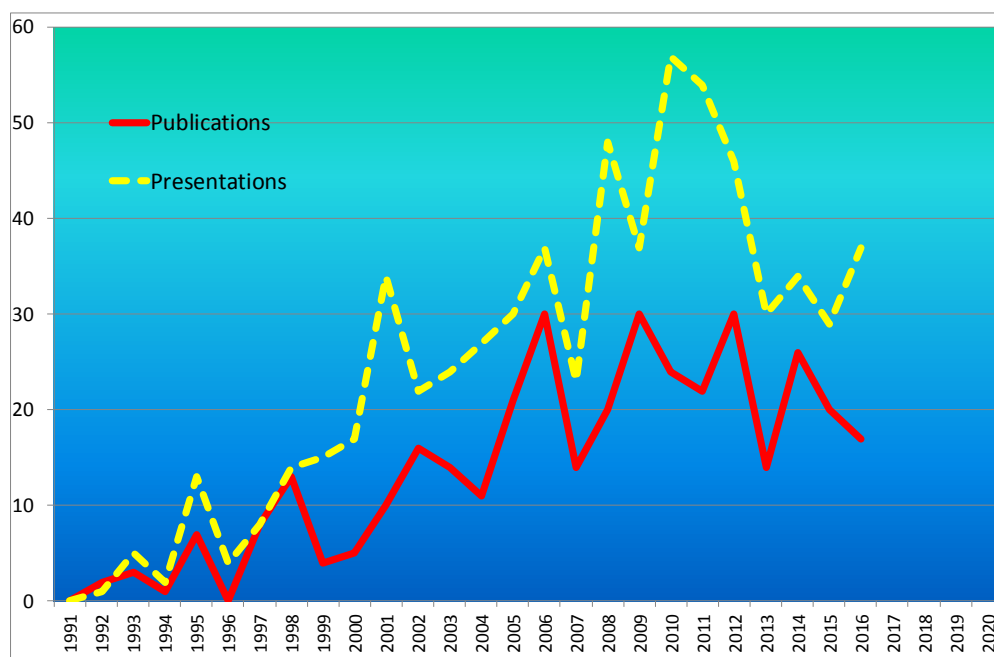


Figure 3 Development of the number of publications and presentations, 1991-2016

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, both 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 2016 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://socialhistory.org/en/hsn/hsn-projects>).

3.1 Genes, Germs and Resources

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



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

The project has several innovative aspects, among which the introduction of genetics into the study of historical mortality as recent advances in human genetics have shown the relevance of the genetic component for longevity and mortality.

The HSN constructed two datasets for this project. The database *Long Lives* started with the HSN basic set from the birth period 1860-1875. Two groups were distinguished: a) persons

who reached the age of 80 years or older and b) a control group with persons who died before the age of 65. Data about the off-spring of these persons were gathered from the population register (second, third and fourth generation). At the end of 2016 the project was closed with the release of a database with about 40,000 persons originating from the modern Dutch population register.

The database *Linked Families* contains data from the LINKS project in a GGR-project related format. It reconstructs the families for provinces with relatively rich indices of the civil registers (Zeeland, Groningen, Drenthe). During 2016 several beta-releases were delivered which were commented on by the researchers and which will be further sophisticated during 2017.

3.2 LINKing System for historical family reconstruction (LINKS)



LINKS entails the development of software for a **LINKing System** for historical family reconstruction and was financed by the CATCH program of the Netherlands Organisation 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 especially for researchers working on the *Genes, Germs and Resources* project (see chapter 3.1).

For more than fifteen years volunteers have been indexing civil records at the Dutch provincial archives, insofar these records are accessible to the general public. Since 2012 the Central Agency of Genealogy (CBG) has taken over the organization collecting all the indexes from the provincial archives in one big data base (previously GENLIAS) and it maintains the website that makes the data accessible to the general public (<https://www.wiewaswie.nl/en/home/>). Nowadays the index contains names from more than 7,000,000 birth certificates of the period 1812-1916, names from more than 3,500,000 marriage certificates of the period 1812-1941 and names from more than 10,000,000 death certificates of the period 1812-1966. These indexed names are a multiple of the number of acts, because the acts are indexed for more than just one name; for marriage acts e.g. not only the names of the bride and groom are indexed, but also the names of both parents.

The LINKS (**Linking system for historical family reconstruction**) project is a cooperation of LIACS, NIDI, the Meertens Institute, the CBG and the organizations behind GENLIAS/WieWasWie (mainly Dutch regional archives) granted by the CATCH-program (Continuous Access To Cultural Heritage) of the Netherlands Organization for Scientific Research. The project started in June 2009 and was finalized at the end of 2014.

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 the last two years we worked on software to fasten the handling, reading and matching of the data.

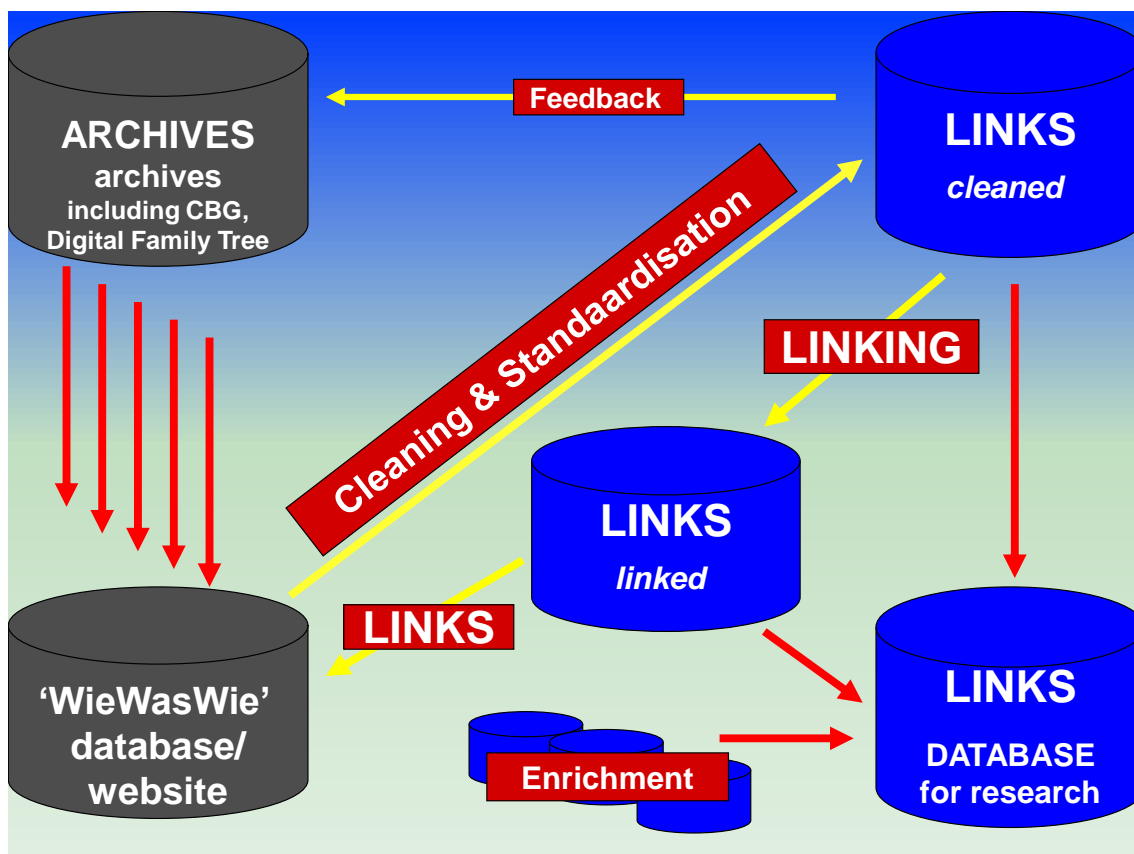
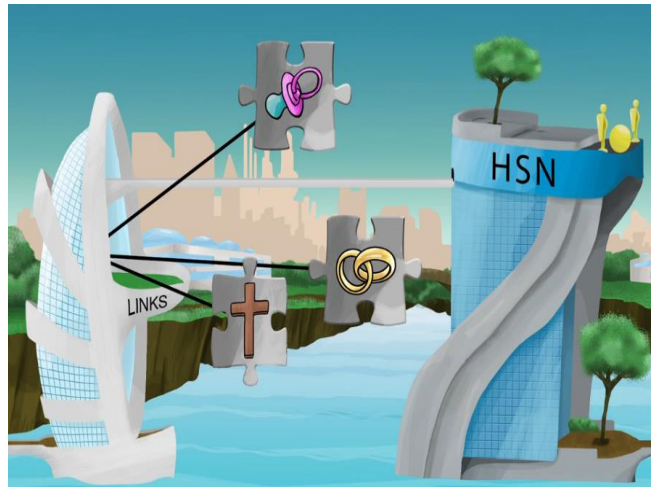


Figure 4 Dataflow LINKS project

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.

The LINKS project was officially closed with a conference in February 2014. The book with conference papers was published by Springer in September 2015: Gerrit Bloothoof, Peter Christen, Kees Mandemakers & Marijn Schraagen (eds.), *Population Reconstruction*.

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).



Still from the HSN CLARIAH Video

3.3 European Historical Population Sample Network (EHPS-Net)

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

In January 2012 the first conference was held in Amsterdam and in September 2012 the second one in Budapest bringing together about 40 scholars and database administrators from all over Europe and Northern America. In September 2014 the third one was held in Alghero. On 20 September 2016, the fourth general meeting of the EHPS-Net took place at KU Leuven, Belgium. The meeting was organized as a pre-meeting to the second conference of the European Society of Historical Demography (ESHD). Twenty-nine participants from seventeen different countries attended the meeting. The meeting was the closing meeting of the ESF program EHPS-Net. 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.



Participants of the fourth general meeting of EHPS-Net, 20 September 2016

During the project period, 2011-2016, the main databases converted their material to the IDS format. In the meantime, data extraction programs for different types of studies (e.g., on migration and fertility) were being prepared in close collaboration between researchers and programmers. The intended system is open, scalable and extendable. New types of analysis can be introduced by adding new extraction modules. Anyone can contribute an extraction module, which will be peer-reviewed and published.

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

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

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

All working groups met each other at least once. Extraction software will be released in the course of 2017. In 2016 several working groups were engaged on all kind of topics relevant for our kind of research and on data gathering and mining. We had summer courses in Cluj and Lund, the E-journal, *Historical Life Course Studies* continued and reached the goal of minimal five articles per year. A winter school was held on the 'Intermediate Data Structure and Extraction Software' at the IISH in Amsterdam at the end of November 2016. The program ran for five years and ended in November 2016. The remaining budget will be used for the continuation of the E-journal and the collaboratory.



Kees Mandemakers teaching at the Nijmegen Summer Course

3.4 CEDA_R project

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



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

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

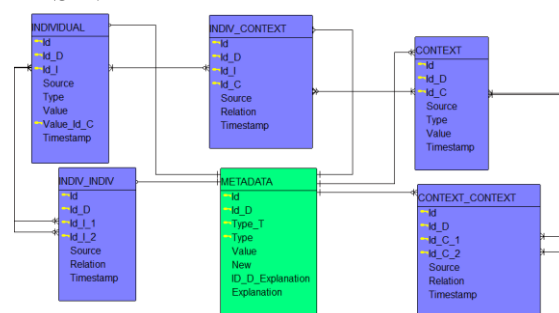
The semantic valorization will be done in two steps: a) a structural improvement of the storage of the data (from spreadsheets into one dataset) and b) harmonizing and disseminating the data by way of RDF-techniques. RDF stands for Resource Description Framework technology. All data from the censuses will be accessible at the level of the cell by way of RDF coding. This will make all kinds of pattern recognition feasible and a much better querying of the dataset (it will be possible to collect for example the number of inhabitants of the village Besoyen for all census years between 1795 and 1947). The team needed to convert about 2,288 spreadsheets in which the Dutch censuses 1830-1947 were transcribed. For this period in the history of the census the results are available at the aggregative level of municipality or province. All spreadsheets were converted into an RDF structure (data cube). The team concentrated on the testing and developing the RDF system and the harmonization of those spreadsheets that contain population and dwellings for each location in the Netherlands. In the spring Albert Meroño-Peñuela, successfully defended his thesis at Free University in Amsterdam. The IISH based PhD, Ashkan Ashkpour, is expected to defend his thesis in 2017. In 2016 he published in *Historical Social Research* the article ‘Source Oriented Harmonization of Aggregate Historical Census Data: A Flexible and Accountable Approach in RDF’

3.5 IDS meso

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

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

CLIO-INFRA were (partly) developed at IISH.

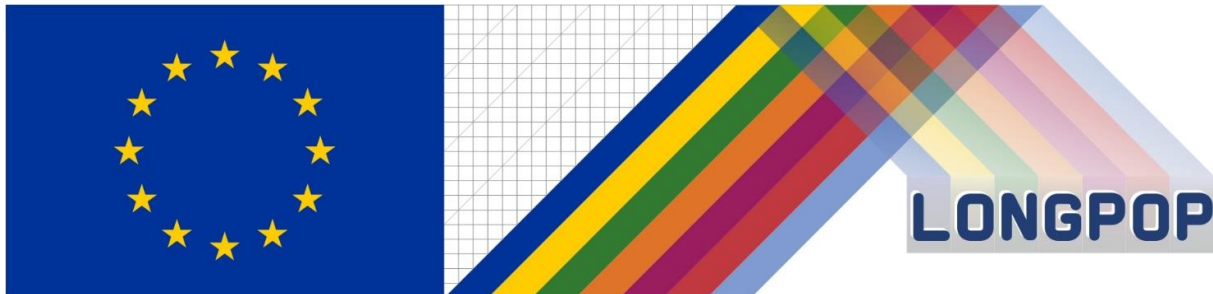


IDS scheme for micro-data

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

The project ended in February 2016 with the development of a demonstrator with a uniform data structure including five different datasets (a.o. unions, strikes, and ships) at the meso level. The project was financed by the KNAW ‘Data advisor’ programme.

3.6 LONGPOP



This project has received funding from the European Union’s Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 676060

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.

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 (ie 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. Francisco Agnuita works on the documentation and extension of IDS related extraction software and Diogo Paiva works on methods of standardizing addresses. Both started at the first of September 2016.

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.

3.7 GIANTS



The GIANTS 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. The study will make use of an extraordinary database, built by the HSN: the reconstructed life courses (occupations, family formation, mobility) of Dutch persons born between 1811 and 1922 (Historical Sample of the Netherlands). The existing HSN database will be enriched with information on heights of a representative sample of about 30,000 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.

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). The database is aimed to be finished at the end of 2018. During 2016 a new module of the data entry programme for entering militia registers was delivered and at the end of the year already 1,000 recruits had been traced (and scanned) in the archives, mainly in Limburg and Noord-Brabant.

4 Staff and Boards HSN

4.1 Staff HSN

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

Four workplaces were available for the work in the archives and the data entry in the office (one WIW workplace and three SWV workplaces). In September Francisco Anguita and Diogo Paiva started 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 and other HSN software. Both are stationed at the IISH Digital Infrastructure department. Kerim Meijer and Richard Zijdemans were involved in the upgrading of HSN data entry software.

At the end of 2016 the total number of HSN-staff, directly and in cooperation with other organizations, was 19 persons (2015: 15 persons). During the year, a total of 20 people worked for the HSN, among whom 8 volunteers, who were engaged in collecting material in archives and in data entry.

Staff in 2016:

F. Anguita, MSc	1,0 fte	Sept. - Dec.
J. Bartman	0,5 fte	Jan. - Dec.
M. Berrier	0,2 fte	Jan. - Dec.
W. Commandeur	0,3 fte	Jan. - Dec.
Th. Dibbets	0,4 fte	Sept. - Dec.
H.J. van Eijden	1,0 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	0,6 fte	Jan. - Dec.
drs. F. Laan	1,0 fte	Jan. - Dec.
prof. dr. C.A. Mandemakers	1,0 fte	Jan. - Dec.
drs. B. Mouwes	0,3 fte	Jan. - Dec.
drs. C. Munnik	0,3 fte	Jan. - Dec.
drs. F. Nijstad	0,1 fte	Jan. - Dec.
D. Paiva, MSc	1,0 fte	Sept. - Dec.
dr. B. Schijf	0,2 fte	Jan. - Dec.
drs. I. de Vries	0,1 fte	Jan. - Juni
M. de Vries	0,8 fte	Mei - Dec.
drs. R. Wasser	1,0 fte	Jan. - Dec.

4.2 Board Foundation HSN

Two members of the board, France Portrait and Vincent Tassenaar, have accepted a new term. In October Hilde Bras resigned from the board. At the end of 2016 the board consisted of the following persons (the year of resignation is between brackets):

Prof. dr. F.W.A. van Poppel, Netherlands Interdisciplinary Demographic Institute (NIDI), Utrecht University, chair (2017)
Prof. dr. A. F. Heerma van Voss, Utrecht University, director Huygens ING, member (2017)
Prof. dr. J. Kok, Radboud University Nijmegen, IISG, member (2020)
Prof. dr. M.H.D. van Leeuwen, Utrecht University, vice chair (2020)
Prof. dr. W.A.F. Maas, Utrecht University, secretary (2018)
Dr. F.R.M. Portrait, VU University Amsterdam, treasurer (2021)
Dr. P.G. Tassenaar, University of Groningen, member (2021)

Official secretary: prof. dr. C.A. Mandemakers.

The board held their meetings on 22 February and 28 September 2016. Main item on the agenda was the progress of the projects.

4.3 Steering Committee HSN

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

The steering committee consists of the members of the HSN board (see foregoing section 4.2) and, on behalf of the IISH, prof. dr. L.A.C.J. Lucassen as head of the research department of the IISH (of which the HSN is a part). The secretary of the steering committee is prof. dr. K.M. Hofmeester. Advisor to the steering committee is prof. dr. C.A. Mandemakers, head of the HSN. The steering committee held their meetings on 22 February and 28 September 2016.

4.4 Scientific Council of Advice

Task of the Advice Council is to provide the board with solicited and unsolicited advice. The board was shocked by the unexpected death of his chair prof. dr. J. Dronkers. In the course of the year there were several informal contacts.

The Scientific Advisory Board consists of:

Dr. P.K. Doorn, head DANS
Prof. dr. 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. H  l  ne V  zina.. There were several mutations in the Board during 2016: prof. Anders Br  ndstr  m left the board after the end of his term at the end of the year. He was replaced as chair by prof. H  l  ne V  zina. Prof. Cameron Campbell entered the board. The composition of the Board was as follows:

Prof. dr. A. Br  ndstr  m, University of Ume  , director Demographic Database Ume  
Prof. dr. C Campbell, University of Science and Technology, Hong Kong
Prof. dr. L. Dillon, D  partement de D  mographie, Universit   de Montr  al
Prof. dr. M. Dribe, Centre for Economic Demography, Lund University
Dr. D. Ramiro-Fari  nas, Instituto de Econom  a, Geograf  a y Demograf  a, Madrid
Prof. dr. H. V  zina, l'Universit   du Qu  bec    Chicoutimi (UQAC)

The Board met in Chicago on 17 November 2016 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

2016

- 370 Gerrit Bloothoof & David Onland, 'Multiple first names in the Netherlands (1760-2014)', *Names* 64 (2016), 1, 3-18, [DOI: 10.1080/00277738.2016.1118860](https://doi.org/10.1080/00277738.2016.1118860) (this paper received the Award for the [Best Article in NAMES 2016](#)).
- 369 Bastian Mönkediek, Jan Kok & Kees Mandemakers, 'The Impact of Family Setting and Local Opportunities on Leaving Home and Migration Destinations of Rural Youths, The Netherlands 1860-1940', *Historical Life Course Studies*, 3 (2016), 43-65, <http://hdl.handle.net/10622/23526343-2015-0005?locatt=view:master>.
- 368 Paul Puschmann, Robyn Donrovich, Per-Olof Grönberg, Graziela Dekeyser & Koen Matthijs, 'Disfavored in Life, Favored in Death? Later-Life Mortality Differences (Ages 30+) between Migrants and Natives in Antwerp, Rotterdam and Stockholm, 1850-1930', *Historical Social Research - Historische Sozialforschung (Köln)* 41 (2016), 4, 257-290.
- 367 Jona Schellekens & Frans van Poppel, 'Early-life Conditions and Adult Mortality Decline in Dutch Cohorts Born 1812-1921', *Population Studies* 70 (2016), 3, 327-343. [DOI: 10.1080/00324728.2016.1223336](https://doi.org/10.1080/00324728.2016.1223336).
- 366 Kees Mandemakers, Bastian Mönkediek & Jan Kok, 'De rol van gezin en bedrijf bij de migratie van Nederlandse boerenkinderen tussen 1850 en 1940', in Paul Puschmann, Richard Paping & Koen Matthijs (eds.), *Familie en levenskansen in het verleden* (Jaarboek Historische Demografie, 2015), Leuven/Den Haag: Acco 2016, pp. 25-56.
- 365 Jan Kok, Hilde Bras & Paul Roterling, 'Courtship and bridal pregnancy in The Netherlands, 1870-1950', *Annales de Demographie Historique* 132 (2016), 2, 165-191, [DOI: 10.3917/adh.132.0165](https://doi.org/10.3917/adh.132.0165).
- 364 Koen Matthijs, Saskia Hin, Jan Kok, & Hideko Matsuo (eds), *The future of historical demography. Upside down and inside out*, Leuven/Den Haag: Acco 2016, 264 pp.
- 363 Kees Mandemakers, 'Historical Population Databases and the Intermediate Data Structure, 1980-2050', in: Koen Matthijs, Saskia Hin, Jan Kok & Hideko Matsuo (eds.), *The future of historical demography. Upside and down and inside out*, Leuven/Den Haag: Acco, 227-232.
- 362 Jan Kok & Kees Mandemakers, 'Life and Death of Singles in Cities, 1850-1940', *Journal of Urban History* 42 (2016), 101-120. First published on 12 September 2015, [DOI: 10.1177/0096144215611098](https://doi.org/10.1177/0096144215611098).
- 361 Hilde Bras & Jan Kok, 'Diverging pathways? Sibling differences in marriage timing in a commercialized rural region of the Netherlands, 1860-1940', in: Fabrice Boudjaaba, Christine Dousset & Sylvie Mouysset (eds.), *Frères et sœurs du Moyen Âge à nos jours* (Population, famille et société, 22 (2016)), Bern: Peter Lang, 189-206.
- 360 Jan Kok, 'Genealogies in historical demography – debates and prospects', *Daedong Munhwa Yeon'gu* 94 (2016), 7-40, [DOI: 10.18219/ddmh..94.201606.7](https://doi.org/10.18219/ddmh..94.201606.7).
- 359 Ashkan Ashkpour, Kees Mandemakers & Onno Boonstra, 'Source Oriented Harmonization of Aggregate Historical Census Data: A Flexible and Accountable Approach in RDF', *Historical Social Research* 41 (2016), 4, 291-321. [DOI: 10.12759/hsr.41.2016.4.291-321](https://doi.org/10.12759/hsr.41.2016.4.291-321).

- 358 Ineke Maas & Marco H.D. van Leeuwen, 'Towards open societies? Trends in male intergenerational class mobility in European countries during industrialization', *American Journal of Sociology* 122 (2016), 838-885, [DOI: 10.1086/689815](https://doi.org/10.1086/689815).
- 357 Antonie Knigge, 'Beyond the Parental Generation: The Influence of Grandfathers and Great-grandfathers on Status Attainment', *Demography* 53 (2016), 4, 1219–1244, [DOI: 10.1007/s13524-016-0486-6](https://doi.org/10.1007/s13524-016-0486-6).
- 356 Julia Efremova, *Mining Social Structures form Genealogical Data*, Dissertation Technische Universiteit Eindhoven (2016), 170 pp, https://pure.tue.nl/ws/files/20451661/20160420_Efremova.pdf.

For the publications in foregoing years see the HSN website: www.iisg.nl/hsn/products/publications.

Appendix B Lectures, symposia and other promotional activities

2016

- 648 Kees Mandemakers, 'International Institute of Social History and Structured Datasets: HSN, LINKS, CLIO-INFRA, Labour Hub and CLARIAH', *First general LONGPOP Meeting*, IISH, Amsterdam, 1 December 2016.
- 647 Antonie Knigge, 'Sources of Sibling (Dis)similarity: Total Family Impact on Status Variation in the Netherlands in the Nineteenth Century', *16th Annual Meeting of the European Academy of Sociology*, Paris, France, 22 October 2016 (invited).
- 646/
645 *41th Annual Meeting of the Social Science History Association*, Chicago, USA, 17-20 November 2016, with the following contributions:
- Ingrid van Dijk & Kees Mandemakers, 'Like Mother, like Daughter. The Intergenerational Transmission of Mortality Clustering in Zeeland 1833-1900', session 'Decease and Mortality'.
 - Kees Mandemakers & Gerrit Bloothoof, 'Linking and Geographical Space. The Consequences of Linking Decisions based on Geographical Distance', session 'Metrics to Assess the Quality of Record Linking'.
- 644 Kees Mandemakers, 'HSN and Modern Data. Connections between HSN and Modern Data and Health (Genetics)', *EHPS Workshop on Modern Data and Health*, Copenhagen, Denmark, 17-18 October 2016.
- 643 Evelien Walhout, 'Were Dutch girls missing? Exploring evidence on female infanticide in the Netherlands, 1625-1910', *International workshop 'Demography and gender'*, Paris, France, 11 October 2016.
- 642 Gerrit Bloothoof, '[A Historic Person Registration](#)' (poster), *Digital Communities Event*, Utrecht, 7 October 2016.
- 641 Kees Mandemakers, 'Historische Steekproef Nederlandse bevolking', *Seminar Nederlands-Indische bronnen*, Nationaal Archief, the Hague, 6 October 2016.
- 640 Antonie Knigge, 'Status Differences between Siblings. The Effect of Birth Order Reconsidered', *ECSR Conference*, Oxford, UK, 22-24 September 2016 (paper).
- 639/
637 *Biennial Meeting of the European Society of Historical Demography (ESHG)*, KU Leuven, Belgium, 21-24 September 2016, with the following contributions:
- Ingrid van Dijk & Kees Mandemakers, 'Like mother, like daughter: intergenerational transmission of mortality clustering in Zeeland 1833-1900'.
 - Rick Mourits, 'Populations where men outlived women: The gender gap in later-life mortality and longevity in the 1812-1862 Zeeland cohort'.
 - Adriana Santacroce, Ewout Depauw & Mariska Meints, 'Trajectories after widowhood. Life-course pathways of widows and widowers in the Netherlands, 1873-1934'.
- 636 Gerrit Bloothoof & Kees Mandemakers, '[The Zeeland Challenge](#)', document for benchmark, presented at workshop *Data Linkage: Techniques, Challenges and Applications* at Isaac Newton Institute for Mathematics, Cambridge, UK, 15 September 2016.

- 635 Gerrit Bloothoof, Jelte van Boheemen & Marijn Schraagen, '[Historical life cycle reconstruction by indexing](#)', workshop *Data Linkage: Techniques, Challenges and Applications*, at Isaac Newton Institute for Mathematics, Cambridge, UK, 12 September 2016, <http://www.newton.ac.uk/seminar/20160912113012001>.
- 634/
633 *Summer Meeting of RC28*, Bern, Switzerland, 29-31 August 2016, with the following contributions:
- Nigel Kragten, Ineke Maas, Marco H.D. van Leeuwen & Harry B.G. Ganzeboom, 'Long-term trends in intergenerational social mobility in the Netherlands, birth cohorts 1760-1980'.
 - Antonie Knigge, 'Status Differences between Siblings. The Effect of Birth Order Reconsidered' (paper).
- 632 Kees Mandemakers, 'Introduction into HISCO and its background', *HISCO workshop, Old and New Historical Population Databases: Queries and Challenges in the Process of Coding Occupations Using HISCO*, Babeş-Bolyai University, Centre for Population Studies, Cluj-Napoca, Romania, 18 June 2016.
- 631 Ingrid van Dijk, 'Death and the family. Early-life mortality clustering in high mortality families', *ANARCHIE Seminar Summer term 2016*, Max Planck Institute for Social Anthropology, Halle (Saale), Germany, 13 June 2016.
- 630 Antonie Knigge, 'Status Differences between Siblings. The Effect of Birth Order Reconsidered', *Dag van de Sociologie*, Tilburg, 9 June 2016 (paper).
- 629 Tim Riswick, 'Differences in Mortality Chances of Brothers and Sisters in the Netherlands, 1860-1910', paper at the *N.W. Posthumus Conference*, Wageningen University, 27 May 2016.
- 628 Paul Puschmann, Robyn Donrovich & Koen Matthijs, 'Salmon Bias or Red Herring? Comparing Adult Mortality Risks (ages 30+) between Stayers, Returnees and Movers in Rotterdam, the Netherlands, 1850-1930', *Economic, Social and Demographic History Seminar*, Radboud University Nijmegen, 13 April 2016.
- 627 Gerrit Bloothoof, '[De grenzen van verwantschapsreconstructie](#)', *Tresoar Grutte Pier Famyljelezingen*, Kimsverd, 3 April 2016.
- 626 Tim Riswick, Presentation of the PhD-project 'Between Rivalry and Support: Differences in the Mortality Chances of Brothers and Sisters in Taiwan (1906-1940) and the Netherlands (1860-1910)', *Program for Historical Demography*, Academia Sinica, Taiwan, 3 March 2016.
- 625/
619 *11th European Social Science History Conference*, Valencia, Spain, 30 March-2 April 2016, with the following contributions:
- Ashkan Ashkpour, Albert Meroño-Peñuela & Kees Mandemakers, 'CEDAR: Harmonizing Aggregate Historical Dutch Census Data: A Flexible Approach'.
 - Ingrid van Dijk, 'Death and the Family'.
 - Paul Puschmann, Robyn Donrovich & Koen Matthijs, 'Health advantage or statistical artifact? A test of the salmon bias hypothesis for domestic migrants in the city of Rotterdam, the Netherlands, 1850-1930'.
 - Tim Riswick, 'Differences in Mortality Chances of Brothers and Sisters in the Netherlands, 1860-1910', paper in session 'Siblings and Life Transitions I'.
 - XingChen ChiaChi Lin & Tim Riswick, 'The Influence of Family Composition on Infant Mortality in Single Parent Families in Taiwan (1906-1945) and the Netherlands (1870-1920)' (paper).

- Ineke Maas & Marco H.D. van Leeuwen, 'Partner choice in the Netherlands 1812-1914: the importance of ascribed and achieved status'.
 - Rick Mourits, 'Later-life mortality and the environment: Geographic clustering of mortality in 1812-1862 cohorts'.
- 618 Kees Mandemakers, 'HSN en LINKS', *Nederlandse Genealogische Vereniging (NGV) & Genealogysk Wurkferbân Fryske Akademy*, Historisch Centrum Leeuwarden, 12 March 2016.
- 617/
615 *EHPS-Net workshop 'Databases'*, Umeå University Centre for Demographic and Ageing Research (CEDAR), Sweden, 16-18 February 2016, with the following contributions:
- Kees Mandemakers & Gerrit Bloothoof, 'Record linkage with the LINKS project: From Marriage Certificates to Pedigrees'.
 - Kees Mandemakers, 'Historical Longitudinal Databases and Quality Standards & Standardizing by HSN'.
 - Kees Mandemakers, 'Intermediate Data Structure (and linkage)'.
- 614 Gerrit Bloothoof, '[De wet van Zipf voor eigennamen](#)', *Grote Taaldag*, Utrecht, 6 February 2016.
- 613 Gerrit Bloothoof, '[Who's who in the Netherlands](#)', pecha kucha presentation at the *Language Science Day*, UiL-OTS, Utrecht, 4 February 2016.
- 612 Rick Mourits, 'Populations where men outlived women: A study of the gender gap in survival after age 50 for the 1812-1862 cohort from Zeeland', *ESDG-seminar*, Radboud University Nijmegen, 20 January 2016.

For the presentations in foregoing years see the HSN website: <https://socialhistory.org/en/hsn/hsn-presentations>.

Appendix C Reports and Working papers

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

2016

- 40 Jelte van Boheemen, *Assembling the pages. A sorting-based approach to historical record linkage*, Bachelor's thesis Artificial Intelligence, Radboud Universiteit Utrecht.
- 39 Max Bosschaart & Tim Lelieveld, *Historische Steekproef Nederlands-Indië. Een haalbaarheidsonderzoek*, Radboud Universiteit Nijmegen.
- 38 Tim Thijs & Marco van Tiggelen, *Verkennd onderzoek naar kohieren van hoofdelijke omslag ten dienste van de Historische Steekproef Nederland*, Radboud Universiteit Nijmegen.
- 37 Bram van den Hout, *Sociale en Ruimtelijke Mobiliteit in Nederlands-Indië. Stageverslag, Internationaal Instituut voor Sociale Geschiedenis, including Handleiding Pilot-Project HSN Nederlands-Indië.*

Appendix D Releases

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

2016 - HSN

- 61 Kees Mandemakers and Cor Munnik, *Historical Sample of the Netherlands. Project Genes, Germs and Resources. Dataset LongLives*. Release 2016.01.
- 60 Historical Sample of the Netherlands (HSN). Dataset *HSN Survival Dates*, release 2016.01 (n=85,334).

For a complete overview of releases see the HSN website, <https://socialhistory.org/en/hsn/hsn-releases>.

2016 - LINKS

- 27 Kees Mandemakers and Fons Laan, *LINKS Zeeland Linked Dataset (Marriages, Births and Deaths), Project Genes, Germs and Resources, Province of Zeeland*, Release 2016_05, beta version.
- 26 Kees Mandemakers and Fons Laan, *LINKS Zeeland Linked Dataset (Marriages, Births and Deaths), Project Genes, Germs and Resources, Province of Zeeland*, Release 2016_04, including IDS format, beta version.
- 25 Kees Mandemakers *LINKS Netherlands Linked Dataset Marriages, Provinces Groningen, Drenthe, Overijssel, Gelderland and Noord-Holland, 1812-1922*, release 2016_01, beta version.
- 24 Kees Mandemakers, *LINKS Zeeland Linked Dataset (Marriages, Births and Deaths), Project Genes, Germs and Resources, Province of Zeeland*, Release 2016_03, beta version.
- 23 Kees Mandemakers, *LINKS Zeeland Cleaned Dataset (Marriages, Births and Deaths)*, Release 2016_01.
- 22 Kees Mandemakers, *LINKS Zeeland Linked Dataset (Marriages, Births and Deaths), Project Genes, Germs and Resources, Province of Zeeland*, Groningen/Drenthe and Limburg Release 2016_02, beta version.
- 21 Kees Mandemakers, *WieWasWie Marriage Certificates Haarlem Railway Personnel, 1903*, release 2016_01.
- 20 Kees Mandemakers *LINKS Zeeland Linked Dataset (Marriages, Births and Deaths), Project Genes, Germs and Resources, Province of Zeeland*, Groningen/Drenthe and Limburg Release 2016_01, beta version.

For the LINKS releases in foregoing years see the HSN website, <https://socialhistory.org/en/hsn/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.

Acronym	Project title
MUT/ASG	Migration in the province of Utrecht
OVF	Reduced fecundity because of maternal high-risk conceptions
RDN	Regional differences in demographic behaviour, the Netherlands, 1900-1960
AKON	General index of death certificates in the Netherlands
TTA	Textile industry workers in Twente
MFZ	Geographic and Social Mobility of Female Domestic Servants in Zeeland, 1850-1950
DUM	Germans in Utrecht: a temporary minority in the 19th century
RCM	Religious differences in infant and childhood mortality, The Hague, 1860-1920
DVI	Settlement determinants for immigrants and their descendants in the Netherlands, 1853-1960
GBW	Family formation and living strategies in the western parts of the Netherlands 1830-1940
ESM	Early-life conditions, social mobility and longevity
RAM	Living Strategies of Born Rotterdammers
VBA	On the move in Amsterdam. Mobility of the Amsterdam poor 1900-1940
LCC	Life Courses in Context (NWO Large investment)
MNI	European migration to the Dutch East Indies
HVL	Marriage certificates Pupils of Dutch Higher Secondary Education
LINKS	LINKing System for historical family reconstruction
LMP	Long Term Mortality Effects of Potato Crisis
JDJ	Jewish Dutch or Dutch Jews?
LHL	Linking Historical Lives (Linked Lives)
MOSAIC	MOSAIC - the Netherlands
EHPS-Net	European Historical Population Samples Network
CEDAR	Census Data Open Linked
HLZ	HSN LINKS Zeeland (CLARIAH seed)
GGR	Genes, Germs and Resources
LINKS val	LINKing System for historical family reconstruction
IDS meso	IDS meso: Intermediate Data Structure for organizations
LONGPOP	Methodologies and Data Mining Big Data based on Longitudinal Population and Epidemiological Registers
GNT	GiaNTs of the modern world. A new history of heights and health in The Netherlands, 1811-1940'
CCH	CLARIAH Curing the Historical Sample of the Netherlands