

Engaged Librarian Forum

Data Case Studies

September 18, 2014

Case studies authored by:

Juleah Swanson, Acquisitions Librarian for Electronic Resources
Ohio State University Libraries

Case Studies: Patrons and Data

We have compiled five case studies involving student and faculty requests, use, or engagement with data in their research or study.

Cases are based on real situations that have been addressed by a number of librarians here at the OSU Libraries. You may recognize some details of these cases as your own, or you may be familiar with similar situations.

Events have been dramatized and all names have been changed from patron names to the names of former child stars of the 1990s and early 2000s who have since grown up and went on to college.

Bonus points for anyone who can identify what made the child star famous and where the child star went to college ☺

Instructions:

- **Break up into five groups.**
- **Read through your assigned case study.**
- **Discuss the case study questions as a group.**
- **Jolt down notes or a summary of your answers on the flipchart**
- **Report out findings to the room**

Case Study Questions

- Summarize some of the nuances within this case study
- What resources or services exist within the libraries? Does this situation need to be referred elsewhere? If so, who?
- How might you assist the patron?
- What can we do better as a library to support this case study?

Case Study #1

Under-caffeinated graduate research assistant Jonathan Taylor Thomas has been asked by his advisor in the Department of Communications to work on a special project to identify *all* the articles faculty members within Communications cite in their publications. He was advised to use Web of Science to look up the Department of Communications faculty publications. His advisor also taught him how to capture each Web of Science result as an html page.

Over the summer, Jonathan spent hours manually looking up thousands of citations in Web of Science and saving thousands of html pages containing the data he needed.

With two weeks left in the summer Jonathan's advisor took off on a well-deserved vacation to hike around the remote Isle Royale National Park with no access to the internet, cell service, or other comforts of modern civilization.

Jonathan is nearly complete with his research on Web of Science, but now needs to parse and compile all of the data from the saved html pages into a format that can be analyzed. He has no idea how to do this and doesn't know who to ask.

In a moment of desperation, Jonathan calls up the OSU Libraries Reference Desk because he figures, librarians are smart people, they can help.

You pick up the phone when Jonathan calls.

Case Study Questions

- Summarize some of the nuances within this case study
- What resources or services exist within the libraries? Does this situation need to be referred elsewhere? If so, who?
- How might you assist the patron?
- What can we do better as a library to support this case study?

Case Study #2

Assistant Professor Haley Joel Osmond is working to build a sophisticated article recommendation and analytics system built on top of rich article metadata. It would be something like a read-alike that would present something similar to a user:

“If you liked this article on Carbon Nanotube-Based Cells for Yield Analysis maybe you’ll like this article on System Level Benchmarking with Yield-Enhanced Standard Cell Library for Carbon Nanotube VLSI Circuits.”

Awesome.

Assistant Professor Osmond needs metadata from tens of thousands of articles to further refine and develop his software. He has heard from colleagues at other universities working on similar research that they have been using metadata from the Association for Computing Machinery (ACM).

Assistant Professor Osmond is smart, so he knows he could easily create a script that would likely be undetectable to the security software that ACM has in place and crawl their website for the article metadata.

But, he also knows that ACM likely has all this metadata in a retrievable data file on their back end, which they could easily send along to him.

Osmund is told by a colleague that he needs to talk to the Libraries in order to work with ACM to acquire the metadata he needs. So he approaches you.

You speak briefly to ACM about the request for a file of ACM article metadata. As a not-for-profit, association publisher, they are more than willing to provide the data to the faculty member for free!

ACM just requests that the faculty member sign a research agreement that mostly stipulates how the data is to be accessed, used, and distributed, but also includes the following clause requesting the faculty member grant ACM a royalty-free, non-exclusive license for the software that is developed based on the ACM article metadata:

“Provide the ACM a copy of any software developed in the course of the research that relates to the management and analysis of the ACM-supplied Material, with a royalty-free, non-exclusive, non-transferrable perpetual license to use it as is, or with modifications, solely in the context of ACM’s Digital Library service.”

Assistant Professor Osmund is open to the idea of a letting ACM use the software he is developing with their metadata, but is unsure about whether he can sign the agreement. He approaches you again for consultation.

Case Study Questions

- Summarize some of the nuances within this case study
- What resources or services exist within the libraries? Does this situation need to be referred elsewhere? If so, who?
- How might you assist the patron?
- What can we do better as a library to support this case study?

Case Study #3

Faculty member Danica McKellar manages a small lab in the College of Veterinary Medicine. Teaching and conducting research at OSU is a second career, as she spent many years in industry managing a much larger lab with many long-term lab assistants.

She has recently become frustrated with the high turnover of graduate assistants in her lab and the lack of basic knowledge and skills that her new graduate assistants have on managing data. Files are poorly organized. Data is being saved all over the place, in secure and unsecure locations. And her students do not seem to understand the importance of formatting, saving, and storing the data for future use and long-term preservation.

She wants to introduce basic data management training into her courses to better prepare her students for current or future work in the lab or field.

You have a meeting scheduled with Danica to discuss her current lab work and research. In this meeting she brings up her frustration and concern over the lack of basic data management skills that she sees among her students.

Case Study Questions

- Summarize some of the nuances within this case study
- What resources or services exist within the libraries? Does this situation need to be referred elsewhere? If so, who?
- How might you assist the patron?
- What can we do better as a library to support this case study?

Case Study #4

Enterprising graduate student Andrea Barber needs to text mine millions of news articles from the past twenty years for her dissertation in Political Science. She is first looking for specific word pairings that exist in news articles and then is looking at the greater context around those word pairings to identify themes and tone, in order to determine changes over time.

Her ideal source for accessing the millions of news articles is the LexisNexis Academic database because it has over 3,000 global newspapers and she was easily able to create a Python script for LexisNexis that crawled for her word pairings and retrieved the data she was looking for.

Andrea launches her Python script, sits back and watches her dissertation data flow in. She is on her way to PhD success and dissertation fame. Maybe she'll even get a book deal from a university press like Harvard or NYU.

Thirty minutes in to running her Python script, the script fails. Or at least it stops running and she's not quite sure why. Andrea sets it aside to try again tomorrow.

The next day, Andrea receives a stern e-mail from the OCIO's Senior Security Engineer informing her that per the OSU Libraries license agreement with LexisNexis, the scripted search she was conducting is not allowable, and has resulted in a temporary suspension of LexisNexis Academic for the entire University.

Andrea writes back:

"I apologize - I didn't know that scripted searches were not allowed. I will stop it immediately and remove the scripts from the machines.

That said, I wonder if I could have the email or phone number of your contact at LexisNexis? The thing is, I need to text-mine millions of news articles in order to do my dissertation. It would be impossible to download everything manually (I need to perform over 1.4 million searches on the LexisNexis database). So I'd like to talk to them, explain my situation, and see if they can make an exception in my case."

The OCIO Senior Security Engineer knows this is something the Libraries should handle, but doesn't know who. He forwards the e-mail along to you for assistance.

Case Study Questions

- Summarize some of the nuances within this case study
- What resources or services exist within the libraries? Does this situation need to be referred elsewhere? If so, who?
- How might you assist the patron?

- What can we do better as a library to support this case study?

Case Study #5

It's Monday morning, it's raining, and you've only had one cup of coffee instead of your requisite two. Things are moving slowly as you enter your office and fire up the computer to check your e-mail.

Unread Message #1

From: watson.754@osu.edu

Sent: Sunday 11:59 PM

Subject: NAV data

Hi I'm a PhD candidate in Political Science and I'm looking for price and net asset value (NAV) data for close-ended, delisted exchange-traded funds (ETFs). Does the library have this data?

Thanks!

Emma Watson

After some swift flexing of your librarian-ing skills you find that this data is available in a module of Global Financial Data that OSU Libraries does not currently subscribe to. Last time you check it was about \$10,000 to add on the module.

You also find that the price data is available through the Datastream terminal at Fisher, but not the net asset value data. The Net Asset Value data is available through Morningstar Direct, another terminal based product that OSU Libraries does not currently subscribe to.

You add this request to your to-do list for the morning and move on to the next e-mail.

Unread Message #2

From: watson.754@osu.edu

Sent: Monday 12:02 AM

Subject: Price index for recyclables

Hi Again!

I'm also working with a faculty member at Fischer so I also need the price index listing market prices for common recycled commodities, such as recycled aluminum. I'd like current and historical prices.

Thanks!

Emma

Again, you flex your librarian-ing-skills but this time come up empty-handed. Nothing you find perfectly matches this request. There's the Global Financial Data database that OSU subscribes to which has dozens of commodity indexes, but nothing specifically on recycled materials. There's the USGS Mineral Industry Survey available online which contains buying prices for scrap metal, and there

is the American Metals Market, which OSU does not subscribe, and you're not sure whether AMM even offers an institutional subscription.

You also put this e-mail on your to-do list to investigate further.

Unread Message #3

From: watson.754@osu.edu

Sent: Monday 12:09 AM

Subject: Foreign Direct Investment by State

*Oh...and I also want Foreign Direct Investment (FDI) data in the US at the state level, state by state. This one is time sensitive. I need the data by the end of the week. Where can I find this?
Emma*

Everyone seems to want Foreign Direct Investment data, so you have a good handle on where this data might exist. Each state's commerce page has data available on FDI, but that would require the student to look at the commerce sites for all 50 states. There's also the resource ForeignDirectInvestment.com that has some freely available datasets, though more enhanced reports and datasets requires purchase.

You also put this e-mail on your to-do list to investigate and to reply.

Case Study Questions

- Summarize some of the nuances within this case study
- What resources or services exist within the libraries? Does this situation need to be referred elsewhere? If so, who?
- How might you assist the patron?
- What can we do better as a library to support this case study?