

# The Norwegian historic population register and migration

**Note no.**

**DART/10/2015**

**Authors**

**Lars Holden, Svetlana Boudko**

**Date**

**1. jun. 2015**

## Authors

Lars Holden, Svetlana Boudko

## Norsk Regnesentral

Norsk Regnesentral (Norwegian Computing Center, NR) is a private, independent, non-profit foundation established in 1952. NR carries out contract research and development projects in information and communication technology and applied statistical-mathematical modelling. The clients include a broad range of industrial, commercial and public service organisations in the national as well as the international market. Our scientific and technical capabilities are further developed in co-operation with The Research Council of Norway and key customers. The results of our projects may take the form of reports, software, prototypes, and short courses. A proof of the confidence and appreciation our clients have in us is given by the fact that most of our new contracts are signed with previous customers.

**Title** **The Norwegian historic population register and migration**

**Authors** **Lars Holden, Svetlana Boudko**

Date 1. jun. 2015

Year 2015

Publication number DART/10/2015

### **Abstract**

This paper describes the building of a Norwegian historical population register that is under construction. It is based on linking together person entries from a large number of sources by linking algorithms and crowd sourcing. There are 9.7 million residents of Norway in the period 1735 - 1964 and 37.5 million events in the most relevant sources. We want to link together as many events as possible for the same persons and families but only include links that have a high probability to be correct.

As the first open national population register, we believe that it will be particular important for migration studies. It is much easier to follow both national and international migration at a national level. Quite many of the international migration sources are national, and place of birth is specified by nationality for persons born abroad.

Keywords Population register, migration, census

Target group Broad

Availability Open

Project number 320517

Research field Smart information systems

Number of pages 17

© Copyright Norsk Regnesentral



# Table of Content

1	Introduction .....	7
2	The three periods .....	8
3	Input data .....	8
4	Linking .....	10
5	IDs .....	12
6	References and the register .....	13
7	Migration .....	14
8	Research and the use of HPR.....	15
9	The data base structure in HPR. ....	16
	Acknowledgment .....	16
	References .....	16

## List of figures

Figure 1. Part of the person page for Christian Michelsen

Figure 2. Illustration of the result of a search for Roald Amundsen in HPR.



# 1 Introduction

Several regions are building and extending historical longitudinal population registers for research, statistical and historical purposes. This paper describes the building of a Norwegian historical population register (HPR) and describes how it can be used in order to keep track of migration. Most of it will be open at internet. With 9.7 million residents of Norway in the period 1735 - 1964 and 37.5 million events in censuses, church books and other sources (Thorvaldsen, 2011), the register will be much larger than the closed Icelandic register (deCode) and the regional registers in Sweden (Demographical data base in Umeå, Rotemannarkivet for Stockholm) and the Netherlands (Historical samples of the Netherlands). Unlike other registers, HPR will build on automatic linking of sources and crowd sourcing at internet.

It is of interest to follow the ascendants and descendants one or two generations for persons that have lived in Norway. This may show family relationship in Norway and the family may migrate several times over generations. It is also of genealogical interest. But it is important that place of birth and death is specified by country in order to avoid any bias in the statistics for Norwegians.

Our ambition is to link as many person occurrences from the 37.5 million events as possible. Two person occurrences should only be linked as long as the probability that they apply the same person is above a certain level. This is a consequence of using crowd sourcing and the ambition of finding all links that are possible to document. This is in contrast to others (e.g. NAPP) where statistical representativity is given more weight at the cost of not using address and family relationship in the linking. When this information is not used, some of the easiest links are not found. When we make statistical estimates for the population e.g. number of birth per women, it is necessary to correct for this bias. We believe our approach will give better estimates when we have obtained a sufficient linkage ratio, since we have a larger data set.

The first version is planned to be available summer 2015 followed by a period of linking data about the same persons for several years. The register will, of course, never be completed. Our aim is to link 80% of the population (i.e. found in minimum two sources) and that the register will be in frequent use, gradually becomes more complete and with improved quality instead of degrading due to wrong links. The need for the longitudinal HPR extending over more than two centuries has been documented in a series of research papers in a wide range of scholarly fields: medicine (Hovig, 2010), demography, history, and the social sciences including economics. Research topics range from the tracing of genetic diseases via name studies to ethnic differentials with respect to cohabitation. This is best documented in the publication list at the Demographical data base in Umeå.

There are many possible extensions of the HPR data base. One of the most important is to include overview over all houses and farms. This will show the development of farms over longer periods which is of interest for studies of the economic development of a region. As a side effect, this will also improve the quality of the HPR data base and make it easier to follow some families for generations. We believe that the more sources and different uses of the data base we are able to include and handle, the more users we are able to attract. This gives a data

base of higher quality and of increased value. We try to make the data base as open and transparent as possible focusing on documentation and references to high quality sources. We want to be important for the scientist, for the experienced genealogist and for the general public that want to find information similar to the use of Wikipedia. We want to connect to, to be a register for and to encourage the use of all the sources about Norwegian historical persons by proper use of references, not to replace the other sources.

The paper (Holden et al. 2012) describes a previous plan for the HPR data base. Now we use a relational data base built on the same principles as the data in the Digital Archive of the National Archives. This gives us direct access to all the updated sources in the Digital Archive instead of copying one and one source into a wiki-database. This gives a more effective data base administration, but less flexibility in building up new uses of the same data.

## 2 The three periods

The register will be divided into three parts due to legal reasons. The Law of Statistics requires that the censuses are closed for a period of 100 years. Information from the municipal censuses may be disclosed after 60 years, but little of this material has been digitized. The Law of Publicity similarly protects some of the information in the church records for 60 years and sensitive issues such as adoptions for 100 years. Privacy regulations (With, 2005) restrict the possibility to link information about people that are still living. Therefore, the register is divided into three parts. The oldest and open part includes only dead persons and censuses before and including 1910. This limit is moved when we are allowed to include more register data in an open register. This part may include some open data of living persons, but it is only possible to link data for dead persons. The second period is closed and continues until 1964 where the National Population Register starts. The National Population Register is the responsibility of the Norwegian Tax Administration and will not be described here.

The register will not contain sensitive information. As an example, we plan to include the death register of Norway from 1928-1960, but leaving out the cause of death. Scientists that need sensitive information may combine HPR with other data bases including this sensitive information. Researchers that need data from the closed periods will get access to this combined with data from other sources, possibly with more sensitive data, in anonymous form from Statistics Norway provided the project is accepted after a scientific and legal evaluation. The National Archive is responsible for the two first periods while Norsk Regnesentral develops the data base for HPR.

## 3 Input data

The register is based on as many full sources as possible: censuses, church books and other sources. The sources in HPR may be divided into four different types:



1. HPR builds on API access to the Digital Archive and the private depository Digital Inn of the National Archive that includes transcribed national and municipal censuses, church books, emigration and immigration records, probate register, prison register, tax records etc. This archive is increased daily and all the data is automatically available in HPR.
2. We also want to include data that is not available in the Digital Archive and store this in a HPR data base with proper references and links to other data bases. We plan to include 2.6 mill. entries from graveyards registered by DIS-Norge.
3. Contributors to HPR are encouraged to register certain types of data in HPR. Currently this includes death, marriages, anniversaries, and events of public interest from newspapers. We have used automatic recognition to register about 0.5 mill. events from Aftenposten, the largest newspaper in Norway, in the period 1950-1999. This may be extended to other newspapers. Contributors to HPR can register new events from newspapers. Norwegian newspapers have a history back to 1763, but this is more relevant to document events from the last 100 years that are not documented in other open sources. This makes it possible to document persons living in the last period and include this in the open data base. It is also possible to register persons directly in an input table provided: (i) the persons are not registered in any of the sources listed above or the sources we expect to include in the near future and (ii) these persons are identified by a trusted contributor like Busetnadssoge (Kjelland). These persons are listed in a separate register sorted by the contributor, source (e.g. Lesja bygdebok) and name of the registered person.
4. For each person registered in the sources described above, it is possible to make references to any other source. This may be church books that are not transcribed, tax records, internet sites like [www.lokalhistoriewiki.no](http://www.lokalhistoriewiki.no) with local history and [www.eidsvollsmenn.no/](http://www.eidsvollsmenn.no/) documenting the ancestors of the persons participating at the first national parliament, family records, local historical books, scientific articles and Wikipedia. The contributors may register any kind of reference. All references are registered alphabetically in a register making it easy to find other person from the same source. It is important that URLs, links and IDs are permanent, otherwise it is of limited use and should not be included.

Each entry of a person name in a source, denoted person source entry, PSE, is a separate unit in HPR. A PSE is a piece in a puzzle that we know should be connected somewhere. Each PSE is only registered once in the data base in order to avoid any doublets. Whenever possible, each PSE has a link to the transcribed source in the Digital Archive or similar data base and ideally also a link to an image of the source. This makes it easy to check the source. If a name or date varies between the sources for the same person, it is easy to check if this is an error in the source or in the transcription. However, the spelling of names may change during the life time of the person. Since we allow any kind of references, a reference is not sufficient to be registered as a PSE. This would lead to doublets of PSE and poor data quality. We try to include as many high quality sources in the first three categories as possible. We have started an ambitious

plan to transcribe all the church books and many other sources and include these in the Digital Archive the coming years. How the different sources are included in HPR, are prioritized in the order they are listed above.

## 4 Linking

The main challenge in HPR is to link all the PSEs regarding the same person. Most persons in HPR will have many PSEs that are independent of each other in the data base until they are linked together. Family relations from the source are maintained. Hence, if we are able to link the baptism of the same person as a child and as a parent, we are able to connect three generations. Family relationship in censuses is slightly more complicated. We use a program from The Minnesota Population Center in Minneapolis that identifies the family relations from the role in the census (Sobek, Kennedy, 2011). But the errors in the interpretation are slightly more frequent than our expectations and we will try to improve this. Since we have internet links to the transcribed source, it is easy to check manually.

We try to link as many PSEs as possible mechanically using a search and matching algorithm. The Norwegian Historical Data Centre at University of Tromsø has developed linking programs over many years. By June 2015 they have found 2.4 million links between and within censuses and church books. The linking program compares PSEs based on names, names of partners, age/year of birth, household address and municipality. Names are compared using Jaro-Winkler making it possible to identify the same name with slightly different spelling. This mechanical linking is performed in a separate Oracle database and the links will be imported into the HPR data base.

When we link two PSEs, i.e. state that this is the same person, we establish a person unit, PU. We also establish a PU if we add any other information to a PSE, a name, date or place of birth or death, a reference or a text/biography. HPR has an internet page for each PSE and for each person/PU. The PSE page only lists the most important information from the source and has links to the transcribed source in the Digital Archive/HPR PSE data base. The person page shows the names, date and place of birth and death, parents and siblings, partners and children, list of linked PSE, a reference or a text/biography. The text may be a biography and/or explain the linking. The list of PSE will ideally include the main information from baptism, confirmation, marriage, baptism of own children and as godparent and death. It may also include other events as tax records, prisons, emigration, immigration, events from newspapers, etc. See Figure 1 illustrating a part of a person page.

Busetnadssoge (Kjelland et al) is a program that is used for making population registers for municipalities. We will also import links from this and similar data bases. The critical part is that the links are identified by the correct IDs for the PSEs in addition to the general requirement that the links have high quality. Such data bases may include persons without any PSE since they are identified by sources that are not transcribed or

not given PSE from the included sources. These persons may be included in the separate person input table as part of HBR PSE data base.

## Peter Christian Hersleb Kjerschow Michelsen

Mann  
 Født: 15.03.1857, Bergen  
 Død: 29.06.1925, Fana, Bergen

[Rediger](#)

[Ok if netet](#)

### Partner og barn

Partner: Johanne Benedikte Boalth Michelsen  
 Barn: Benny Wallendahl

NR	Dato	Kilde	PFID	Rolle	Navn	Fødselsdato	Fødested	Bosted	Fam. St.	Sivst	Yrke	Far/brudgom	Mor/brud	Partner	Barn/andre
1		FOLK, Bergen, 1865	pf01038249015898	folketelt	Peder Christian Michelsen	1857	Bergen	Bergen: 14-29	son	u					
2		FOLK, Fana, 1910	pf01036693004940		Christian Michelsen	1857-03-15	Bergen	Fane: Gammelhaug	hf	e	skipsreder				
3		MINI, Bergen, 1867-1882	pv00000001121068	brudgom	Peter Christian Hersleb K Michelsen	15.03.1857					Uk overrettsagfører			Johanne Benedikte Boalth Wallendahl	
4		FOLK, Bergen, 1885	pf01053295001211		Christian Michelsen	1857	Bergen	Bergen: Alleegade 59			Overrettsagfører				
5		FOLK, Bergen, 1875	pf01052295014857		Christian Michelsen	1857	Bergen	Bergen: 14-29	Søn	ug	Student				
6			pd00000007656353	far	Peder Christian Hersleb K Michelsen	1857					Overrettsagfører			Johanne Benedicte Boalth Wallendahl	

### Merknader

Christian Michelsen var skipsreder og norsk statsminister i perioden 1905-1907.

[Lagde endringer i merknader](#)

### Referanser

- [Wikipedia, Michelsen, Christian](#)
- [Norsk biografisk leksikon, Michelsen, Christian](#)

Figure 1. Part of the person page for Christian Michelsen that was prime minister in Norway. The third column "kilde" describes each of the sources. The sources are from censuses 1865 and 1910, marriage in 1881, censuses 1885 and 1875 and father in baptism in 1887.

In HPR it will be possible with manual linking by contributors as crowd sourcing at internet. It is necessary to be registered in order to be a contributor. We will start with a small group of contributors and will gradually increase the number of contributors when we are sure that we are able maintain a high quality on the contributions. All contributions are registered by the name of the contributor and the time of the contribution. It is possible to identify all contributions by each contributor. This makes it possible to encourage persons with many contributions, but also to shut out contributors that violate or have contributions of low quality. We also need to get experience with a large number users and contributors in our HPR data program. The open part of HPR will be available on the internet, but we will not encourage use before we have more experience and increased the number of links. We expect a really large

number of users and potential contributors when we gain some experience, as genealogical research is a major activity on the internet. It is of great importance to succeed in attracting a large number of contributors to HPR.

HPR uses the same search program and data as is used in the Digital Archive. But in HPR this also includes hits in HPR PSE data base. The result of a search is a list of data with the vital information from the PSE/PUs that match the search. See Figure 2. From each search result it is possible to put a PSE/PU into a basket. This is similar to a basket/chart in an internet shop. In an internet shop the basket/chart contains the items that you plan to buy. In HPR the basket contains the PSE/PU that the contributor wants to link, i.e. state that this is information about the same person. From the person page, it is also possible to put the person into the basket or perform a search for similar persons.

In addition to linking the PSE/PUs in the basket, we may instead register the link as a family relation, a linkage candidate or a negative linkage. Register as a family relation may be based on a manual interpretation from a census or when we are certain about the relationship but are not able to document it in a source available in HPR. Linkage candidates are PSEs that may be the same person, but the documentation is regarded as insufficient. Negative linkage is PSEs that have similar attributes, but are not the same person. Negative linkage gives information to search routines and other contributors, that these PSEs should not be linked. All information regarding a PU is stored at the person page.

All links are given a score between -1 and 10. 10 means that all vital attributes are the same and 0 that no of the attributes are the same. So far, we are not able to utilize the frequency of the different names. -1 means that there is a conflict f.ex. if we link two different PSEs present in the same census. We know that there may be errors in sources f.ex. that the same event is registered twice. Hence, links with score -1 may be correct.

We apply star linking of the PSE in a PU. This means that all links are recalculated as a link between each PSE and the primary PSE in the PU. Scores are calculated as the lowest score from the sequence of links between PSE that is replaced by the direct link between a PSE and the primary PSE.

## 5 IDs

All PSEs in HPR is given a unique ID. The Digital Archive gives all PSEs in their archive a unique, non-informative 16 digit ID. HPR gives PSEs from newspapers a unique, informative 20 digit ID. When we link PSEs into a PU, one of the PSE is denoted primary PSE and the PU gets a unique ID generated from the primary PSE. The persons making the links determine which PSE that is the primary PSE in a three structure. Hence, the contributor decides which of the primary PSEs in the two linked PSE/PUs that becomes the primary PSE in the new PU. Search for IDs to PUs that no

longer are active (f.ex. due to a merge between two PUs) will be redirected to the active PU that includes the primary PSE to the non-active PU.

## Søk

PFID	Navn	Kjønn	Rolle	Fødselsår	Fødested	Bosted	Stilling og stand	Hendelsesdato	Kilde	Kildekategori
pf01053257046634	Roald Amundsen	m		1872	Borge	Kristiania: Uranienborgveien 9	Skolediscipel	1885-12-31	FOLK, Oslo, 1885-1885	ft
pf01036371002571	Roald Amundsen	m		1872	Borge Smålenene	Nesodden: Uranienborg	hf ug kaptein, polarforsker	1910-12-01	FOLK, 0216 0217, 1910-1910	ft
pf01037045153788	Roald Amundsen	m		1872	Borge i Smålenene	Kristiania kjøpstad: Professor Dahls Gade 45	ug Sømand	1900-12-03	FOLK, Oslo, 1900-1900	ft
pf01052055035483	Roald Amundsen	m		1872	Borge Sogn	Kristiania: Uranienborgveien 3		1875-12-31	FOLK, Oslo, 1875-1875	ft
pe00000000554744	Roald Amundsen	m		1872		Oppegård	u Forsker	1922-03-17	EMIP, Oslo, 1867-1930	em
pe00000000561184	Roald Engebret Amundsen	m		1872		Oppegård	G Skibsf.	1924-09-27	EMIP, Oslo, 1867-1930	em

Figure 2. Illustration of the result of a search for Roald Amundsen in HPR. The hits are from top: municipal census from Kristiania from 1885, national census in 1910, national census in 1900, municipal census from Kristiania in 1875 and then two registrations from the emigration register from 1922 and 1924 respectively.

## 6 References and the register

Use of references is a method to utilize any other kind of information that is available at internet, in books, archives or any other places but not included in the Digital Archive and HPR PSE data base. This opens for very important and relevant information from a large number of sources. It may also give information of poor quality since we don't want to make strict regulations but are open for a diversity of sources. We don't include private gedcom files with family records into HPR. This would give us a large number of links of high quality, but also many doublet and data of poor quality. However, we encourage references to family records at internet or in books. Here, we may find links and family relations that we (so far) have not included in HPR. Since we want to document everything in HPR with high quality sources, we need references to all other kind of information also where the quality and completeness of the data is more uncertain. These other sources may have references to primary sources that we don't have access to in HPR and correct links that we are not able to document in HPR. We expect the users of HPR are able to handle this multitude of sources of varying quality.

All references are listed alphabetically in the register. This makes it easy to find references to the same source. This may be different persons in the same family record, passenger at the same ship or known persons with a biography in Wikipedia. Then we are also able to identify if two different PSE or PU in HPR both refer to the same person in f.ex. an American census. Hence, we are able to avoid doublets references to the same source almost in the same manner we avoid doublets of the PSE in HPR. This makes it possible to increase the number of links and reduce the number of doublet PUs in HPR. As an illustration we give this list of possible references:

Geni.com, Astri Brun (Genealogical database with 89 mill. profiles world wide)  
Eidsvollsmennenes etterkommere Jacob All eftk.: Nils Hofman Aall  
Heimen, 2011, b48. Eide, Thorvaldsen: Andreas A. Svalestuen  
Lokalhistoriewiki: Nansen, Fridtjof  
Vindern historielag, Medl. blad., 2015/1, Sars: Fredrik Barth  
Wikipedia: Nansen, Fridtjof

A challenge here is to find a unique way to write the different references. If the same reference is written differently, we may get doublet references in HPR. Hence, the list above is important also as examples on how to write references. We will automatically generate references between the biographies in a wiki for local history (Lokalhistoriewiki.no) which currently have more than 10.000 biographies. We also encourage users to write biographies here.

## 7 Migration

There are many regional population registers covering a municipality. It is possible to obtain a high linkage ratio in remote places with limited mobility like Rendalen. At places with more migration like the mining city Røros and in larger cities like Oslo, it is not possible to obtain the same linkage ratio. It is easier to follow events for persons through their life and families for generations if they live at the same farm or municipality, than persons and families that migrate. It is easier to follow the persons that marry within the local community than persons that find their partner far away. Persons that don't have a permanent address are of course even more difficult. Hence, all population registers will be biased by poorer coverage for people that migrate. Continuous or stepwise migration is more difficult to follow than a single migration.

Migration is a major motivation for a national population register. This makes it possible, but still difficult, to follow persons and families when they move. Place of birth is specified in most sources. For migration inside Norway we have PSE for own baptism and death and possible baptism as parent and censuses in between that is not linked before we are able to make the connection. In principle, migration in and out of Norway should also be registered several times. Also this gives us PSEs that may be connected to the other non-linked PSEs. When we have made all the easy links in HPR, it will be easier also to find the more complicated ones. Linking candidates may be used when we have found a likely link but when we cannot be sure.

There are a large number of sources for migration from Norway, see Eide et al 2011. The police in the emigration harbor made a register on all emigrants from 1865 which is included in the Digital Archive. Only from Oslo this register has about 470.000 names on emigrants. The priest made a register over migration in/out from the parish from 1813. The Norwegian censuses from 1910 and 1920 have information about migration. About 20.000 migrated back to Norway from America in the period 1910 to 1920. The American censuses have information on place of birth for the person and in the period 1900-1930 also for the parents. As an example, 114.000 persons with Norwegian nationality are registered in the American 1870 census. Also these sources are part of/planned included into the Digital Archive. There are also several other sources like list from ships. None of these sources are complete, but together they make it possible to link a large number of persons giving a better understanding of the major migration from Norway in the period 1825-1940. There will be a close cooperation between the development of HPR and the international cooperation in the North Atlantic population project, NAAP (Thorvaldsen, 2011). Hopefully, we will be able to read in the links made in the NAPP-project between Norwegian and American census.

The strategy is clear. We will try to include as many primary sources into the Digital Archive or directly as part of HPR. This includes sources with emigrant like passenger lists, emigrant records, Norwegians in other national censuses etc. either as transcript of the original source or as a newspaper list. All these sources give a large number of PSEs that is a piece in the large HPR puzzle. Each PSE has a unique position in the large puzzle.

Sources that are not included as described above, may be used as references. As described above, we will be able to avoid doublet references since these will be identified in the register that is sorted alphabetically. This may f.ex. be links to regional population registers in Sweden (f.ex. Rotemannsarkivet) or Denmark (f.ex. Odensedatabasen). When HPR has gained experience and obtained many contributors, the register will give a good overview over the multitude of possible sources. Hopefully, owners and users of other sources and internet sites want their favorite sources to be included in the register. HPR want to encourage use of other sources by proper use of references and URLs with links if possible, not replacing these other sources. We will make HPR as transparent as possible.

Typical references on migration:

Danmarks Adels Årbok, Bugge: Christian Bugge

Ellis Island, Kristianiafjord 1913: Henrik Hansen (Immigration to US)

Odensedatabasen: Peder Hansen ID 190001, d. 1743

Rotemannarkivet: Ole Andreas Edvard Olsen b. 1856-09-26

WeRelate.com: Karelius Olsen (1) (genealogical database with 2,6 mill, Minnesota, US)

## 8 Research and the use of HPR

It is important that HPR is able to attract the scientists and other experts (employees at museums, genealogists...) and make these contribute to HPR. Then we need to make HPR valuable for them. The main purpose of HPR is research. This is where we have obtained the major funding. Hence, the focus will be to give information in a form that

is valuable for scientists. We will provide statistics and graphics that will cover the most typical needs of the users. But this output is more likely to satisfy the average user of HPR than the expert. The expert will probably like to export data from HPR into their own private data base. Therefore, export of data is a central feature of HPR. For HPR it is important that this use also leads to improvement of HPR itself. We will encourage expert users of HPR to improve HPR before data is exported. We are also able to import links from other programs into HPR where we know that the links is of high quality. The imported links need to refer to the unique IDs provided by the Digital Archive for each PSE. HPR may also be used as documentation of research result since it may be document historical persons and events and have permanent IDs for the entire historical population.

For HPR it is also important to be attractive to the large population of hobby genealogists. Many of these persons have a thorough knowledge of the sources. This group is likely to provide the majority of the links in HPR. We would like to attract persons with a variety of interests: local history, follow a family or farm, interest in particular groups of persons or associations and different time periods etc. We will establish close relationship with the relevant associations in order to make HPR a valuable tool for these persons and associations and to promote contribution of high quality to HPR.

## 9 The data base structure in HPR.

HPR is using a MySQL relation data base where the most important tables are:

**Person units (PU)** with ID, name, birth and death time and place, biography, link references.

**Person links** with the ID to PSE and PU, time of linking, score and contributor

**Family relations** with IDs to PSE, family relation, time of linking, contributor.

**References** text, URL, time of registration, contributor.

We only store family relations that are not specified in the source. Both person links and family relations may be negative specifying that these persons are not the same/or not a given family relationship. There will be some graphics and statistics in HPR, but this is not determined so far.

## Acknowledgment

The authors thanks the Research Council of Norway for funding and valuable comments from partners in the project.

## References

1. deCode, [www.decode.com](http://www.decode.com)
2. The Demographic Database (DDB) at Umeå University <http://www.ddb.umu.se>



3. Eide, Ole Jone and Thorvaldsen, Gunnar, Historisk befolkningsregister: Oversikt over emigrasjonen. Heimen (ISSN 0017-9841). 48, 2011.
4. Historical Sample of the Netherlands, se: [www.iisg.nl/hsn/](http://www.iisg.nl/hsn/)
5. Holden, Lars; Thorvaldsen, Gunnar; Bråthen, Torkel Rønold. Historisk befolkningsregister og DNF 1814. Heimen (ISSN 0017-9841). 49 pp 399-414. 2012.
6. Hovig, Eivind. «Norske gener?» i Genialt, nr 1/2010, s. 14–15.  
[http://www.bion.no/filarkiv/2010/07/Genialt\\_1\\_2010\\_norske\\_gener.pdf](http://www.bion.no/filarkiv/2010/07/Genialt_1_2010_norske_gener.pdf)
7. Lokalhistoriewiki.no. a wiki for local history
8. Kjelland, Arnfinn. Busetnadsoge tilsett.hivolda.no/ak/BSS/Busetnadssoege.html
9. NAPP, [www.nappdata.org/napp](http://www.nappdata.org/napp).
10. Odensedatabasen, persons and real estate 1740-1921.  
<http://www.odensedatabasen.dk/>
11. Rotemannsarkivet, persons in Stockholm 1878-1926,  
<http://www.stadsarkivet.stockholm.se/>
12. Sobek, Matthew and Kennedy, Sheela. Identifying Family Interrelationships in the World's Largest Census Microdata Collection. Int. Statistical inst. Proc 58<sup>th</sup> World Statistical Congress, Session CPS070, Dublin, 2011.
13. Thorvaldsen, Gunnar. Using NAPP Census Data to Construct the Historical Population Register for Norway, Historical Methods. 44 (1) 2011, s. 37–47.
14. With, Christian. Slektforskning og personvern. Slekt og Data 1, p. 11-13, 2005.