Integrating Data Collections with Open Data Resources: a communication-centered approach

Thomas Bourke
UNECE, 30 June 2017
European University Institute

- Founded in 1976
- Located near Florence
- 55 Professors & 30 Part-time Professors
- 100 post-doctoral Fellows
- 575 PhD candidates
- Economics, Law, Sociology, Political Science, History
Multiple Communities

- Official Statistics producers
- Data publishers (commercial)
- Data users
- Researchers as data creators
- Data repositories
- ICT community
- Library community
Research library priorities

• Centrality of data user
• Heterogeneous data resources
  – Relevance
  – Quality
Dissemination

• Research libraries re-disseminate

  (i) official statistics (including micro)
  (ii) commercially-subscribed data
  (iii) open research data
<table>
<thead>
<tr>
<th>Restricted – free</th>
<th>Open – free</th>
</tr>
</thead>
<tbody>
<tr>
<td>eg. micro data</td>
<td>eg. official statistics</td>
</tr>
<tr>
<td>Restricted – paid</td>
<td>Open – paid</td>
</tr>
<tr>
<td>eg. financial data</td>
<td>eg. fremium</td>
</tr>
</tbody>
</table>
Restricted / Open

Restricted

Open

Finance & Micro

Macro
Financial & micro data
Universities are working to create synergies between traditional 'data collection' roles and newer 'data service' roles.
## Traditional research library roles

<table>
<thead>
<tr>
<th>Acquisition/deposit</th>
<th>Serials maintenance</th>
<th>Paper / CDs DVD / Databases</th>
<th>Cataloguing/classification</th>
<th>Terms &amp; conditions of access and use</th>
<th>Reference support</th>
<th>Preservation</th>
</tr>
</thead>
</table>

---
New library roles

- ‘Mixed’ portals - paid and free resources
- Data literacy
- Recognising researchers as data creators
- Research data management / DMPs
- Metadata assistance
- Repositing / preserving
Data Discovery

Data repositories

Metadata for research datasets

Research Data Management / DMPs

Software / ICT support

Data literacy

Acquisition / Commercial databases / Open data

Data portals

Terms of access and use
Economic and Data Information Services

- For information, contact Thomas Bourke: econlibrary@eui.eu

Economics

- Library Catalogue
- Economics Journals
- Economics Books
- Book Delivery Service
- Economics Working Papers
- Departmental Library Services
- Books & Data Manuals in Departmental Library

Statistical Data

- Data Portal
- Macro Data
- Micro Data
- EU & Euro Area Data
- Data Services and EUI ResData
- Terms and Conditions of Use
- Library Welcome Page for 2017-18 Economists

Global Financial Data

EUI members now have access to the Global Financial Database providing long-range financial data on stocks, bonds, bills and other instruments; covering approximately 200 countries from the medieval period to present. Full details are on the Library’s EUI GFD resource page.

Research Data Services

Research data services are introduced on this Library page. The workflow for the (beta) EUI ResData repository for sharing and archiving EUI-generated datasets is now available online. The 2017 edition of the Library Research Data Guide was published on 8 May. For weekly news about data resources, sign up for the Data e-Bulletin. Databases are accessible via the Library Data Portal.
Library Data Portal, Data Services and ResData Repository

For research data support please contact Thomas Bourke

**DATA PORTAL**

The EUI Library Data Portal provides access to Library-licensed macroeconomic, micro-socioeconomic and Europe-related databases. 76 resource guides provide information on data variables, coverage, user networks, online manuals and terms and conditions of use. The Data Portal also provides details about the Library’s restricted micro data server.

**DATA SERVICES**

Click below for EUI Library data services including: research data management (RDM); data management plans (DMPs); data support for Horizon 2020 projects; metadata and data documentation; preservation and open data. For data support, please write to econlibrary@eui.eu or visit one of the two information desks: (BF-085) or (VLF-035).

**EUI ResData**

From May 2017, EUI members can submit their research datasets for inclusion in the EUI ResData repository - a solution for sharing and archiving EUI research data outputs. ResData is based on the DSpace infrastructure. Full details - and the online form for submitting details about a dataset for repos - are on this Library page.

1. **EUI Library Data Portal and Data Discovery**

The Library maintains an indexed Data Portal providing access to both (i) licensed data resources for EUI members and (ii) open data resources relevant to EUI research themes. There are three sub-directories: macroeconomic data; micro-socioeconomic data and Europe-related data.

- **Macroeconomic** databases provide statistics on national, regional and global economic and political-economic developments from international institutions and data publishers, including Thomson Reuters, ECB, Standard & Poor’s, Eurostat, OECD, IMF and World Bank.
- **Micro-socioeconomic** databases provide individual, family, household-level and company data observations. Major providers include GESIS, ICPSR, Bureau van Dijk, UKDS, Eurostat and DIW. Access to micro data on the EUI Library restricted data server requires additional registration.
- **European, EU and Euro Area** data resources provide macroeconomic, micro-socioeconomic and financial data for research on pan-European topics, EU states, European sub-state regions, the Euro area and Europe in global context.
## Economic and Social Data Portal

*For support please contact Thomas Bourke*

### Macroeconomic data

<table>
<thead>
<tr>
<th>Macroeconomic data</th>
<th>Micro-socioeconomic data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Macroeconomic data</td>
<td>Micro-socioeconomic data</td>
</tr>
<tr>
<td>Banking data - SNL Financial</td>
<td>Adult Education Survey - Eurostat</td>
</tr>
<tr>
<td>Central and Eastern European data</td>
<td>Amadeus European company data</td>
</tr>
<tr>
<td>ComExt: Intra- and Extra-EU trade</td>
<td>Competitiveness Research Network (CompNet) Europe</td>
</tr>
<tr>
<td>Comparative constitutions data</td>
<td>Consumer Expenditure Survey (USA)</td>
</tr>
<tr>
<td>Comparative Political Data Sets</td>
<td>Current Population Survey (USA)</td>
</tr>
<tr>
<td>Cross-National Time-Series</td>
<td>DNB Household Survey (Netherlands)</td>
</tr>
<tr>
<td>Data Repositories Registry - re3data.org</td>
<td>DReMM migration data</td>
</tr>
<tr>
<td>Datasync - Thomson Reuters</td>
<td>ECRi European household lending data</td>
</tr>
<tr>
<td>Demographic Yearbook - UN</td>
<td>EU-LFS: Labour Force Survey</td>
</tr>
<tr>
<td>Direction of Trade Statistics - IMF</td>
<td>EU-SILC: Statistics on Income &amp; Living Conditions</td>
</tr>
<tr>
<td>Electoral and parliamentary data</td>
<td>EUDO data on European democracy</td>
</tr>
<tr>
<td>EUI open data registry</td>
<td>Eurobarometer surveys</td>
</tr>
<tr>
<td>European Capital Markets Institute (ECMI) data</td>
<td>EUROMOD tax and benefits simulator</td>
</tr>
<tr>
<td>European Central Bank data</td>
<td>European Community Household Panel</td>
</tr>
<tr>
<td>European Data Portal (EC)</td>
<td>European Social Survey</td>
</tr>
<tr>
<td>European regional (sub-state) data</td>
<td>European Values Study</td>
</tr>
<tr>
<td>Eurostat data</td>
<td>Eurostat Micro public use files</td>
</tr>
<tr>
<td>Eurostat Yearbook</td>
<td>German Socio-Economic Panel data</td>
</tr>
</tbody>
</table>

### Micro-socioeconomic data

<table>
<thead>
<tr>
<th>Micro-socioeconomic data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Classification of data</td>
</tr>
<tr>
<td>Conditions of use</td>
</tr>
<tr>
<td>Data protection</td>
</tr>
<tr>
<td>Data deposit (beta)</td>
</tr>
<tr>
<td>Data services</td>
</tr>
<tr>
<td>Data software manuals</td>
</tr>
<tr>
<td>Data news @ sign up</td>
</tr>
<tr>
<td>Economic research resources</td>
</tr>
<tr>
<td>EUI research data registry</td>
</tr>
<tr>
<td>European, EU &amp; Euro Area data</td>
</tr>
<tr>
<td>Horizon 2020 data support</td>
</tr>
<tr>
<td>List of data resources (Excel)</td>
</tr>
<tr>
<td>Loading data for Windows</td>
</tr>
<tr>
<td>Macroeconomic data directory</td>
</tr>
<tr>
<td>Micro-socioeconomic data directory</td>
</tr>
<tr>
<td>One-stop guide</td>
</tr>
<tr>
<td>Open data</td>
</tr>
<tr>
<td>re3data repository registry</td>
</tr>
<tr>
<td>Registration for micro data</td>
</tr>
</tbody>
</table>
EUI ResData repository: archiving and sharing datasets

- EUI members can submit their research datasets for inclusion in the EUI ResData repository (launched 8 May 2017).
- The first step is to complete the form below. The Library will make an appointment for data transfer later.
- By submitting this form, EUI members acknowledge that the dataset for deposit is the output of original data collection and elaboration; or the output of significant, value-added, elaboration of pre-existing sources.
- This form may also be used to submit a dataset for inclusion in the EUI's Zenodo data collection. The Library will provide case-by-case advice on which repository is suitable.

GUIDELINES

Dataset Creator(s)

- Datasets presented for inclusion in the EUI ResData repository must be the output of research by a current EUI member or research team.
- The name of the Principal Investigator, researcher(s), and - if applicable - technical collaborator(s) who created the dataset, must be provided. EUI email contacts must be indicated.
- If the project is undertaken in the context of a consortium, the name of the Data Manager should be provided (if different from the Principal Investigator).

Dataset as original work and source(s) of data

- Dataset creators must certify that their work complies with the Code of Ethics in Academic Research of the European University Institute.
- Scholars will be required to sign a declaration that they are creators of the dataset being presented for deposit, and that the dataset is the output of original data collection and elaboration; or is the output of significant, value-added, elaboration of pre-existing sources.
- The source(s) of the data must be indicated. If the dataset is the output of original data collection and elaboration, details must be provided. If the dataset is derived from pre-existing sources, those sources must be clearly indicated (data creator, institutional source, publisher).

Permission to share dataset as open data

- Either open access will be provided (where disclosed by the EUI ResData repository) or data will be distributed by

Dissemination, use, elaboration, new dataset outputs
Challenges

• Communication
• Dissemination
• Infrastructure
Quality

• Eurostat and the broader European Statistical System "see quality as the basis of our competitive advantage in a world experiencing a growing trend of instant information which often lacks the necessary proof or quality."

Walter Radermacher

Conference of European Statistics Stakeholders (CESS)

Budapest, 20 October 2016
Communication

- Strengthen the dialogue:
  - Official statistics producers (including micro)
  - Commercial data publishers
  - Research data / open data community
Dissemination

- Strengthen the relationship:
  - Official statistics agencies
  - Data publishers
  - Research repository community
In order to adapt to the evolving professional and technical demands, the necessary competences and skills can be acquired by training current personnel; by collaborating with universities and other networks; or recruiting individual data scientists.

Conference of European Statisticians (19–21 June 2017)
Infrastructure

• International registry / union catalogue?
  – Official statistics resources
  – Commercial data resources (abstracts)
  – Research data repositories

• Cross-indexing synergies:
  • Primary
  • Derivative files
Integrating Data Collections with Open Data Resources: a communication-centered approach

Thomas Bourke
UNECE, 30 June 2017