Data Sources for Research: Discovery, Access and Use

John Southall
Bodleian Data Librarian
Subject Consultant for Economics, Sociology, Social Policy & Intervention

Catherine Scutt
Subject Consultant for Education

Michaelmas Term 2018
Research data at Oxford

– We’re considering data as a secondary resource
  • i.e. not data you are collecting yourself
  • Already gathered data available for re-use
  • Some disciplines see this as novel others as routine
  • Quantitative or qualitative

– Described as databases, datasets or archives
Understanding the language

– A database or dataset
  – A collection of information; numerical, textual etc.
  – Structured – excel spreadsheets, Word et al
  – Indexed
  – Derived from a particular research project, source or programme
  – Often downloaded or accessed
  – May be updated
  – Intended for academic use
Be aware of options

Apparently Wikipedia doesn’t count as a literature review

Use the best tools
Google Scholar Preferences

Library Links
(what's this?)

oxford

Find Library

e.g., Harvard

Show library access links for (choose up to three libraries):

- University of Oxford - Find it @ Oxford
- Oxford Brookes University - LinkSource@Brookes
- NOAA Oxford Marine Library - FullText via SwetsWise
- OXFORD BROOKES UNIVERSITY - ProQuest Fulltext

Online access to library subscriptions is usually restricted to patrons of that library.

Bibliography Manager

- Don't show any citation import links.
- Show links to import citations into

- RefWorks
- BibTeX
- EndNote
- RefMan
- WenXianWang

Save your preferences when finished and return to search.

Save Preferences
Databases

• Factiva: Business news sourced from Newspapers and business reports
• Proquest Historical newspapers
• Amadeus: Pan-European company information derived from annual reports
• Proquest Dissertations and Theses
Understanding the language

– A data archive
  • A large collection of datasets or even databases
  • Structured, indexed and documented
  • Curated by an organisation; cared for and future-proofed
  • Sometimes discipline based
  • Sometimes based around a specific institution or organisation
  • Content is ordered or downloaded
  • Static and unchanging
  • National archives – Institutional repositories
Data Archives and repositories

- ORA and ORA-Data: http://ox.libguides.com/home/researchers
- UK Data Service: http://ukdataservice.ac.uk/
- ICPSR: https://www.icpsr.umich.edu
- CESSDA (European Social Science Data Archives): http://cessda.net/
- Research Data Repository Registry: http://www.re3data.org/
UK Data Service

Welcome to the UK Data Service

Your resource for quality social research data

Survey and UK funded social science data.
Advice on finding data deposits on the UK Data Service site.

Can browse and read documentation before ordering data.

International macro data from Eurostat, OECD, UN, World Bank, IMF, IEA.
Open data portals

- Recent innovation
- Usually concentrate on statistics
- Making data publicly available via web browsers
- May seem to duplicate what’s in archives and databases
- Q – does it differ?
Statistical data portals

- UNdata: http://data.un.org/
- World Bank data: http://data.worldbank.org/
- U.S. Census: http://www.census.gov/
Statistical data portals

- Project specific?
- Oxford Poverty & Human Development Initiative:  
- The Oxford Roman Economy Database:  
  http://www.romaneconomy.ox.ac.uk/
- Charles Booth’s London:  
  https://booth.lse.ac.uk/
Exercises

—Time to look at some examples

• OECD iLibrary
• UK Data Service
Browse/Search

– The sheer amount being made available creates problems of discovery
– Interested in seeing what’s available
  • Browse for them
– Increasingly datasets are cited
  • E.g. Iterate – political terrorism database, British Household Panel Survey, Bloomberg Professional
  • Search for them using public and university search engines
Search/Browse

- Search engines include data alongside books, journal articles etc.
  - SOLO, Databases A-Z
- Encouraged by publishers like Proquest
  - Creating bibliographic databases
- Often available via subscription
  - Asked for payment? Enter via SOLO etc with single-sign-on
  - Additional requests to sign in?
What to use?

– SOLO
  • Integrates with searches for traditional resources
– Databases A-Z
  • Tighter focus on e-resources
– Subject Libguides
  • Curated by your subject librarian
– Bodleian Data Library
  • Additional content on data management, analysis
– Same resources but different shop windows
Evidence Based Social Intervention: Getting Started

WORK IN PROGRESS:

Some of the links in this LibGuide are broken due to a system-wide change. We are aware of this, and are fixing the links. In the meantime, you will find all of the Bodleian Libraries subscription resources listed at A-Z Databases and SOLO.

How to find

- Articles & Reports
- E-Books & Reference
- Theses & Dissertations
- News Sources
- Data & Statistics
- Managing research data

New and Trial resources

See the latest print and digital resources in Social Policy & Social Intervention, purchased by the Bodleian Social Sciences Library.

Browsing Guides & Collection Policies

- Subject Browsing Guide Social Policy and Social Work Books
- Subject Browsing Guide Sociology Books
- Collection Policy for Social Policy and Social Work
- Collection Policy for Sociology
- Guide: Remote access

Subject Guide

John Southall

Email Me

Contact:
Bodleian Social Science Library
Manor Road, Oxford, OX1 3UQ
+44 (0)1865 205980

Subjects:
Economics, Economics - General
Data and Statistics for Social Sciences: Media, Information and Communication

**Subjects:** Development Studies, Economics, Politics and International Relations, Social Policy and Social Work, Sociology

---

**Pew Research Center**

**Internet, Science & Tech**

The Pew Research Center's Internet Project offers access to raw data sets from their research which are made available as single compressed archive files (.zip files).

All uses of this data should reference the Pew Research Center as the source of the data and acknowledge that the Pew Research Center bears no responsibility for interpretations presented or conclusions reached based on analysis of the data. Pew Research is interested in learning about other ways that scholars use our data. If you publish something based on our data, please let us know by sending us an email. Individuals will need to fill out a brief registration before downloading data.

Questions concerning the data sets may be directed to the Pew Research Center.

---

**Online resources**

**BBC Monitoring International Reports**

Full text from 1979-present available via Lexis Library. Provides access to broadcasts from 120 countries, and emphasizes information from countries for which other news sources are not readily available.

**FBI - Foreign Broadcast Information Service Daily Reports 1975-1995**

Selected broadcasts, newspaper articles and government statements from around the world, translated into English and issued by an agency of the US government.

**World News Connection Archive 1895-2013**

Newspaper articles, broadcast transcripts and datelines from around the world, translated into English by the US government.

---

**UK Data Service**

**Information and communication**

Information and communication technologies are reshaping the world, transforming the way in which we communicate, govern, work, manage crises, do business and spend our free time. These changes are, in turn, driving new policy development to address the societal impacts of digital technologies. The UK Data Service holds data on a wide range of ICT topics, such as mobile communication, telework, social media, radio, television, Internet use, media management, infrastructure security and trust. This key research area helps us to understand how the fast evolving world of digital technologies impacts people's lives and communities, changes the way we interact, affects our cultural experiences and could shape our future society and economy.

Go to Discover to search and browse data collections, series, case studies, and ESRC outputs on Media, communication and language.

---

**VIIPT database**

World Association of Newspapers and News Publishers

**World Press Trends (VIIPT) Database**

A dynamic continuously updated database that allows users to produce their own, customized reports, which can be exported to PDF or to Excel format. Available in PDF format via Weblearn.

**European Audiovisual Observatory Yearbook**

The Yearbook provides a wide range of statistical data on the four main industry branches of the audiovisual sector in Europe: the film industry, television, video/CVOD, non-dramatic audiovisual services. The Yearbook provides three sub-sets of data, delivered as it becomes available: Detailed tables and figures; country profiles; key trends reports. The figures are available in Excel tables and graphs, and the text is available in PDF version.

Username/password are required for this resource. See Weblearn for details.
Data Archiving and Sharing

Find out about securely storing and sharing your research data.

More...

News: UKDS: Online seminar: Data sources for research students - Released 8th September by UKDS

Data Analysis
A selection of tools for making best use of the data you have collected for your research.

Data Obligations
What you need to know about using data at Oxford. Your responsibilities and how to comply with data use terms and conditions.

Research Data Management
Find out more about effectively organising data as part of your research.
Exercise

- Introduction complete
  - Questions?
- For ten minutes try some searches
  - Identify three resources relevant to your discipline – why?
  - Say how you found them
  - Are they public or subscription based
Caveats & annoyances

– Search engines collate available resources and present them in a unified way
– Are they seen as ‘valid data’ in your discipline
– How you access, view and use them can vary
  • Be clear what is being made available
  • Be clear in what format
  • Are there usage restrictions
– Extra hurdles
  • Time delays
Common examples

— Registration suggested or required
  • Subscription database environment
  • Common with external data archives
  • Providing authentication

— Restrictions
  • Sensitivity dealt with by making access harder
  • Archived data discoverable but not accessible
Depositing Data

— Becoming the norm
  • Growing expectations by funders and publishers
  • University policy
  • Data and contextual documentation
  • An aid in preserving your own data
  • Become part of the research record

— ORA
  • ORA-Data
  • Research Data Oxford
More Caveats

– Restrictions on how you use it
  • For analysis only - cannot be quoted
  • Cannot be shared
  • Supplied tabulations cannot be altered
– Make sure you are aware of these!
– May need to analyse data within their browser/interface
– Data may be available from multiple sources but with differing documentation and support
Restricted/Microdata

– Very sensitive or includes personal identifiers
  • E.g. economic data, population data
– Only available pending application approval
  • From e.g. Eurostat, Statistics Canada, UKDS
    Secure lab, ICPSR
  • Open data portals will not supply it
– Support making applications via Bodleian Data Library
Restrictions on how you use it

• For analysis only cannot be quoted
• Cannot be shared
• Supplied tabulations cannot be altered

Make sure you are aware of these!

May need to analyse data within their browser
More Examples

– Restrictions on how you use it

• For analysis only cannot be quoted
• Cannot be shared
• Supplied tabulations cannot be altered

– Make sure you are aware of these!
– May need to analyse data within their browser
Thank you

– Any final questions or observations?
– Follow up this session with your own subject librarian (or any other library specialist)
  • See libguides
– Data Librarian: John Southall
  john.southall@bodleian.ox.ac.uk