

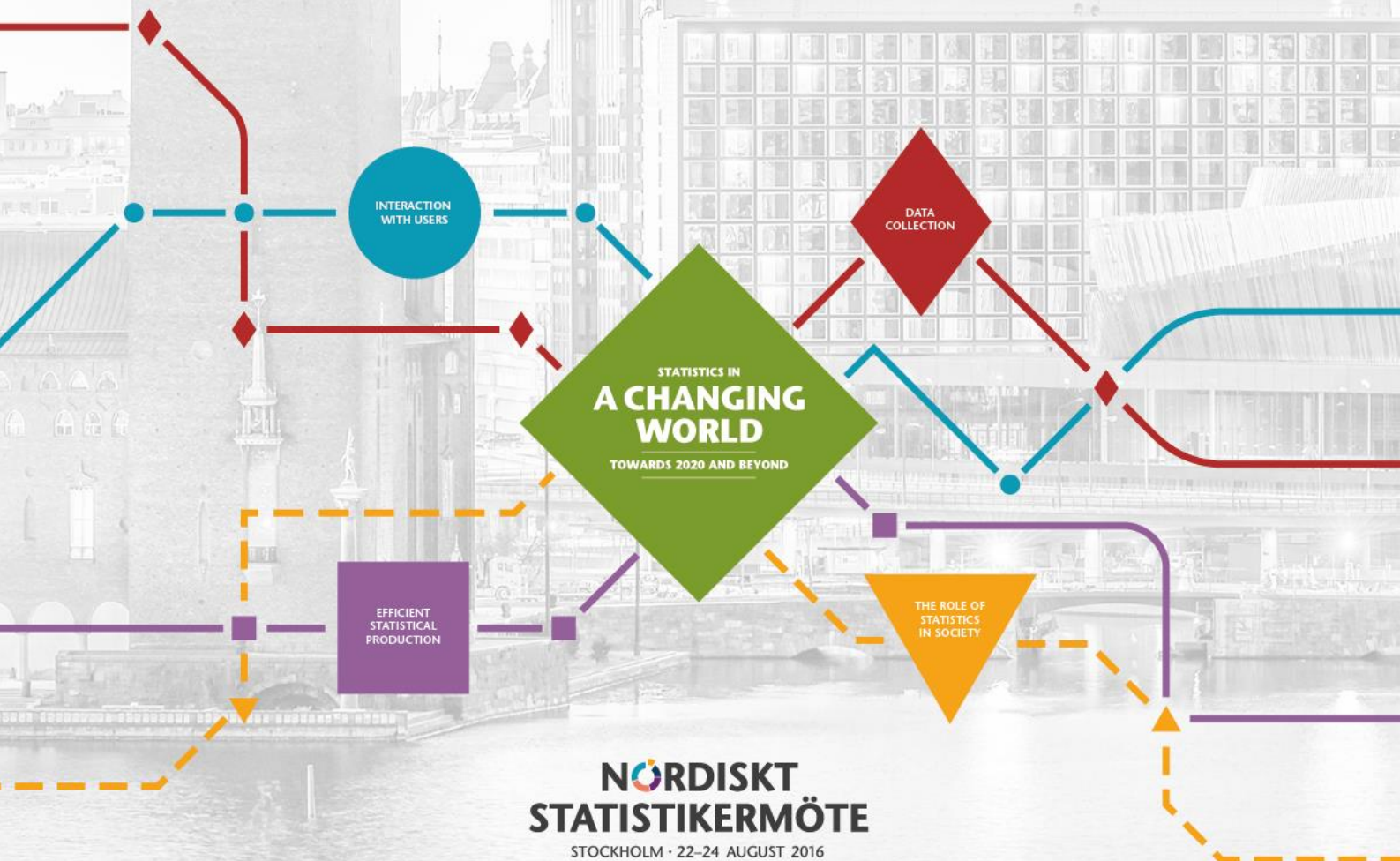
NordMAN: The Nordic Microdata Access Network

C-G Hjelm

Statistics Sweden

Anton Örn Karlsson

Statistics Iceland



NordMAN: The Nordic Microdata Access Network

C-G Hjelm¹, Anton Örn Karlsson²

The Nordic National Statistical Institutes (NSI's) have over the years collected and stored a huge amount of microdata for official statistical production. This microdata is also a valuable source for researchers in the Nordic countries. Although each of the Nordic NSI's have a well formed process with regards to researcher's access to microdata, no cross-national processes or protocols dealing with data requests involving two or more NSI's. Because of interest from the Nordic research community the Nordic NSI's will start a two year project in January 2016 with two main goals: 1) To test an implement a common Nordic model of microdata access for researchers and 2) propose a common metadata framework for cross-Nordic research projects.

The presentation will be aimed at describing the project, its background, main goals and its structure. It is foreseen that this project will increase the ease of applying for cross Nordic microdata access for research purposes and thus increase the satisfaction of users from the research community.

Key words: microdata, metadata, researcher access, cross-Nordic cooperation.

1. Introduction

The NordMAN project is cooperation platform for the Nordic National Statistical Institutes (NSI's) in which the aim is to provide better service to the Nordic research community by improving the access to cross-Nordic data, both by providing a unified access point to cross-Nordic microdata and by providing harmonized metadata across the Nordic NSI's. The project has two broad goals: 1) Test a model for cross-Nordic microdata access which was proposed in the Feasibility Study Regarding Research Access to Nordic Microdata; 2) Propose and implement a common framework for metadata accompanying microdata, available for research purposes across the Nordic NSI's. The project is funded by NordForsk.

¹ Statistics Sweden, Senior Advisor, Örebro, Sweden, Email: claus-goran.hjelm@scb.se

² Statistics Iceland, Methodologist, Reykjavik, Iceland, Email: anton.karlsson@hagstofa.is

2. Background

The background to NordMAN project can be traced back to the meeting of the Nordic Chief Statisticians in Rosendal in August 2013. One theme of the meeting was the exchange of confidential data at a national and at the EU-level, including data exchange for research purposes between the Nordic countries. This particular issue was raised because the Nordic NSI's had been approached by the working group of the NORIA network, which was a project launched by NordForsk. The meeting decided that this matter would best be served by sending it to the Nordic network for microdata access which would carry out a feasibility study on joint Nordic research access to microdata in the Nordic countries.

2.1 Feasibility study regarding research access to Nordic Microdata

The Nordic network for microdata access started its work on the feasibility study in December 2013 with a meeting in Copenhagen. The work proceeded fast, with the group working on comparing legal frameworks and national practices regarding researchers' access to microdata within the Nordic countries.

The final report of the network was presented to the Nordic Chief Statisticians' meeting in Helsinki in September 2014. Among the main conclusions of the group was that the legal situation was similar in the Nordic countries as well as the security situation. The main differences were related to different access systems in the countries as well as differences in how access is given in practice. This highlighted the fact that for researchers applying for access to cross-Nordic data, they have to deal with each of the NSI's, which have different practices with regards to microdata access and different technical solutions for providing access to microdata. Therefore, the group recommended that a common cooperation model of cross-Nordic microdata access would be tested, evaluated and adopted by the Nordic NSI's.

2.2 NordForsk grant application

During the feasibility study, an idea of further work within this area was discussed. There was for example a known issue of the comparability of data between the Nordic countries, i.e. although the countries may have the same registers or surveys, it is not always clear to what extent variables are the same across countries. This problem is exacerbated by the fact that metadata for microdata can also differ between the countries, leaving researchers working with cross-Nordic data in a difficult position. Therefore the idea of a metadata structure for cross-Nordic microdata for research purposes (an idea that had already surfaced at the Nordic Chief Statisticians' meeting in Rosendal) was discussed within the network and the possibility of a constructing a project aimed at building such a structure for cross Nordic data with the help of NordForsk.

On the basis of this a grant application was sent to NordForsk, which was accepted by its board in the beginning of 2016.

3. The NordMAN project

3.1 The organization of the project

The NordMAN project is divided into the following six work packages:

- WP1: Planning and administration (leader: Sweden)
- WP2: Implementing and testing the common model of cooperation (leader: Denmark)
- WP3: Assessment, needs and decision of a future common metadata framework (leader: Sweden)
- WP4: Development and implementation of an IT solution (leader: Sweden)
- WP5: Filling contents into the Nordic metadata model (leader: Finland)
- WP6: Final report (leader: Iceland)

Statistics Norway only participates in WP2.

3.2 Ongoing initiatives

Two of the work packages of the project have already started and meetings have been held in the various cities of the Nordic countries.

Two meetings have been held in WP2 where delegates working within the research services from all the Nordic NSI's have met and discussed practical issues regarding the application process for cross-Nordic data. In order to test the model for cross-Nordic microdata access the work has focused on applications which are based on research projects funded by NordForsk. According to the terms of reference for WP2, its main purpose is the forces and weaknesses of the proposed model, monitor legal changes within the participating countries with regards to microdata access, describe developments in prices of access to microdata and inform each other on developments in technical solutions. The group will provide a document describing routines for microdata exchange, security agreements and communications between NSI's. It will also document outlining researcher routines related to the application form and the researcher agreements as well as communication with research communities

WP3 is led by Statistics Sweden. The group has met four times where participants from the Nordic NSI's (excluding Statistics Norway) and from the Swedish Research Council have discussed metadata. The main issue for WP3 is a technical description and framework of the metadata structure which will be used for describing

microdata from the Nordic NSI's. For NSI's where metadata for microdata is already produced according to a harmonized structure or system, the main issue is to find a structure where metadata can easily be moved from one system to the next. For countries where no harmonized structure for micro-metadata is at hand, the task is much easier as they can use the cross-Nordic harmonized structure in their own work.

At the current stage of work within WP3 a combination of two standards for producing metadata DDI (Data Documentation Initiative) and GSIM (Generic Statistical Information Model) are being assessed with the aim being on the re-usability of the concepts in the structure for increased flexibility.

3.3 Future initiatives

The work on WP4 is expected to start in the fall of 2016 or as soon as WP3 will finish as the metadata structure decided upon by WP3 will be the basis for work in WP4.

WP5 is however both linked to both WP2 and WP4. The metadata that will be filled into the Nordic metadata model will be based on the needs of the NordForsk-backed research projects which are being dealt with in WP2 and also based on the technical solution which will be decided in WP4.

3.4 Additional effects of the project within the Nordic NSI's

Although the project is first and foremost aimed at providing better service to the Nordic research community, it also has an effect beyond that for the NSI's taking part in it. It is especially well suited to increase the modernization of statistical production within the Nordic NSI's. The main reason for this is the impact the project has on the internal way of thinking within the Nordic NSI's and the clear connection to one implementation of GSIM.

4. Conclusion and next steps

The NordMAN project is set to end in the summer of 2017. This is an ambitious timetable and therefore the winter of 2016-2017 will be used for much of the work.

By the fall of 2017 it is foreseen that a metadata framework for cross-Nordic will be up-and running and that results of the pilot project for common access of cross-Nordic microdata will be available.