

Helsinki

5. April 2011

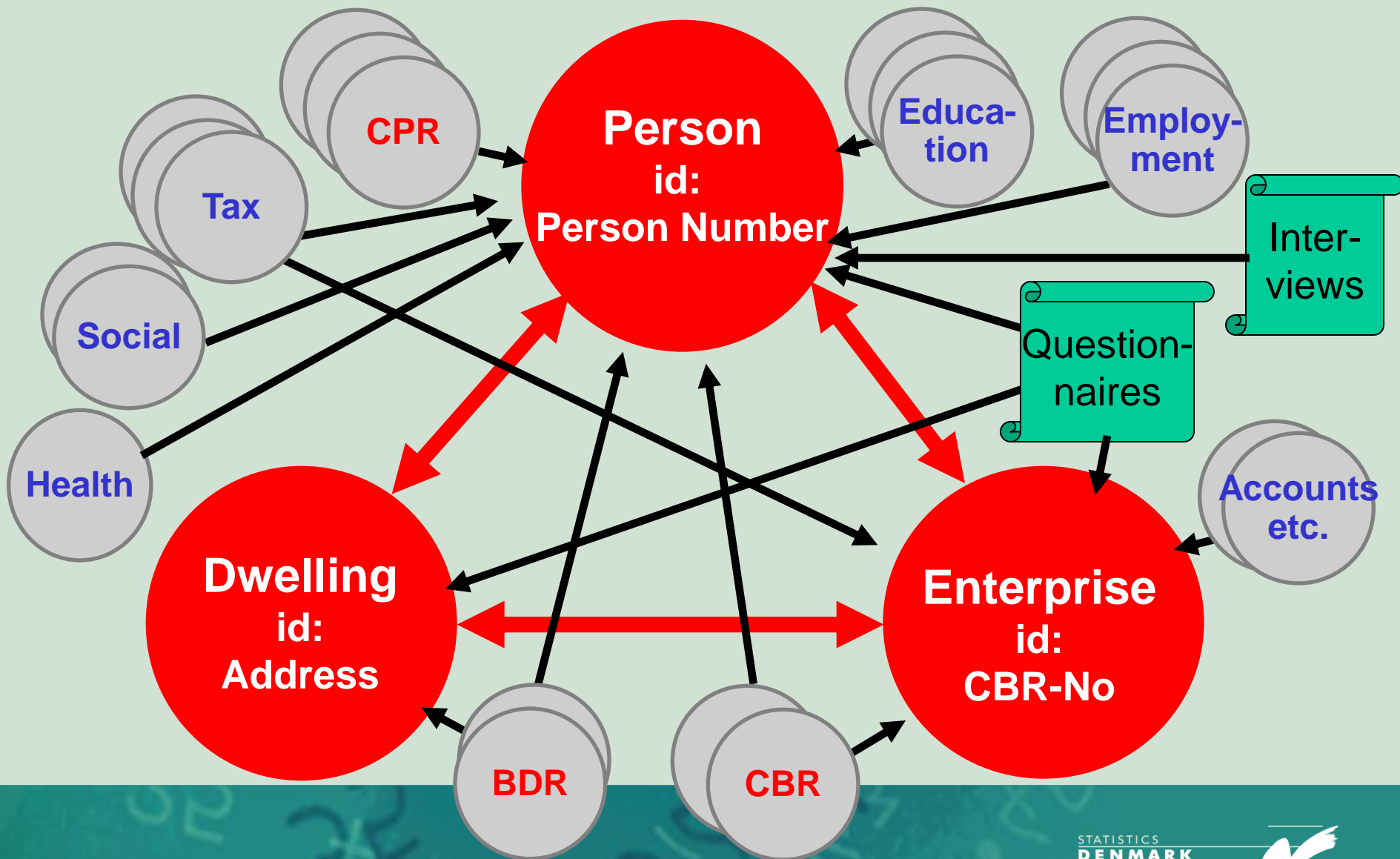
**Researchers Access to Micro-data in
Statistics Denmark**

Ivan Thaulow

**Head of Division
Research Services
Statistics Denmark**



The Danish Statistical info System

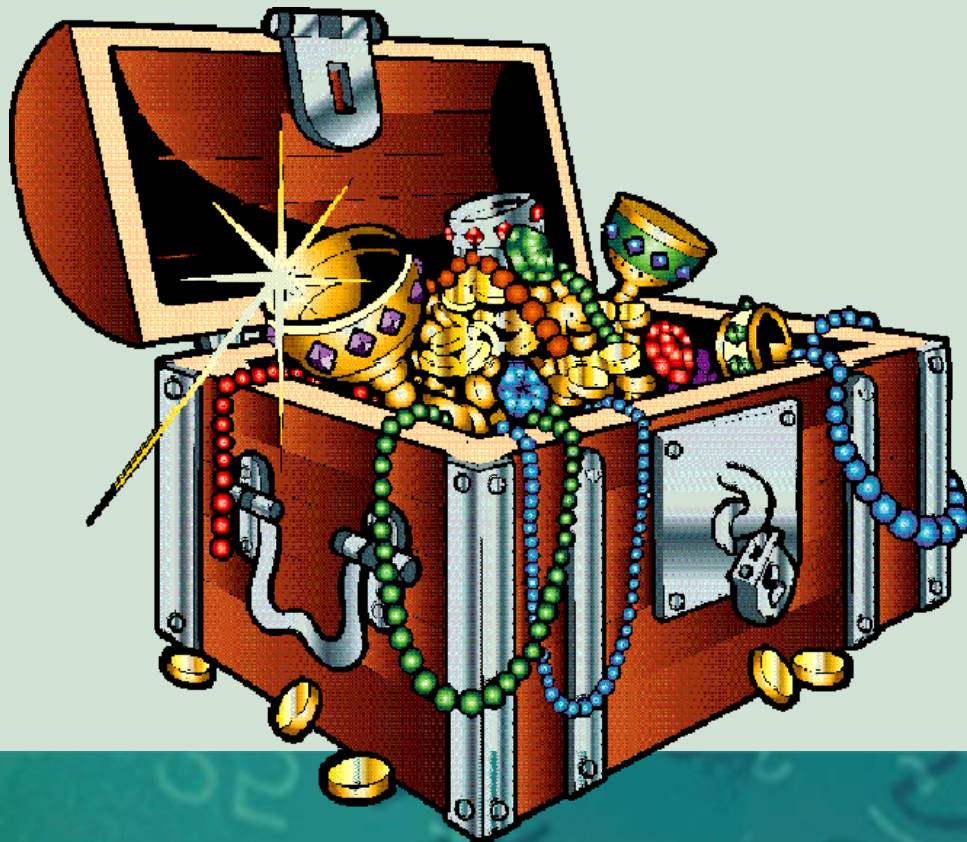


Data linkage

- Registers linked by various keys:
 - Persons – Families - Dwellings
- Possibilities of combining data are infinite
- Also possible to combine data with own surveys or registers outside SD
 - i.e. Register of Medicinal Products
- Unique research potential
- Illustrated by many research projects in Economy, sociology, demography, epidemiology etc.



A gold mine of data!



- High data quality
- Comprise entire population
- Registers cover all years back to 1980.
- Longitudinal studies

The Research Service Division

- All dataaccess goes trough a specialized unit in DS: The Research Services Division
- The division has 17 employees
- It coordinates all research related services and data from all divisions in Statistics Denmark etc.
- It creates all datasets for researchers
- It advises researchers – on new as well as old research projects
- It administrates the remote access arrangement, authorizations, contracts, tokens, passwords etc.



Who can get access to microdata?

- Only researchers from specially **authorized institutions** (stable institutions with a responsible manager and a number of researchers/analysts)



Security

- The Head of SD signs all Authorizations: Responsibility & supervision
- The researcher signs declaration to follow rules for access to microdata
- Possible to close an authorization if rules for access is not followed



Limitations on data-access

- Access to data according to a "need to know" principle
- So, SD can reduce data if necessary (samples and grouping)
- Single companies cannot have access to microdata with data on business units



The Practical Solution

- De-identified microdata always stays in SD
- Special research computers in SD for remote access
- One basic public paid system
- Private computer-solutions is an option
- Separated from the statistical production



Practical solutions (continued)

- Log-on via tokens and strictly personally password
- Software: Flexible but most used is SAS, STATA, GAUSS, SPSS



Output

- Researchers are not allowed to print individual records and cannot download data or results
- Output with statistical results is forwarded by e-mail
- All mail is logged and manually random checked by SD



All access to data is online

- On-line access is possible for foreign researchers all over the world. They just have to be affiliated to a authorized Danish research environment
- On-site arrangement closed down



The end

