DATA STORIES: USING NARRATIVES TO REFLECT ON A DATA PURCHASE PILOT PROGRAM

Anita Foster, Electronic Resources Officer
Gene Springs, Collections Strategist
The Ohio State University Libraries
NASIG 32nd Annual Conference
June 9, 2017
No vendors were harmed in the making of this presentation.
PROLOGUE
Translational Data Discovery Theme in 2014
Launch of Translational Data Analytics Program in 2015
Interdisciplinarity across campus
Research Commons space envisioned
CLUSTER HIRE OF 2014

3 intended positions related to Data Services

- Research Data Management Librarian (Hired & Started 2/14)
- Digital Humanities Librarian (Closed 2/14; Re-opened 3/15, Closed 7/15; Re-opened 1/16, Hired & Started 7/16)
- GIS Librarian (Closed 2/14, Re-posted 3/15, Not filled)
Cluster Hire of 2014

- 3 Subject Librarians Hired (Started 1/14)
  - Head, Fine Arts Library
  - Head, Architecture Library
  - Business Librarian (Gene!)

- Collection Strategist (Gene’s predecessor) Hired (Started 1/14)
Purchasing Data

- GeoLytics requested as a New Resource Proposal in FY15
  - DVDs of geospatial data
- National Association of Insurance Commissioners (NAIC) data
  - Purchased historical annual datasets via flash drive for research group at Fisher College of Business
- Business-related data purchased by/for Fisher College of Business
  - WRDS platform, Capital IQ, Bloomberg, DataStream
Research Commons

- Services for Research Commons begin in Summer 2014
- Physical space opens in January 2016
  - Collaborative space for graduate students and faculty
  - Computer lab with statistical data software
  - Data visualization lab
  - Expansion of services, campus launch
- Became the point of need space for data services
New Vice Provost and Director of Libraries, Damon Jaggars arrives 2/1/16

Series of “listening tours” with campus and Libraries communities

Data is a repeated refrain from all communities
RESEARCH DATA IN THE LIBRARY TASK FORCE

- Long-gestating task force focused on research data services charged and deployed in June 2016
- Charge:
  - Conduct an environmental scan to reflect current OSUL services and practices that support research data
  - Include a statement on the challenges and complexities which currently prevent us from offering support
  - Recommend solutions and next steps for delivering a robust suite of services around research data
- Representation from: Research Commons, Subject Librarians, Archives, Acquisitions, Institutional Repository, Health Sciences Library, (Collection Strategy)
STRATEGIC COLLECTION AREAS OF FY17

- 3 Collection Areas prioritized for one-time funds in FY17
  - Streaming Media (continuation from FY16)
  - Business Resources
- DATA
- Pilot Programs
FY17 DATA PILOT

- Parameters – none!
- Budget – no specifics!
- Process & workflow – we’ll learn!
- Multi-year, pay in advance agreements preferred!
- Communication – inclusive!
OTHER FACTORS

- New Resource Proposal Process
- Submitted Research Data in the Library Task Force report (11/16)...
- ...Re-envisioned Research Data in the Library Task Force (1/17)
- Re-submitted Research Data in the Library Task Force report (4/17)...
- ...Creation of Data Working Group (5/17)
AND SO OUR TALES BEGIN
RESOURCES INVESTIGATED

- Purchased: 7
- Not Purchased: 4
- Deferred: 2
INFOGROUP

- Purchased Historical Business Data and Historical US Residential Data, full file + 4 years of updates
- Initial Library Contact/Communication: 2015
- Licensing Highlights:
  - Auditing language
  - Library publicizing resource
- Lesson Learned: Know what you’re buying

- March 2015
- May 2016
- August 2016
- October 2016
- December 2016
- February 2017
- Fall Semester 2017
Request for Gallup Analytics Data Access at The Ohio State University

Gallup Analytics is a cutting edge, innovative data tool that could advance research and teaching efforts at The Ohio State University.

Gallup has decades of experience producing high-quality, reliable, global population-representative social science data. Ohio State faculty, graduate students, and researchers could use these data to advance science by testing research questions that have heretofore remained unexamined.

For example, the incidence rate of same-sex couples in the population is low. Jones and Gates (2015) examined Gallup data from June 27, 2015, when same-sex marriage was legalized in all 50 states, to November 23, 2015, and estimated that there were 0.220,000 American married same-sex couples (or 0.10% of the 20 million in same-sex cohabiting unions). Thus any scholar wanted to examine the health and well-being of same-sex couples is going to be hard-pressed to find data that will allow them to have an adequate sample-size to conduct analyses. Because Gallup conducts a daily survey asking 1000 US adults about various political, economic, and wellbeing topics and is constantly refreshing the Gallup Analytics portal, scholars at Ohio State could use the cutting edge Gallup data with their large, population representative samples, to examine the wellbeing of individuals in same-sex marriage and compare it before and after the US-wide legalization of same-sex marriage. Further, because the Gallup Analytics portal is global and represents more than 90% of the world's population, it is also possible to do global comparisons. No other data source like this exists.

Usage of the Gallup Analytics Data

Research
As was stated above, there are many research questions that can be answered with the Gallup Analytics data, and importantly, the data would be able to answer research questions in several disciplines across the university, from sociology and communication to political science and public health. Professor Paul Behr of Sociology says "I am studying employment trends among prisoners and am always looking for new ways to examine local contexts including economic conditions at the MSA level. This data would open new doors in terms of exploring the impact of broader economic conditions." Professor Chris Gelch of Political Science says "These data would enable my research on American public opinion and foreign policy." Associate Professor Claire Kamps Dush of Human Sciences would use the data to test research questions related to same-sex marriage in the US. Graduate Student Margaret Hansen of Political Science highlighted that these data would be useful in her dissertation that focuses on Political Independence.

Teaching
These data could be highly useful in teaching both graduate and undergraduate students at Ohio State. As Associate Professor Donna Bebbitt-Zeher mentions "I think access to these data would be useful for several purposes. I could see undergraduates using this for independent research projects, I also envision faculty using this for their own research, or for incorporating into the classroom when teaching research methodology, statistics, or on any of the subtopics covered in the data." Associate Professor Gerd Koscil says "I typically teach Public Opinion and Communication for my school at the undergraduate and graduate levels. I also teach the Survey Practicum, which is one of two required courses to receive our Graduate Interdisciplinary Specialization in Survey Research Certificate, which is open to students all over campus who complete the requirements. I often direct students to The Roger Center for ICPSR to find data or questions or whatever. This sounds like an interesting resource." Thus, undergraduate and graduate students could analyze these data as part of their coursework, and importantly, for honors theses, masters theses, and dissertation research.

Conclusion
The Gallup Analytics data is a state of the art data source that could advance both research and teaching efforts at The Ohio State University. These data are also an excellent example of "Big Data" which is very popular in science right now and is related to the Translational Data Analytics Discovery Theme at Ohio State. Our peers at Indiana and Northwestern Universities and aspirational (Carnegie, Stanford, and Princeton Universities) institutions already have access. We request Ohio State acquire these data.

Signatures

[Signatures]

1. Professor, Environmental Sociology, College of Food, Agricultural and Environmental Sciences
2. Associate Professor, Sociology, College of Arts and Sciences
3. Associate Professor, Human Sciences, College of Education and Human Ecology
4. Associate Professor, Sociology, College of Arts and Sciences
5. Director of the Institute for Population Research and Professor, Sociology, College of Arts and Sciences
6. Professor, Political Science, College of Arts and Sciences
7. Associate Graduate Student, Political Science, College of Arts and Sciences
8. Professor, Sociology, College of Arts and Sciences
9. Associate Professor, Environmental Health Sciences, College of Public Health
10. Associate Professor, School of Communication, College of Arts and Sciences
11. Associate Professor, School of Communication, College of Arts and Sciences
12. Professor, Political Science, College of Arts and Sciences
Purchased Gallup Analytics and Microdata data files

Initial Library Contact/Communication: September 2016

Licensing Highlights

- Export Control
- Two Licenses for One Deal

Lesson Learned: Know what you buying
WEB OF SCIENCE

Patron → Subject Librarian A → Subject Librarian B → Collection Strategy/Electronic Resources Librarians

Subject Library A/B → Research Commons Staff → Data Management Services Librarian → Interim Head of Acquisitions

Electronic Resources Officer → Data Management Services Librarian → Patron → Need met ???
Purchased Web of Science Core Collection Data Set + 3 years of data refreshes

Initial Library Contact/Communication: November 2016

Licensing Highlights:
- Statement of Work Agreement

Lesson Learned: Involve appropriate parties as early as possible
Purchased Trucost Academic Portal – Three Year Access

Initial Library Contact/Communication: October 2016

Licensing Highlights:
- Keep an eye on language changes
- Vendor Industry may not be the same as the University/Library Industry

Lesson Learned:
- New to Academic Libraries?
- Teaching and Learning Aspects as well as research
OTHER DATA PURCHASES

- **BioCyc: Licensing Highlights:** Problems can sometimes come from within

- **Data from Foreign Vendors:**
  - Holidays can impact timing
  - What part of licenses can be modified

- **Data from Government entities:**
  - May not be able to do multi-year agreements
  - No License may be required – easy peasy
THINGS WE DIDN’T BUY (YET)

- Things that had been purchased previously
  - Updates to existing data set
- Clarification around format, delivery and access methods
- Ran out of time
- Recommendation: Requestors submit through New Resource Proposal program in following fiscal year
THINGS WE DIDN’T BUY

- Couldn’t license for all of the university
  - Restricted to Grad students and faculty
- Decided to buy content, not tools
- Content didn’t meet requestor need
OVERALL LESSONS LEARNED

- Local hosting issues
  - Identify who can or will host any purchased data
  - Is there enough server space?
  - Front End interface development
- Patience really is a virtue
  - Licensing will take time and a lot of back and forth
- Communication
  - Keep people in the loop
  - Once the Libraries takes on a request, we become the owner of the process and timeline
WHAT NOW??

- Describe processes
- Develop criteria
- Budget??
- Assess pilot and purchases
  - Have researchers used the data?
  - Development of metrics/measurement
    - Including impact on Research Commons staff and services
Swanson, J., & Rinehart, A. K. (January 01, 2016). Data in context: Using case studies to generate a common understanding of data in academic libraries. The Journal of Academic Librarianship, 42, 1, 97-101; http://dx.doi.org/10.1016/j.acalib.2015.11.00

Research Commons: https://library.osu.edu/researchcommons/

Translational Data Analytics Institute: https://discovery.osu.edu/tda/
FOR MORE INFORMATION

Anita Foster (foster.1037@osu.edu)
Gene Springs (springs.14@osu.edu)