

## Access for scientific purpose

- **Achievements after 6 years of implementation**
  - 2 business microdata sets (CIS and SES) available through Eurostat safe centre
  - 4 anonymised microdata sets available on CD Rom under licence contract
  - 23 non-EU research bodies declared admissible
  - More than 200 researchers covering all major themes of socio-economic research and policy with now about 90 research contracts per year

1

## Confidentiality provisions

### – In general (2)

- **Statistical disclosure control**
  - Obligation to ensure protection
  - Obligation to harmonise
  - Implementing powers given to the Commission (Eurostat)



## Policy on access to RMFs

- The CSO **may** provide access
- Access can only be granted for statistical purposes
- First option will be to see if the provision of aggregates will suffice
- In general only one version of an RMF for a survey instance
- Responsibility for confidentiality rests with individual (s) appointed “Officer of Stats”
- CSO reserves right to put outputs into public domain

## An heterogeneous and complex system at national and European levels

- Information for procedures and applications in different places and not always easily accessible on websites
- Differences in arrangements (who is doing what)
- Differences in criteria for accrediting and decisions not systematically relying on advices by committees involving researchers
- A great variety of model end user licences for same type of microdata

# Framework for decisions

1. Legal
2. Researcher needs
3. Data Strategy
4. Cultural
5. Technical
6. Partnerships
7. Money
8. Methodology
9. Governance
10. Inertia / familiarity

5

Economic and Social Research Council

## Next Steps



- Early stages – over the next two years need to establish:
  - relationships with potential data providers (especially the UK Office for National Statistics)
  - scale of future demands on the service – both current and projected
  - links with other relevant research resource (ESDS, ADLS, VML, ‘Safe-settings’ in other countries, similar developments from other research funding agencies)

## Conclusions

- It is possible to collect and provide data in a more useful way than often done.
- Education information should be collected in all countries in all required details to allow valid representation of relevant educational differentiation.
- EU-databases such as EU-LFS or EU-SILC should include in their microdata releases education codes in the most detailed ISCED-version.
- In addition to the harmonised codes, the dataset should also include the detailed national information in order to
  - allow further validation research to improve international comparability of education measures
  - enable researchers to construct further typologies that are most adequate to the theoretical needs of their research.
- Problems similar to those found for education likely exist in other areas. To improve both data protection and usefulness of data for research and statistical reporting it is urgent to institutionalize further intensive exchange between research data users and data providers.

7

## Barriers in access to microdata

Perspective of Statistics Netherlands

Jan Mol  
Centre for Policy Related Statistics  
Statistics Netherlands (SN)



	Public use Files	Scientific use Files	Extracts (subsets)	Public tabulations	Special (bespoke) tabulations	Secure remote access / execution	On-site safe centres
Population Census, Register, or Microcensus	3-5 countries	6-8 countries	> 9 countries	> 9 countries	> 9 countries	6-8 countries	6-8 countries
Household Surveys	> 9 countries	> 9 countries	6-8 countries	> 9 countries	> 9 countries	6-8 countries	6-8 countries
Data from Administrative Registries	3-5 countries	3-5 countries	3-5 countries	6-8 countries	6-8 countries	6-8 countries	3-5 countries
Business Data	3-5 countries	6-8 countries	6-8 countries	> 9 countries	> 9 countries	6-8 countries	6-8 countries
Other economic and financial data	3-5 countries	3-5 countries	3-5 countries	6-8 countries	3-5 countries	3-5 countries	3-5 countries

**Table 1:** Current modes of access, regardless of distributor, in European countries.

> 9 countries
  6-8 countries
  3-5 countries
  < 3 countries
 ▶

cessda

## Functions of CESSDA Archives

- *Direct* dissemination to researchers:
  - Anonymised data and (sometimes) confidential data, as shown above.
  
- In addition, *indirect* mediation services:
  - Documentation, metadata preparation (GR, FR...)
  - Anonymisation (DE, NO, ...)
  - Archiving, deposit of government data (UK, ...)
  - Covering costs for access (FR, ...)
  - Information and support to data users (SE, SI...).

2/16/2009
CESSDA Data Archives

## Advantages

### Data Producers

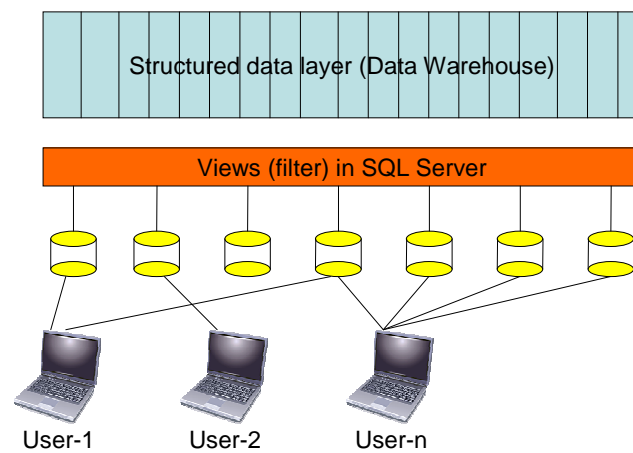
- Reduce pressure/work load
- Supplement direct access as an efficient and cost effective data provision mechanism
- Extensive and qualified use
- Quality control
- Visibility and usability
- Support, trust and legitimacy

### Research Communities

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

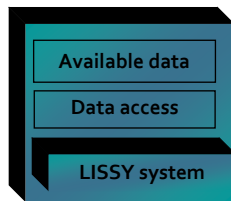
11

## Structure in MONA



12

## LISSY - A Remote Access System Its Implementation towards security issues (2/2)



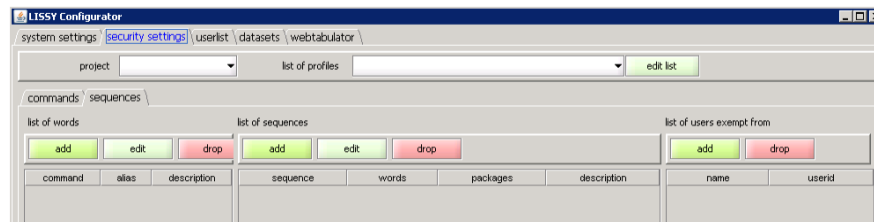
Lissy is flexible enough to configure a level of confidentiality that matches the requirements of your institution

- Restrictions of queries based

statistical package, project, user profile, datasets

All

incoming jobs and outputs are stored



## Secure Data Service

- Two-year pilot began 1 October 2008
- ESRC-funded and ONS-approved
- Remote access to sensitive and disclosive data held centrally
- All computation on secure central server; no data on local machines
- Outputs vetted for disclosure issues before release
- Expecting tens not hundreds of users in first instance

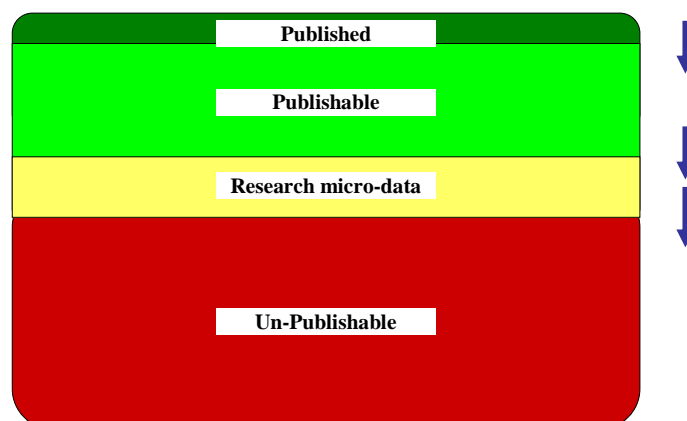


### 3. Network of national safe centres

- **Planned ESSnet project:**
  - „Decentralised access to EU microdata sets“
- **Development of a „Decentralised access to EU microdata sets“ within a researcher from a certain MS can use European datasets in his own MS**
- **The advantage of the network of safe centres is that every single MS can widen the supply of microdata**
- **In addition to national datasets the researchers can use EU microdata in their MS (without traveling to Luxembourg)**

© Federal Statistical Office, Research Data Centre, Maurice Brandt

## The Information Levels



16

# The Trade-offs

