Keeping Pace with Research:

Maintaining Effective Data Services at Academic Libraries

Cameron Riopelle, PhD
Head of Data Services
University of Miami Libraries

Goals of presentation

PRIMARY:

I will share my thoughts on developing new skills and knowledge needed to maintain and support Data and RDM services in a rapidly evolving ecosystem.

SECONDARY:

I will provide an overview of how the University of Miami has developed robust research data management services in response to the needs of researchers.

- This is both a subjective case study and an opportunity for discussion and mutual learning
- Every institution is distinct and is composed differently than others
- University of Miami is a medium-sized private university with a strong marine/atmospheric science program and a medical campus

Goal of Data Services

- Data services triangulates between data curation/management, software instruction, and research methods.

- Goal: To support researchers working with data within different disciplines, software, and research methods.

- *So how do we work toward this goal?*

Data Services at the University of Miami Libraries

Our Research Data Management bundle:

- CURATION: data management plan (DMP) review, DOI assignment, assistance with data deposit process, Esploro (Ex Libris) RIS (in progress)

- EDUCATION: Workshop series (GIS, data management, research methods software, visualization)

- EXPERTISE: Consultations, Digital Scholars' Lab (specialized computing lab), Learning Commons, open office hours

Our history:

Data Services for 2 years, GIS Services for 4 years, newly merged

Our staffing:

- Head of Data Services
- GIS Librarian
- GIS & Data Specialist
- Biomedical Data Librarian* (search open)
- Research Data Scientist (joint appointment with Center for Computational Studies)
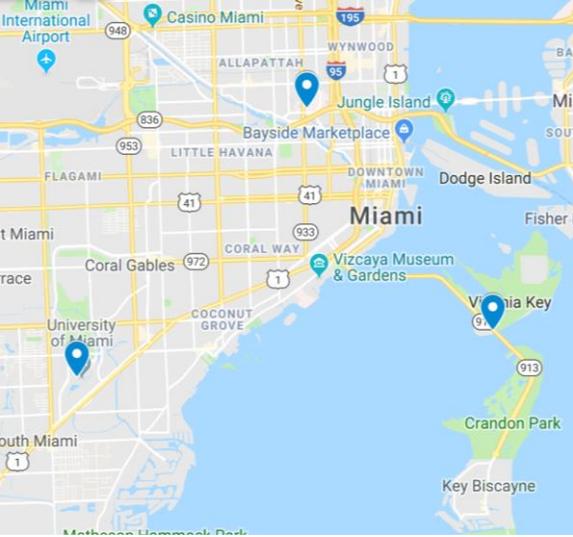
Spaces of Data Services

The State of Data Services at the University of Miami

Things we are doing well

Stats + survey workshops, GIS course supports, consultations (+400 data services a year, +800 GIS including license administration)

Things we are working on

Improving DS lab, increasing staffing (Biomedical Data Librarian), unifying data collections for assessment, collection development plan

Things we are not doing (but should)

Workshops on qualitative research software

More data visualization that is narrative/storytelling in nature (pilot Data Viz program)

A Philosophy of Data Services

Data services operates at intersection of data curation, research methods, and disciplinary knowledge

DATA CURATION: "traditional" approach to data services and still important

RESEARCH METHODS: bringing in applied methods to make DS continuously relevant

DISCIPLINARY KNOWLEDGE: The knowledge the researcher/student/faculty members brings to the table and not necessarily your job to know

The right approach can differ from field to field.

It's ok not to wear the expert hat here.

Becoming Data Services

FREE VS. COST-RECOVERY

- Cost-Recovery: Better tech, better capacity to upgrade
- Free: Can be egalitarian, create equity, given social and cultural context, word of mouth
- Learning how to set boundaries and standards of workflow

WELCOMING: Go for "Meet Halfway" Attitude, keep in mind positionality, possibility of microaggressions, communications, insignificant findings

USEFUL: People must want to come back a second time. Not a one-time service.

- Relevant: Keep up to date, know something to start
- Comprehensive: Must expand knowledge at or near pace of research world, must have awareness of what is happening in research

Keeping up-to-date with Research

The Universe Metaphor

- Research constantly expanding, boundaries expanding.

The Surfing Metaphor

- We exist at edge of these boundaries, as that's where uncertainty is. Finding balance.
- Must be comfortable with uncertainty.
- Must be comfortable with not being "expert."

Assessing Need

- **Examples of Direct requests:** Nvivo, Tableau, REDCap

- **Examples of Indirect requests**: help with data management, help with learning a stats program, help with replicating a study

Training

- Make room for training
- Research day?
  - Use the summer for extra training
- Projects are good ways to learn
  - Low stakes
  - Known short timeline until completion
- Use methodology sections of peer-reviewed articles

- Try to be your own best practices
  - Reinforce general good computing skills and data management practices
  - Using fake data can be very useful for learning and for situations in which you cannot use or see the real data (e.g., human subjects)

A consultative process

POSSIBLE SCENARIOS

1. Meet (evaluate, listen) → immediate resolution
2. Meet (evaluate, listen) → prepare, study → Second meeting → … →
3. Meet (evaluate, listen) → refer
4. Meet (evaluate, listen) → decline to help (possibly refer)

- The prep and study increases possibility for future immediate resolution

- The act of referral can lead to further acts of referral back to you

- Resolution often leads to a second visit anyway for new and emerging issues

A Situation

"I am currently studying the movements of large coastal sharks in Biscayne Bay, FL via passive acoustic biotelemetry and would like help quantifying the: i) residency; ii) site fidelity; and iii) activity spaces (minimum convex polygons) of these animals utilizing R software."

Where do you begin? What do we know?

- Solution must use R

- Solution has something to do with measuring residency, site fidelity, and activity spaces (minimum convex polygons)

- Two paths: start with R packages, or start with the literature

Two paths

Analysis of Animal Movements in R:
the `adehabitatLT` Package

Clement Calenge,
Office national de la chasse et de la faune sauvage
Saint Benoist – 78610 Auffargis – France.

March 2019

https://cran.r-project.org/web/packages/adehabitatLT/vignettes/adehabitatLT.pdf

"We used a hierarchical, first-difference, correlated, random-walk, switching SSM (hDCRWS) [43]. This model allows for estimate parameters jointly across multiple individual tracks."

https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5567640/

https://www.whoi.edu/cms/files/patterson_movement_state_space_models_tree_57192.pdf

Finally, leads to:

https://rdrr.io/cran/bsam/man/fit_ssm.html

Both paths should be made clear to the researcher and ideally should reinforce each other

Another Situation

"I am looking for some help in designing and conducting a propensity score analysis study with administrative data we have collected and to link for children in early learning programs, their participation, attendance and school outcomes."

Questions:

- What program(s) do they know?

- What program(s) do they have?

- How can data services help?

After the first meeting, it turns out the researcher has Stata and wants to confirm *that this is possible*

The meeting ends up being a demo of how to propensity matching in Stata using sample data following: https://www.stata.com/meeting/italy14/abstracts/materials/it14_grotta.pdf

I was not allowed to look at the data since it was human subjects (and children) and I was not on the IRB

Approaches to working with researchers

- Take an attitude of "meeting them halfway"

- Don't overpromise

- Don't *solely* take on the expert identity

- The first consultative meeting should be framed as evaluative in nature

The question of value

- Value to peers

  - Demystifying data

- Value to researchers

- Value to research enterprise

Closing Thoughts

Back to the Philosophy of Data Services

- Universe metaphor
- Surfing metaphor

Stay afloat. Let others stay afloat.

- Keep on training
- Make room for training
- The invisible labor of training should be encouraged by department and administration

Questions?

THANK YOU VERY MUCH!