

Register Based Research Exploiting a Goldmine of Knowledge

Ivan Thaulow, Lars Thygesen

► **To cite this version:**

Ivan Thaulow, Lars Thygesen. Register Based Research Exploiting a Goldmine of Knowledge. Christian Gram; Per Rasmussen; Søren Duus Østergaard. 4th History of Nordic Computing (HiNC4), Aug 2014, Copenhagen, Denmark. Springer International Publishing, IFIP Advances in Information and Communication Technology, AICT-447, pp.385, 2015, History of Nordic Computing 4. <10.1007/978-3-319-17145-6>. <hal-01301616>

HAL Id: hal-01301616

<https://hal.inria.fr/hal-01301616>

Submitted on 12 Apr 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Register Based Research

Exploiting a Goldmine of Knowledge

Ivan Thaulow^(✉) and Lars Thygesen
Danmarks Statistik, Copenhagen, Denmark
{ ith, lth } @dst.dk

Summary

Researchers in Denmark have for many years had outstanding opportunities for exploiting data from a wide range of registers. They can carry out research based on a goldmine of information on social and demographic characteristics of the population, labour market participation, health data and medical treatment, and more. All these data can be combined, and longitudinal analysis can be made.

Researchers can access anonymous micro-data tailored to their needs in a secure setting at the Research Service of Statistics Denmark, who manages the service. Basic data originate from Danish official statistics that are largely based on administrative sources. As the statistics produced by Statistics Denmark cover mainly social and economic domains, cooperation arrangements have been agreed with the State Serum Institute, who manages the Danish health data. Other external data can be integrated into the datasets that are created for researchers, e.g. special surveys carried out among specific parts of the population.

The real challenge of such a service is balancing, on the one hand the need for exploiting the enormously rich data bases for the benefit of society, on the other hand the need to protect the individuals who are registered. For researchers it is of course important that access is easy and allows them to do all the things that will give them the desired results. Legislation on data protection sets limits to access. ICT systems and organizational measures must reconcile these concerns.

The lecture explains how the Research Service has been set up and how the challenges are managed; it also describes the performance of the arrangement and attempts to highlight results and benefits. Experience will be viewed from the researcher's as well as from the data manager's point of view. Finally we will present some of the challenges for the future: How to accommodate needs for cross-border research, and how to include data from emerging sources, commonly labelled Big Data.