From the SelectedWorks of Jennifer Allison

Spring 2015

Research with Non-Legal Databases: Class Handout

Jennifer Allison

Available at: https://works.bepress.com/jennifer_allison/53/

Research Class
Beyond Lexis and Westlaw: Non-Legal Databases

April 10, 2015
Harvard Law School Library
Jennifer Allison
Librarian for Foreign, Comparative, and International Law
jallison@law.harvard.edu

In this class, we are going to explore doing some initial research on a multi-disciplinary topic using Harvard's non-legal databases. It will involve the following tasks:

1. Brainstorming the topic
2. Using Google Scholar for research
3. Looking for articles in the ScienceDirect database
4. Looking for articles in the PsycINFO database

What is the research question?

How have psychological and sociological factors influenced our feelings about the regulation of data protection and privacy for financial and medical data?

Task #1: Brainstorm

Journals covering which academic subjects may include information that would be helpful in researching this topic? List a few below.

- Psychology
- Sociology
- Data privacy and protection (maybe IT)
- Law
- Political Science
- Medicine (specifically discussing medical records)
- Finance (specifically discussing financial data)

Task #2: Google Scholar

Because our research is multi-disciplinary, we want to start by casting the widest net possible. Google Scholar is an excellent source for us to use for this. It's easy to search, and it includes a lot of journals, books, working papers, etc.

When we use Google Scholar, we want to make sure that when we click links we are taken directly to the Harvard-subscribed databases.

1. If you are **on the Harvard network (on campus)**, you don't have to do anything except go directly to http://scholar.google.com.
2. If you are **not on the Harvard network (off campus)**, you have to use the special Harvard version of Google Scholar:
    a. Go to **Hollis+** = http://hollis.harvard.edu.
    b. Sign in with your **HUID and password** (use the link in the top right corner of the screen).
    c. Search for **Google Scholar**. In the first record, click the **View Online** link.

Now you're ready to do some searching. One option, of course, is to type a bunch of keywords in the box, like this:

psychological sociological factors data protection privacy

If we enter all of those words in the search box, we get more than 100,000 results. Theoretically, everything in this results list has all of these words in it, sorted by relevance.

From here, we can either start going through the results, clicking on links that interest us, or we can limit the results. A good limiter may be the date option. It might be good to start with more recent articles. So click the **Since 2011** link on the left.

Let's examine a few of these results in a little more depth. We can use this process to get a survey of the lay of the land, so to speak, for this type of scholarship. Use a chart like this one to record some key information about sources that look like they'll be helpful for your research.

Title	Author	Year	Publication Name	Database	Additional information (citation, summary, keywords, etc.)
Privacy and Human Behavior in the Age of Information	Acquisti, Brandimarte, Lowenstein	2015	Science	Science	v. 347, no. 6221, p. 509-514, January 30, 2015. Summarizes and draws connections between diverst streams of empirical research on privacy behavior. *Part of a special issue: The End of Privacy*
Information Privacy Research: An Interdisciplinary Review	Smith, Dinev, Xu	2011	MIS Quarterly	ACM Digital Library	v. 35, is. 4, Dec. 2011, p. 989-1016. This paper provides an interdisciplinary review of privacy-related research in order to enable a more cohesive treatment. With a sample of 320 privacy articles and 128 books and book sections, we classify previous literature in two ways: (1) using an ethics-based nomenclature of normative, purely descriptive, and empirically descriptive, and (2) based on their level of analysis: individual, group, organizational, and societal.

Title	Author	Year	Publication Name	Database	Additional information (citation, summary, keywords, etc.)
Research Note—Effects of Individual Self-Protection, Industry Self-Regulation, and Government Regulation on Privacy Concerns: A Study of Location-Based Services	Xu, Teo, Tan, Agarwal	2012	Information Systems Research	Informs PubsOnLine	v. 23, is. 4, 2012, pp. 1342-1363. This study seeks to clarify the nature of control in the context of information privacy to generate insights into the effects of different privacy assurance approaches on context-specific concerns for information privacy. *Focuses on Singapore*
What Is Privacy Worth?	Acquisti, John, Loewenstein	2013	The Journal of Legal Studies	JStor	v. 42, No. 2, June 2013. Understanding the value that individuals assign to the protection of their personal data is of great importance for business, law, and public policy. We use a field experiment informed by behavioral economics and decision research to investigate individual privacy valuations and find evidence of endowment and order effects.
Privacy, Security, and Trust in Cloud Computing (Book chapter)	Pearson	2012	Privacy and Security for Cloud Computing (Book)	Springer	Book. We assess how security, trust and privacy issues occur in the context of cloud computing and discuss ways in which they may be addressed. *Note: Harvard has this as an eBook.*
Privacy Self-Management and the Consent Dilemma	Solove	2013	Harvard Law Review	HeinOnline	v. 126, p. 188. Introduction to a symposium issue on Privacy and Technology. *Note: Should read all the articles in this issue!*
"To disclose or not to disclose, that is the question": A structural equation modeling approach to communication privacy management in e-health	Seung-A	2012	Computers in Human Behavior	ScienceDirect	v. 28, is. 1, Jan. 2012, p. 69-77 Driven by communication privacy management (CPM) theory and regulatory focus theory (RFT), this study examined important antecedents of information withholding and truthful disclosure in the novel context of e-health communication. *Notes: want to search this journal and database directly for more articles, want to read more about CPM.*
Theories in online information privacy research: A critical review and an integrated framework	Li	2012	Decisions Support Systems	ScienceDirect	v. 54, is. 1, Dec. 2012, p. 471-481. This study reviews fifteen established theories in online information privacy research and recognizes the primary contributions and connections of the theories. *Notes: looks like a good broad overview of privacy theory (such as "privacy calculus"), could be helpful in developing additional searches.*

Issues with this search:

1. It looks like Google Scholar may have returned articles I didn't want – for example, articles that had terms like "social network" (probably instead of "sociological") and "mental" (probably instead of "psychogical"). Our original question didn't contemplate social networks specifically.
2. Too many results did not appear, on the surface, to have much to do with data privacy.

Task #3: ScienceDirect

Two articles that we entered into our article spreadsheet above came from the same database: **ScienceDirect**. This indicates that it might be a good option for more in-depth research.

Let's go there: http://www.sciencedirect.com.ezp-prod1.hul.harvard.edu/.

At this point, we have seen some problems with a broad all-text search with a lot of keywords. So let's try a more targeted search this time.

1. Click **Advanced Search**.
2. Enter **"data privacy"** in quotes in the search box, and in the corresponding drop-down, select **Abstract**.
3. Click **Search**.

This is a very manageable number of search results. We know that these articles discuss data privacy in some depth, because we required that the article have that phrase in its abstract. Let's chart out a few of these articles also.

Title	Author	Year	Publication Name	Database	Additional information (citation, summary, keywords, etc.)
A comparison of Proposed Legislative Data Privacy Protections in the United States	Barclay	2013	Computer Law & Security Review	ScienceDirect	v. 29, is. 4, Aug. 2013, p 359-367 The leading model for data privacy protection is the 1980 Organization for Economic Cooperation and Development (OECD) Guidelines on the Protection of Privacy and Transborder Flows of Personal Data. This article examines two leading legislative privacy proposals in the context of the OECD principles. This examination concludes that, although the proposals do not provide sufficient comprehensive privacy protections, they do fill significant gaps in current U.S. privacy laws. *Notes: When you look at this in ScienceDirect, click the "Recommended Articles" link on the right side to view a list of other articles that appear to be really relevant.*
A standardised graphic method for describing data privacy frameworks in primary care research using a flexible zone model	Kuchinke et al.	2014	International Journal of Medical Informatics	ScienceDirect	v. 83, is. 12, Dec. 2014, p. 941-957. Discusses development of a model that allows analysis of data privacy and confidentiality issues for research with patient data in a structured way and provides a framework to specify a privacy compliant data flow, to communicate privacy requirements and to identify weak points for an adequate implementation of data privacy. *Notes: EU focus, not US.*
Protecting health data privacy while using residence-based environment and demographic data	Rodgers et al.	2012	Health & Place	ScienceDirect	v. 18, is. 2, March 2012, p. 209-217. Spatial analyses of environment and health data are often made using point address data, despite the risk of identity disclosure. We describe how geospatial environment and non-spatial health data can be linked anonymously, thereby maintaining geoprivacy. *Notes: Geoprivacy! That may be a sub-aspect of privacy that is especially relevant in medical data. Does the government regulate that? Should they? Do people even care about that? Should they?*

More emerging issues:

1. Is our topic too broad? We've seen general articles but we've seen a lot that focus on medical data. Will the research be easier if we limit just to that? Or have too many people already written on this topic?
2. *Computer Law & Security Review* could be an especially helpful journal for this research. We may want to browse it to see the themes it has covered over the past few years.

We're still really interested in learning about the intersection of psycology and data privacy. How do people feel about data protection? Does that influence the creation of data protection laws?

Now, let's take a look at other databases in Harvard's collection to see if there's one that focuses on psychology.

Task #4: PsycINFO

Because at this point we're not sure what Harvard has in terms of psychology-related databases, let's browse a list of some databases that are popular and see if any match our criteria.

1. Go back to **Hollis+**, and click the **Databases** tab at the top of the screen.
2. View the list of databases in the **selected major resources** menu.
3. Select **PsycINFO**, and click **Go**.

Per its website, PsycINFO is "an expansive abstracting and indexing database with more than 3 million records devoted to peer-reviewed literature in the behavioral sciences and mental health, making it an ideal discovery and linking tool for scholarly research in a host of disciplines." This sounds good if we want to focus on issues of psychology and and data privacy!

In PsycINFO, which is one of the databases in a database collection called **EbscoHost**, let's try the same search we did earlier – "data privacy" in the abstract.

This returns 30 results – a very manageable number. Let's chart a few of these out also.

Title	Author	Year	Publication Name	Database	Additional information (citation, summary, keywords, etc.)
Human-oriented design of secure machine-to-machine communication system for e-healthcare society	Saleem et al.	2014	Computers in Human Behavior	PsycINFO	In this paper, we propose a Machine to Machine (M2M) Low Cost and Secure (LCS) communication system for e-Healthcare society. The system is designed to take into consideration the psychological issues related to all actors in the e-Healthcare society such as: stress due to high workload, anxiety, and loneliness. The system is capable of performing most of the tasks in an autonomous and intelligent manner, which minimizes the workload of medical staffs, and consequently minimizes the associated psychological stress and improves the quality of patient care as well as the system performance. *Notes: Not really about patient psychological issues, but health care workers. Is this regulated anywhere? Could be an interesting twist on this topic.*
Privacy-preserving data integration in public health surveillance	Hu	2014	PhD Dissertation?	PsycINFO	We propose new privacy preserving data integration protocols for public health surveillance, identify a set of privacy preserving data integration patterns, and propose a supporting framework that combines a methodology and architecture with which to implement these protocols in practice. Our work is validated with two real world case studies that were developed in partnership with two different public health surveillance organizations. *Notes: What is this? If it's a PhD dissertation, it should have a good bibliography.*

Title	Author	Year	Publication Name	Database	Additional information (citation, summary, keywords, etc.)
A practical framework for data management processes and their evaluation in population-based medical registries	Sariyar et al.	2013	Informatics for Health & Social Care	PsycINFO	v. 32, no. 2, Mar. 2013, p. 104-119. We present a framework for data management processes in population-based medical registries. The standardisation of data management processes in medical registries is important to guarantee high quality of the registered data, to enhance the realisation of purposes, to increase efficiency and to enable comparisons between registries. Our framework is destined to show how one central impediment for such standardisations—lack of practicality—can be addressed on scientific grounds. *Notes: Another one about geography-specific privacy issues in the medical field.*
Legal and Ethical concerns of Collecting Data Online	Sturgill and Jongsuwanwattana	2007	Handbook of Research on Electronic Surveys and Measurements (note this is a book)	PsycINFO	International data collection offers the possibility of greater external validity for studies. However, using the Internet to collect information internationally brings concerns related to data privacy for respondents. Nations vary on their privacy protections for citizens' Internet use. Beliefs about data privacy, official policies notwithstanding, can also affect respondents and response rates. This chapter explores these issues and lists issues that should concern the international survey researcher. *Notes: Good international survey? Gutman and Widener both have this; Hollis # 009762391.*

More emerging issues:

1. Are there even articles on data privacy and psychological issues for financial data? Is "psychological issues" even the right term or concept? Maybe I should be looking in more financially-oriented databases for those types of articles?
    a. One option for this may be to use the **Ebsco Business Source Complete** database. This is also available through the selected major resources menu on the Hollis Databases page.
2. In the last research task, we found both a PhD dissertation and a book. I know to look up a book in Hollis, but how do I find PhD dissertations?
    a. Harvard has many databases that include PhD dissertations, including **ProQuest Dissertations and Theses** - http://id.lib.harvard.edu/aleph/007300670/catalog. Remember, PhD dissertations can be a great source of sources – they have to list all of the books and articles the author used.

Ongoing Tasks and Suggestions

1. While you're researching, you will want to keep a running list of keywords, and try multiple searches using various combinations of these keywords. We really stuck to "data privacy" in these examples, but you should not limit yourself like this.
2. You can limit searches to abstracts, like we've done in the two examples here. But some articles don't have abstracts. So also try limiting searches to article titles. Those are two ways you can help keep the results lists smaller and probably better limited to only relevant sources. However, if you use those types of search limits, you may miss some articles. So full-text searching can also be helpful.
3. During your searches, pay attention if you start to see the same authors over and over. For example, we saw two articles on data privacy by Alessandor Acquisti during the Google Scholar portion of the research. Check out website at http://www.heinz.cmu.edu/~acquisti/. You'll see that he has written a lot about data privacy and security, and even maintains a database with sources on the economics of privacy: http://www.heinz.cmu.edu/~acquisti/economics-privacy.htm. Take advantage of research that has already been done by smart people!
4. Take a look at later articles that cited a relevant article that you've read. This can be done in Google Scholar by clicking the **"Cited by (number)"** link below the article excerpt.
5. Eventually you're going to start reading sources, and you are probably going to figure out your topic is too broad. This is okay! You want to start broad and find a lot of sources. Once you start reading them and educating yourself, it will be easier to narrow your topic.
6. Explore, explore, explore. You're not going to break anything by trying out different kinds of searches and databases. And don't click too fast! Read what's on the screen first before you move along to hyperlinked material.
7. Use library research guides. Librarians have created guides that explain how to use their specific databases. You can browse the research guides created by all of the Harvard Libraries by topic here: http://guides.library.harvard.edu/by-subject.
8. If you feel like you're struggling, contact a reference librarian for some judgment-free assistance. The law librarians are more than happy to help you, either at the reference desk or in a private research consultation meeting. Visit http://asklib.law.harvard.edu to find out how to get in touch with us.