



# **Seminar on Register-based Research in the Nordic Countries - Possibilities and Practices**

## **How Norway Make Valuable Registers Available for Research**

**Vigdis Namtvedt Kvalheim, NSD**

National Institute for Health and Welfare,  
Helsinki, April 5<sup>th</sup>, 2011



## Status in Norway

In 2008 the RCN conducted survey to find out if Norwegian researchers, research institutions and data producers comply with the requirements of the OECD guidelines concerning data management and sharing.

### Questions:

1. **Policy and procedures for data archiving**
2. **Policy and procedures for data dissemination and sharing,**
3. **Information about policy and procedures to staff members,**
4. **Perceived need to change procedures (Universities and colleges, Research Institutes University hospitals, Statistics Norway and various data producers)**

### Major findings:

- Institutional support to the principle of open access and data sharing
- See no need to change institutional practice
- Absence of rules and procedures across institutions and research fields.
- Points to the procedure with NSD as a national data archive (required by the NRC grant)
- Points to the procedures for sharing data from National Health Registers and Statistics Norway (required by law)
- Points to the need for national standards and infrastructure
- Ask for a more active role on behalf of the Ministry and the NRC



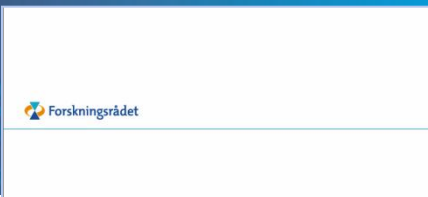
## Status in Norway

**Data situation is unique – world class data**

**The possibilities for research is enormous**

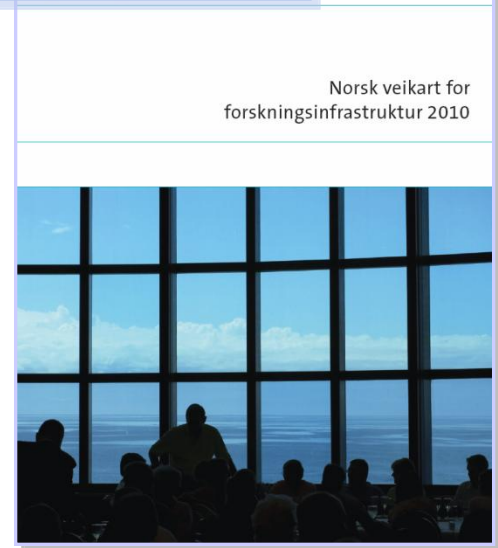
**Existing data is heavily underused**

- All fields need improved access to data
- All fields need better access mechanisms that ensure efficient use and exchange of information and data
- The Medical and Health Sciences and the Social Sciences are well placed with regard to developing common solutions for data management and sharing



Verktøy for forskning

Nasjonal strategi for forskningsinfrastruktur (2008 – 2017)



Norsk veikart for forskningsinfrastruktur 2010



“Data is the single most important component for a science based understanding of society. To promote and facilitate access to these data is to promote research”

**CESSDA-PPP, 2007**

**A Social Science Research Infrastructure Project, supported by the European Roadmap for Research Infrastructures, ESFRI**



## Facts About NSD

### What?

- A National Research Infrastructure, established as part of the Research Council of Norway in 1971
- From 2003 owned by the Norwegian Ministry of Education and Research
- Main funding from
  - Research Council of Norway
  - Ministries
  - EU
  - Universities and colleges

### Why?

- Facilitate empirical research that primarily depends on access to relevant and high quality data

### How?

- Ensure easy access to data and support services for the research community
- To provide guidance for researchers and students in the fields of data collection, data analysis, data methodology, protection of privacy and research ethics
- Remove financial, technical, knowledge-based, legal and administrative barriers between users and data resources





## 50 years later the context is changing

*"the value of data lies in their use"* US NRC

A common understanding that data constitute the key element of a functioning research infrastructure (OECD, ESFRI, Roadmaps)

A common recognition that in spite of huge investment in data production and research, research contributions and returns of public investments are limited

A growing acknowledgement by governments and data producers that existing access mechanisms do not work or are not sufficient to meet research needs



## Major Barriers to Use of Registers

- Institutional and legal bottlenecks
- Technical bottlenecks
- Financial impediments
- Linguistics impediments
- Data quality and coverage problems
- Lack of metadata and documentation
- Qualifications and experience



## Harmonised European Legal Framework – Diversified Access Policies and Legal Practices

A recognition that structural problems (institutions, people, policies and practices and capital resources) restrict opportunities for more research and thus the progress of science and society

Political will to use strong incentives, changing legislation and establishing new funding schemes **to support and enforce data access and exchange**





## Status in the Nordic Countries

### Comparative advantage

- Data resources and possibilities of data merging
- Legal framework by and large in place
- Potential for national coordination and cooperation

### Comparative disadvantages

- Research resources (money and number)

National coordination to ensure an optimal division of labour between research and support services



## The Norwegian Model – Collaboration

- **Research Council of Norway**
  - Funding of existing archives that contains data on individuals, regions, and political and administrative systems
- **Data Protection Agency**
  - NSD as a broker between the Agency and the scientific community. Today NSD function as Data Protection Official for research
- **Statistics Norway**
  - Since 1976, a formal agreement on dissemination of data for research purposes. Allowing for extensive use of aggregated and individual data; for instance surveys and register data
- **Research Community**
  - Close contact with researchers in different scientific fields

*"Specialised support services should be considered as means to ensure the cost-effective production, use, management and archiving of research data." (OECD)*



## Cooperation with Statistics Norway

### The purpose of the agreement (1976, formal agreement)

- To give researchers and students the easiest possible access to the largest possible amount of Statistics Norway's data, without violating the current regulations and without relaxing the protection of individual data

*The term researcher refers to researchers affiliated with a university or college in Norway or abroad, and the institutional sector in Norway.*

### The scope of the agreement

- Aggregated (table) data for municipalities and census-tracts
- Anonymous from Statistics Norway's sample surveys or other anonymous individual data (persons and establishments)  
**Scientific Use Files, <http://www.nsd.uib.no/solr/nsu>**
- **De-identified data to be rendered anonymous before dissemination (1997 extended cooperation)**



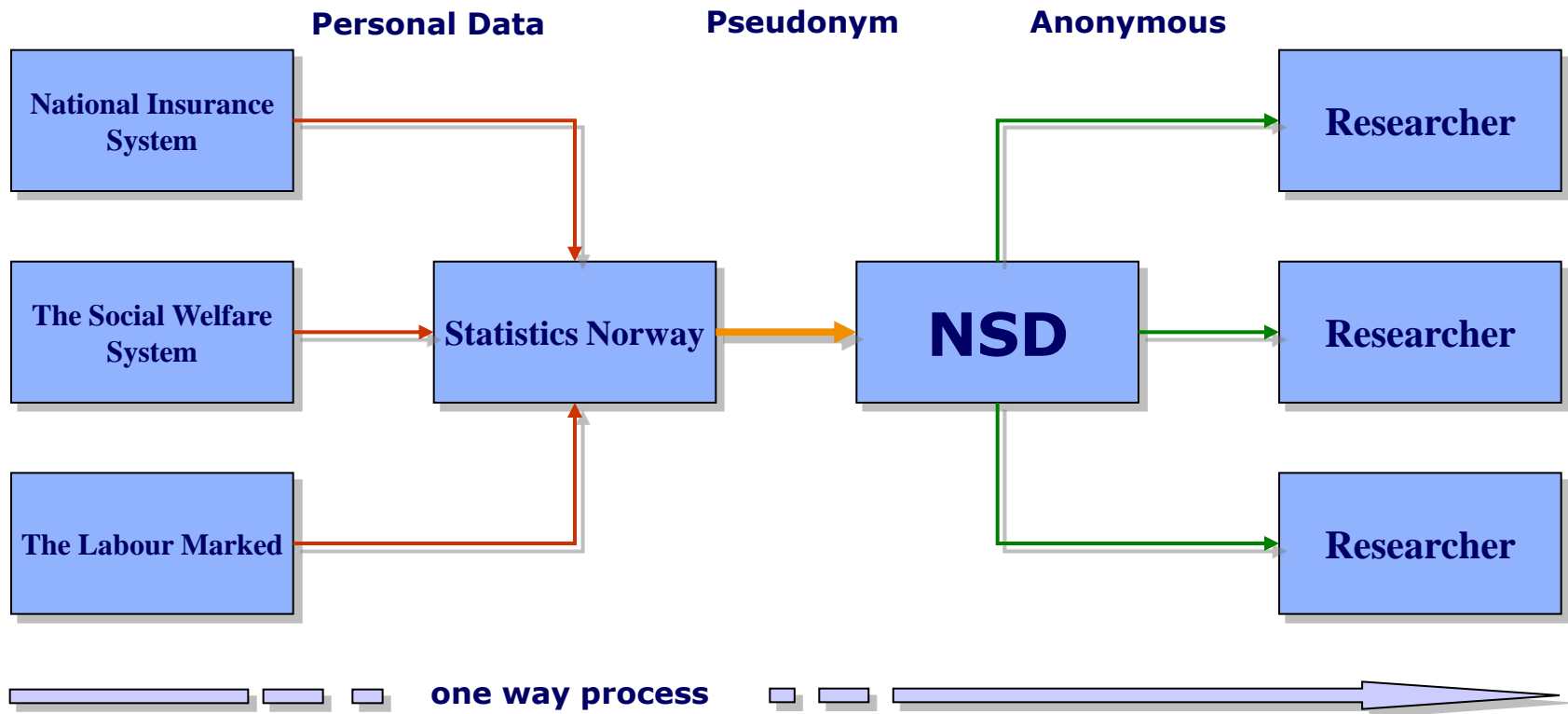
## The Data Archive as a Safe Data Access Mechanism

### KIRUT-project – Extended cooperation - NSD act as a data broker on behalf of various public data producers and register owners

- Event history database covering eight years
- Data about individuals contact with the welfare system and the labour market
- **10% yearly up-dated sample** of all Norwegians between 16 and 67
- Data from a number of governmental registers
- Merged and made pseudonymous by Statistics Norway
- Transferred to and stored at NSD
- NSD manage access to anonymous and (in special licensed cases) de-identified data for researchers and students
- Used in hundreds of research and student projects



# Access Arrangement – Using a NSD as a Data Broker





## FD-Trygd's<sup>1</sup> First Period

- FD-Trygd replaced KIRUT in 1999
  - Covers the whole Norwegian population from 1992 to present
  - Covers all major social security schemes
  - Administrative data from The Norwegian Labour and Welfare Service (NAV) and Statistics Norway
  - SN deliver non-anonymous individual level data
- NSD as a support service:
  - Intermediates in the application process for researchers and students
  - Deliver anonymous data on behalf of Statistics Norway
  - Received non-anonymized data from SN for disclosure

**<sup>1</sup>FD-Tygd is Statistics Norway's events database, which includes details on demography, social conditions, social security, employment, search for work, state employees, income and wealth. The data covers from 1992 to present.**



## FD-Trygd's Second Period

- FD-Trygd was a success
- Increasingly capacity problems and in 2005 waiting time for users six months or more
- Data from FD-Trygd was free of charge through funding from the Research Council until 2007
- Result; high cost, delays, increased barriers for students and researchers



## Data from Registers – Possible Future Solution

### New Agreement between Statistics Norway and NSD:

- Statistics Norway and NSD are in the process of finalizing a new agreement
- Increase the use of FD-Trygd data and secure free and equal access to data
- Discussions are based on the Norwegian legal framework, the Statistics Act (1989) and the Personal Data Act (2001)
- Building on the experience from more than 30 years of close cooperation
- Transferring pseudonymous or de-identified personal data from Statistics Norway to NSD
- NSD should be able to deliver anonymous data from the **20% FD-Trygd sample** before this summer, and not later than September
  - De-identified with approval from register owners and the Data Protection Official/Data Inspectorate





## Advantages

### Data Producers

- Reduce pressure/share the work load
- Supplement direct access as an efficient and cost effective data provision mechanism
- Extensive and qualified use
- Quality control
- Visibility and usability
- Support, trust and legitimacy

### Research Communities

- Easy and equal access to rich data resources
- Cost and time efficient
- Qualified data and support service
- Quality control



## **RAIRD - Remote Access Infrastructure for Register Data**

Statistics Norway and NSD have applied to RCN's infrastructure program to establish a national research infrastructure providing remote access to large amounts of rich high-quality statistical data for scientific research while at the same time managing statistical confidentiality and protecting the integrity of the data subjects.

### **The focused collaboration between SN and NSD will:**

- Increase the total dissemination capacity in this field
- Make the two institutions by concerted action able to handle an expected increase in requests for data in an efficient way
- Contribute to increased use of quantitative data in health, welfare- and other socio-economic research both nationally and internationally.



## **RAIRD - Remote Access Infrastructure for Register Data**

### **Short term the goals**

- Develop and implement a full scale sustainable data dissemination service for micro-oriented data set up as a Secure Data Service
- Offer equal opportunity to access metadata and applications that makes it possible to carry out data analyses and produce anonymized analytic results online.
- Develop a data access system based on a formalized data access policy and procedures for authorized users and uses of data.
- Develop and implement a system for controlling analytic outputs
- Investigate, propose and implement solutions to a range of technical, procedural, legal, methodological and documentation



## **RAIRD - Remote Access Infrastructure for Register Data**

### **Long term goals**

- To stimulate cooperation between researcher and research projects across projects, regions and countries.
- Contribute to the development and dissemination of knowledge
- Enlarge the accessible data content and investigate the possibilities and conditions for covering further data resources and a wider range of researchers.
- Implement a procedure for access to customised micro data files in the SDS for projects needing data not available in the “standardized” micro data collection.



## Why are both Statistics Norway and NSD focusing on services based on extended cooperation in this area?

- Register data constitute a gold mine for many scientific fields; economy, sociology, medicine, demography, heavily underused
- For SN data provision through NSD (as safe data service) is seen as an important tool to meet increasing demands and to fulfill its obligations to research and society as set out in the Statistics Act
- Complimentary to other access mechanisms. Admission that existing access mechanisms are not sufficient to meet demands.
- Sharing of labor and costs, e.g. involving anonymisation and services
- A rich data resource that potentially gives Norwegian researchers a clear advantage compared to other (non-Nordic researchers)
- There is a clear potential for close Nordic cooperation within this area



## Future Models for Data Access - ESFRI Process

CESSDA ERIC - Develop CESSDA into robust European wide institutional framework on sound financial and legal footing.

Data Without Boundaries (DWB) – Partnership CESSDA and Statistical offices

Partnership a potentially efficient and cost-effective data provision tool for data producers (national statistical offices and other governmental bodies)

Managing access to a variety of data sources for research purposes and at same time protecting the privacy interests and the integrity of the data





## The Norwegian Experience

- Cooperation and partnership work (served as a model for making data from the national health registers available for research (Ot. prp. Nr. 5 on the Health Registers Act 1999-2000))
- It is possible to provide access and at the same time maintain the trust of the public
- Laws and regulation do not necessarily lead to restricted or no access, on the contrary
- An adequate legal framework is important to protect privacy as well as access interests



# Thanks for listening!