Data Sharing in Taiwan: Policies and Practice

Ruoh-rong Yu
Research Fellow and Executive Director
Center for Survey Research
Research Center for Humanities and Social Sciences
Academia Sinica, TAIWAN
Current Status of Data Sharing in Taiwan
Data Sharing in Taiwan: SRDA

- Survey Research Data Archive (SRDA)
  - Center for Survey Research (CSR), established in 1993, originally aimed to implement surveys for researchers at Academia Sinica
  - In 1994, CSR set up the SRDA to preserve and disseminate the survey data collected

~ Now has the largest collection of government and academic survey data in Taiwan
Data Sharing in Taiwan: Government

Some government organizations have own data centers, for instance,

- Directorate General of Budget, Accounting and Statistics (DGBAS), established in 1931
  
  - Censuses and surveys conducted by DGBAS

- Collaboration Center for Health Information Application, established in 2011
  
  http://biostat.cmu.edu.tw/~nhip/indexn.html
  - Administrative records of National Health Insurance
  - Registration records of Taiwan aborigines
Data Sharing in Taiwan: Survey Projects

- Some large-scale survey projects maintain their own data archives, for example,
  - Taiwan Social Change Survey (TSCS)  
  - Taiwan Integrated Postsecondary Education Database  
    https://www.cher.ntnu.edu.tw/
  - Taiwan Election and Democratization Survey  
    http://www.tedsnet.org/cubekm2/front/bin/home.phtml
Data Sharing Policies in Taiwan
Data Policies: Administrative Records

- Open data is a hot issue for Taiwan government in recent years
- Collaboration Center for Health Information Application, established in 2011, is the first official data center for administrative records
- However,
  - Other administrative record data (e.g., tax, labor insurance, entrance exam to college) are not accessible for most researchers
  - Linkage of administrative records with survey data are often prohibited
Data Policies: Research Funded by NSC

- National Science Council (NSC)
  - As the main funder of researchers in Taiwan, since 1999 NSC requires the survey projects subsidized by it should deposit their data in SRDA within a certain number of years
  - However, the requirement is not mandatory

- For survey projects funded by NSC,
  - SRDA would contact PIs for data deposit and dissemination
  - Under the subsidy of NSC, SRDA provides workshops and courses on data cleaning and metadata preparation for PIs and their RAs
Data Policies: Research Funded by Others

- Other government organizations
  - For research projects subsidized by other government organizations, there are no common or consistent data policies

- Universities or research institutes
  - For the survey projects subsidized by universities or research institutes (e.g. Academia Sinica), there are no clear policies on data sharing
Data Policies: Ethical Ground

- Academic associations
  - Some associations have setup their ethical codes in recent years
  - These ethical codes are primarily related to data collection and utilization

- IRB
  - In these two years, IRBs on humanities and social sciences have been founded under the request of NSC
  - IRB review is mainly concerned with data collection and data linkage
Data Policies: Law Ground

- Personal Information Protection Act
  - Implemented since October 2012
  - Contains strict regulations on collecting, processing and utilizing of personal information
  - This makes many government organizations and survey institutes hesitate to release their data to researchers, especially when the data is confidential or sensitive
Introduction to Survey Research Data Archive (SRDA)
Survey Research Data Archive (SRDA)

- Established in 1994 by Center for Survey Research (CSR)
- CSR is one of the five research centers at the Research Center for Humanities and Social Sciences (RCHSS), Academia Sinica
- Academia Sinica
  - As the nation's premier research institution, is directly responsible to the President of Taiwan
  - It supports research in a wide variety of disciplines, ranging from sciences to humanities and social sciences
Welcome to SRDA!!

Please beware that this website is only works in Internet Explorer !

Any questions? Skype csr_209!! (message only)

Mon-Fri: 8:30am-5:00pm(TST) [GMT+8]
Tel: 886-2-27871829

New Release

<table>
<thead>
<tr>
<th>Date</th>
<th>Project Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>2012/08/01</td>
<td>青少年聴覚疾患の発症の研究（Ⅱ）資料開放...more</td>
</tr>
<tr>
<td>2012/08/01</td>
<td>精神分裂症と躁鬱症の合併影響是否影響患者對疾病認知程度？資料開放...more</td>
</tr>
</tbody>
</table>
SRDA’s Missions

Academic Surveys

Government Surveys

SRDA

Data Users
Data Acquisition

- Surveys
  - Implemented by CSR
  - Subsidized by NSC
  - Conducted by government

- Censuses

- Administrative district-level statistics
  Collaborated with Center for GIS at RCHSS
  - Demographic Statistics (since 1946)
  - Urban and Regional Development Statistics (since 1975)

- In-house processed data
  - Merged Manpower Utilization Survey Data
Data Processing

- Check for risk of disclosure, and decide whether to release or not
- For releasable data,
  - Clean data and delete sensitive information
  - Prepare required metadata
  - Create data files of different formats (e.g., SAS, SPSS, STATA)
- Downloadable files include raw data, questionnaire, codebook, user’s guide, etc.
Number of Datasets Archived by SRDA

Based on the statistics of May 2013,

- **Academic surveys**
  - Acquired data: 1,434
  - Released data: 1,135

- **Government surveys**
  - Acquired data: 863
  - Released data: 434

- **Censuses**
  - Acquired & released: 34
Data Access: Membership

- Two-year membership
  - Researchers or teachers at colleges, research institutes, or government organizations
- One-year membership
  - Graduate and undergraduate students, research assistants at colleges or research institutes
- Academia Sinica membership
  - Researchers at Academia Sinica
~ Free membership
  1,383 effective members in the end of this May
Data Access: Means of Access

- Most data can be downloaded from SRDA website directly
- Few data are of restricted access, and require review process
  - On-site access
  - Remote access
- Free access
- From May 2012 till the end of April 2013,
  - # downloads of non-restricted data is 13,578
  - # users of restricted data is 366
Urgent Topics to Develop Social Science Data Sharing in Taiwan
Urgent Topics with Data Sharing

- Even though both government and researchers recognize that open data is an emerging trend, there are some problems with the data sharing infrastructures for social science data
  - Government survey/census data
  - Government administrative records
  - Academic survey data
  - Ethical and law issues
Issues with Data Sharing (1/3)

- Problems with government survey/census data
  - Some gov’t organizations are conservative and reluctant to release their data to researchers
  - Some gov’t organizations did not preserve their data well
  - Some gov’t organizations are not aware how sensitive their data are
  - Metadata of survey data are often inadequate or incomplete
  - Different gov’t organizations have different data sharing policies (e.g., how and what to release)
  - There is no unified data center for the survey data collected by government
Issues with Data Sharing (2/3)

- Problems with government administrative records
  - Linkage with survey data are usually prohibited
  - Most administrative records are not accessible

- Problems with academic survey data
  - NSC’s policies on the subsidized survey projects are not mandatory
  - Some researchers view the data collected by them as private goods and hesitate to share the data with others
Issues with Data Sharing (3/3)

- Ethical and law issues
  - Researchers’ ignorance of ethical concerns when using data
  - Disincentive for data sharing induced by the Personal Information Protection Act
Thank you for your listening!

yurr@gate.sinica.edu.tw