Finding your niche: Exploring the data ecosystem as an instruction librarian

Kelly Grossmann Getz, MS Information, MS Bioinformatics
Assistant Professor, STEM Librarian
Eastern Michigan University
kgrossm3@emich.edu
MiALA Data Literacy Boot Camp, March 8, 2019

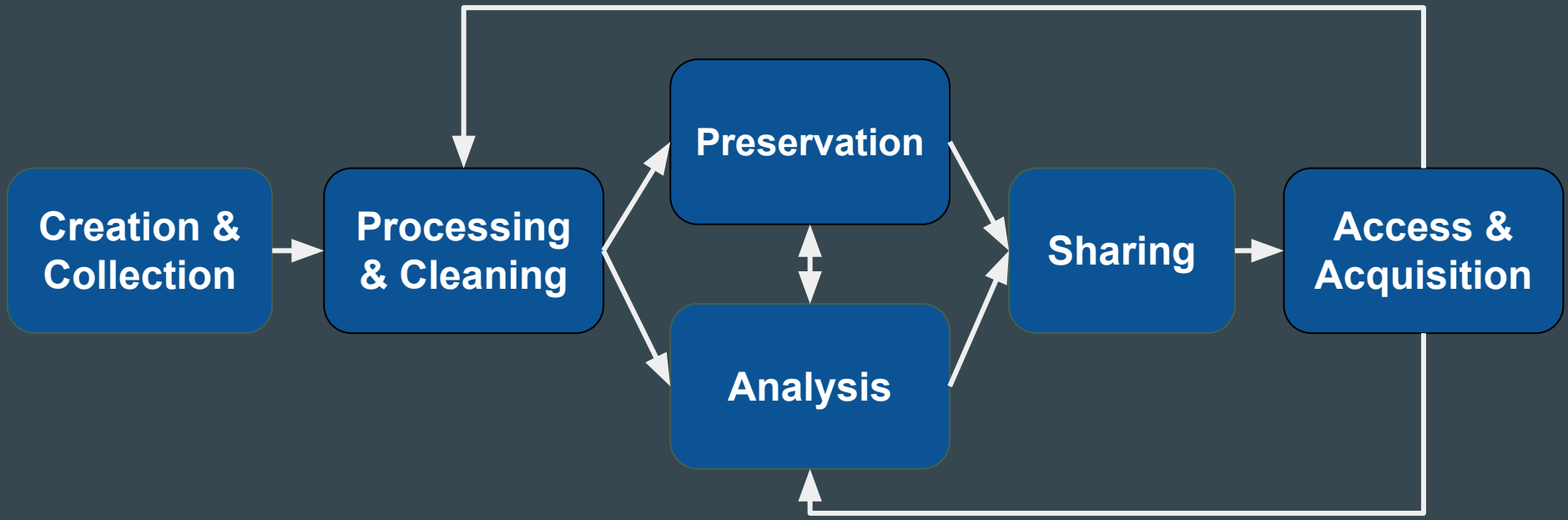
“The accidental data librarian” - Cindy Severt, IASSIST, 2005

The accidental data instruction librarian

Photo used with permission from: *lilothehusky/Instagram*

Fitting in
In the Research Data Ecosystem

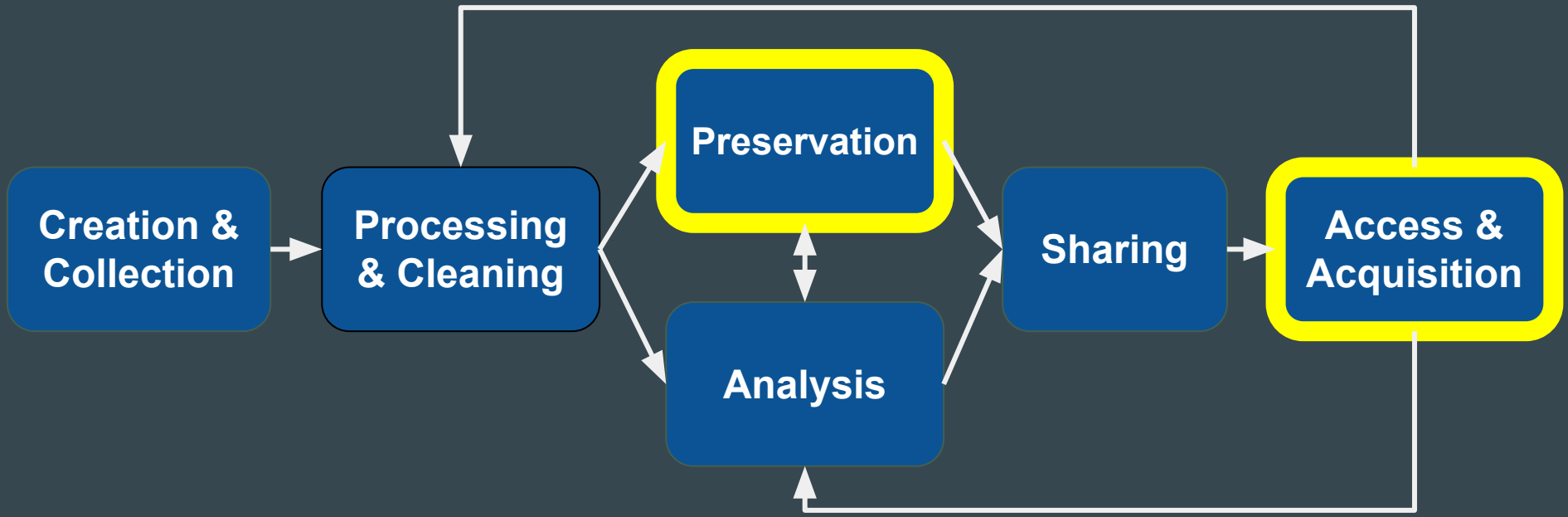
Data Lifecycle



Inspired by and adapted from:
https://www2.usgs.gov/core_science_systems/access/summer_2013/article-12.html

Librarian assistance

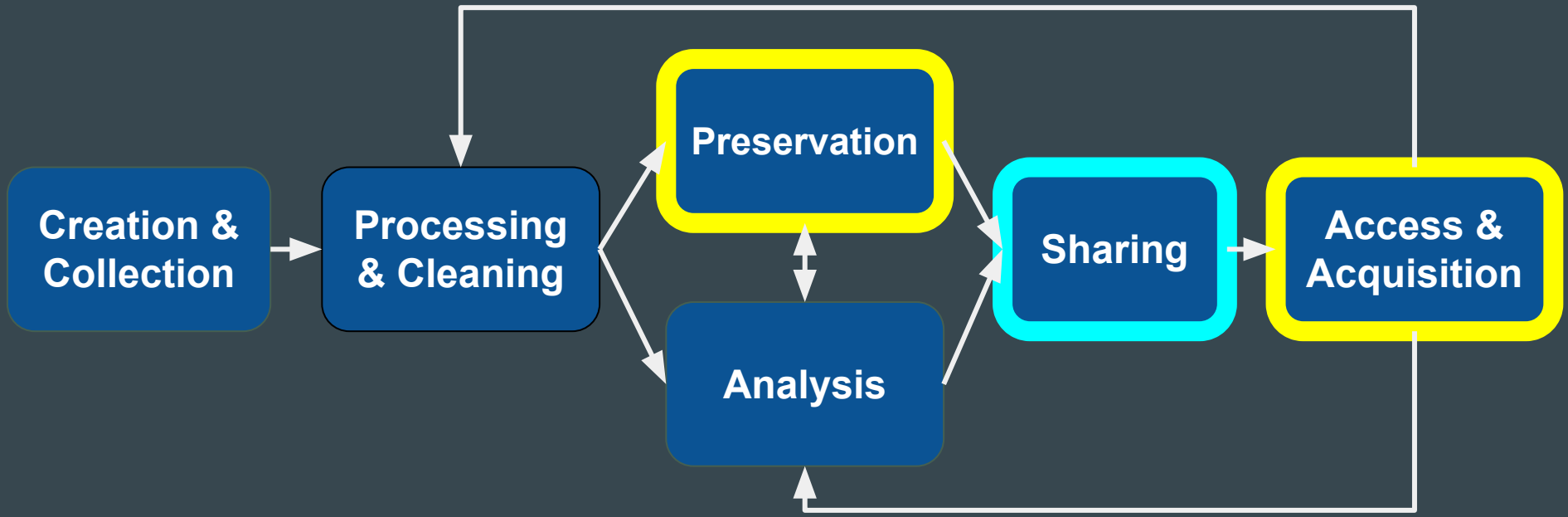
Librarians as stewards



Librarian assistance

Librarians as stewards

Librarians as facilitators

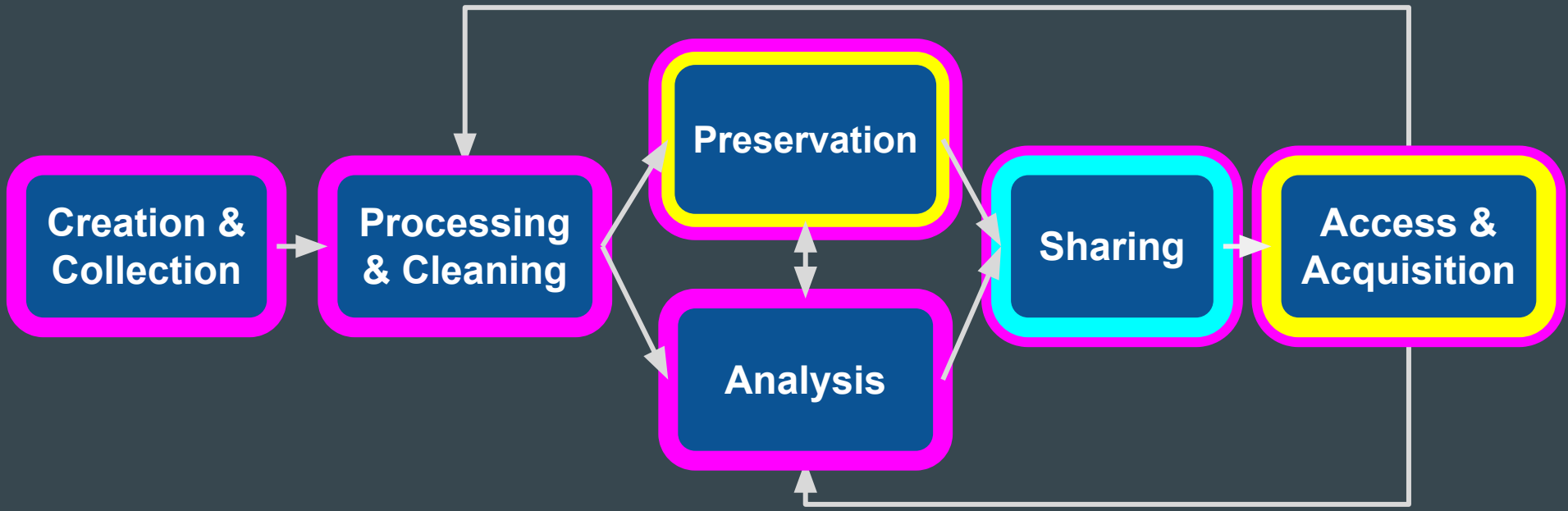


Librarian assistance

Librarians as stewards

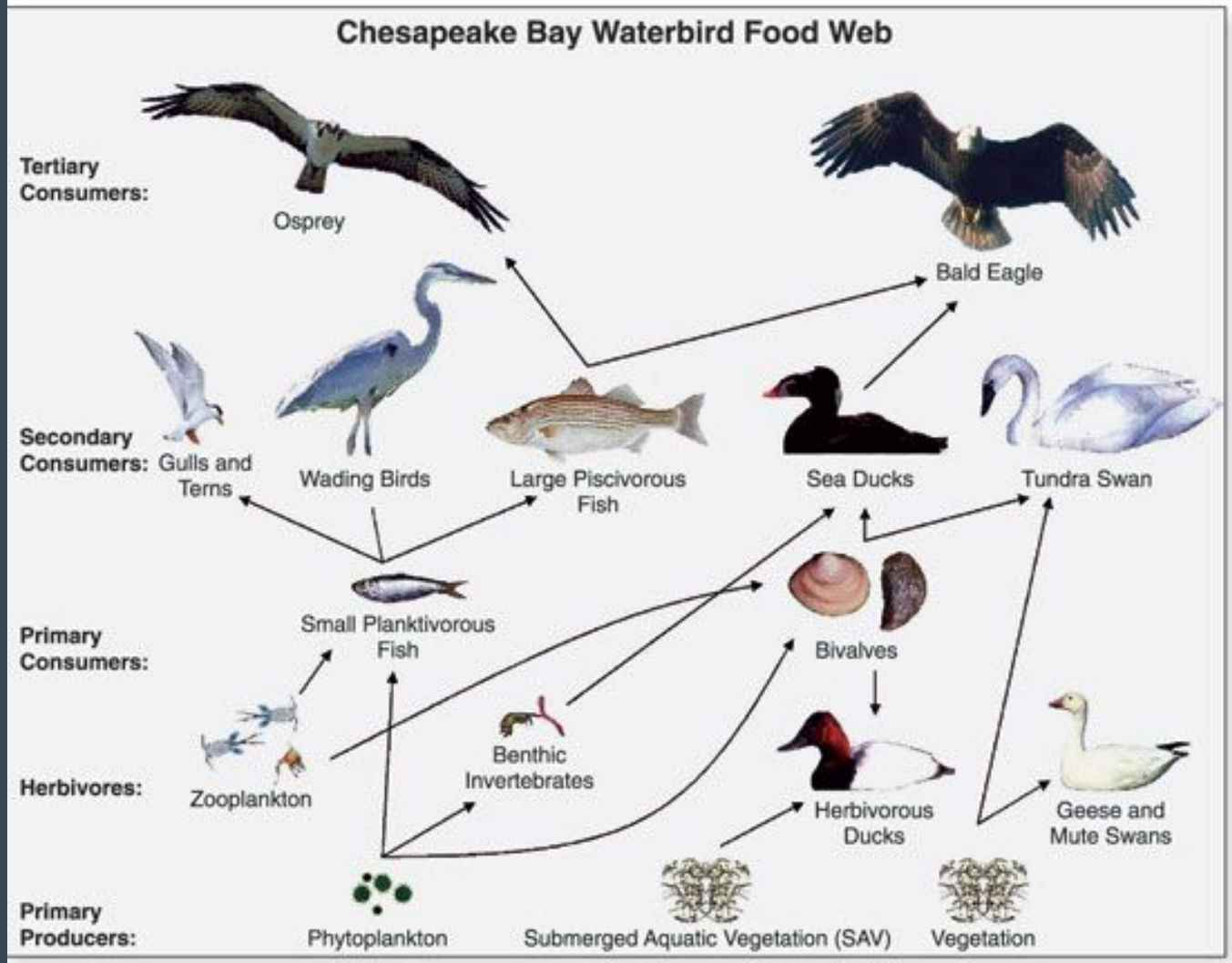
Librarians as facilitators

Librarians as educators



Food Web

Demonstrate producers and consumers in an ecosystem



Research Data Ecosystem

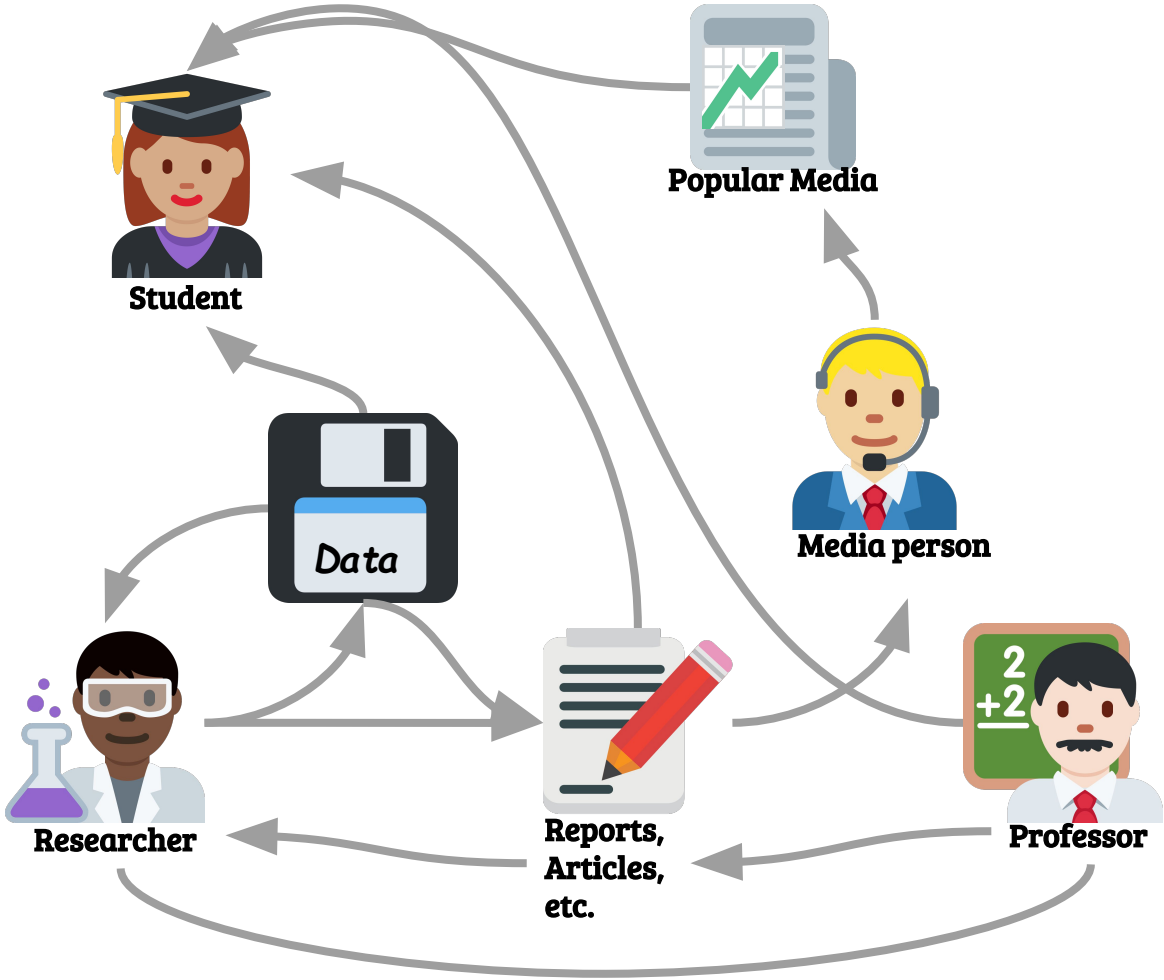
Researchers produce data

Researchers analyze others' data & their own

Students learn to analyze

Students consume media

Students read articles



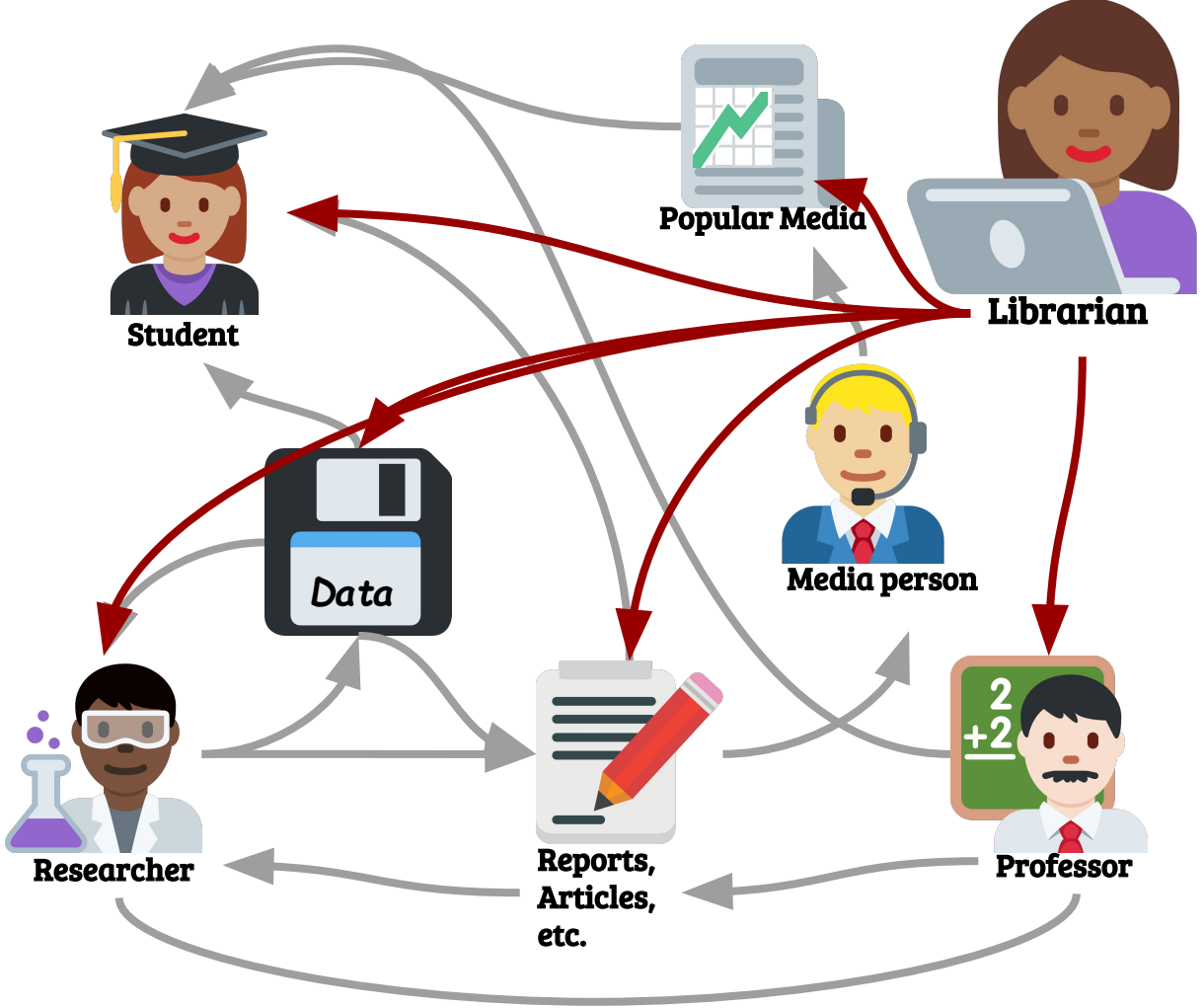
Research Data Ecosystem

Librarians help access data

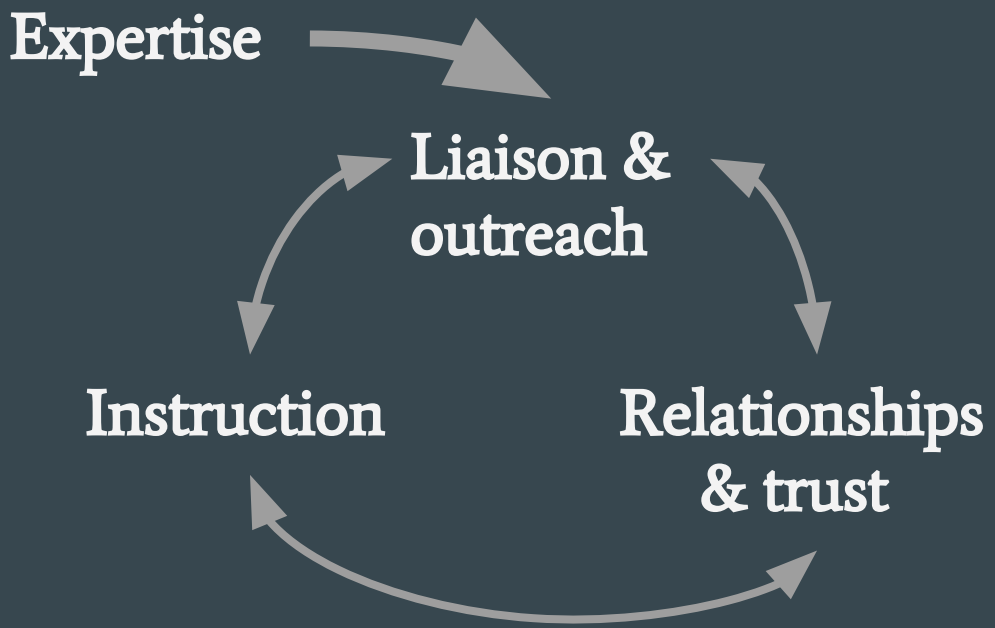
Librarians help professors store & describe data

Librarians teach professors about data

Librarians teach students about access, analysis, and evaluation



Expertise is our way into the “circle of trust”



Different types of users have different needs

The roles of the user

Consumers

Analyzers

Generators

Data Users as Primary, Secondary, and Tert. Consumers

Who are they?

Everyone! People who view and interact with data visualization and representations on an everyday basis either in popular media or academia

Researchers using others' published datasets

What do they need?

Help understanding and evaluating the information they are receiving.

Data Users as Analyzers

Who are they?

Individuals who analyze data, create statistics, and design visualizations of data.

Students in undergraduate laboratories;
Faculty who use large published datasets;
Graduate students creating GIS maps; Etc.

What do they need?

Space; Software; Information about analytical tools; clean datasets to work from!

Data Users as Data Generators

Who are they?

Individuals who collect original data

Undergraduate students doing research;
Graduate students doing original research;
University faculty & researchers; Etc.

What do they need?

Help with DMP; Education about clean formatting ; Infrastructure for preservation; Assistance with metadata creation

Data Needs differ in Different Disciplines

Sciences

Social Sciences

Business

Sciences

Types of data used:

- Big quantitative data sets (geospatial, genomic, etc.)
- Protected or private data (GWAS, etc.)
- Student lab data
- Images & records for natural history

Needs:

- Help locating, accessing, storing, and sharing big quantitative data
- Software for analyzing & visualizing data
- DMP assistance

Social Sciences

Types of data used:

- Census data
- Geospatial data
- Interview results
- Survey results

Needs:

- Help storing confidential data
- Help locating, accessing, storing, and sharing big population data
- Software for analyzing qualitative data
- Software for maps
- DMP assistance

Business

Types of data used:

- Market research
- Statistics
- Qualitative research from customers
- Industry data

Needs:

- Assistance locating and acquiring pricey data
- Help evaluating published reports

Activity: Map user needs to your skills

Complete some of the table from the handout.

User	How does the user interact with data?	What is the user's role? (Consumer/ Generator/ Analyzer)	What does the user need?	What is the librarian's role? (Steward/ Facilitator/ Educator)	How can I help? (Know how/ Can learn/ Will refer)
Undergraduate Students	*Read news media about climate change*	*Consumer*	*Strategies to evaluate data centric arguments and images*	*Educator*	*I know how to do this.*
Researchers	*Collect and publish geospatial data for ecological study*	*Generator*	*Help constructing metadata to make data findable. A place to store a sizable dataset.*	*Facilitator*	*I know how to help construct metadata; Will refer to scholarly comm librarian for storage*

Align your skills to users' needs

- Talk to faculty to identify gaps in knowledge and needs

Expand and adapt

- Search syllabi for data instruction opportunities

Example Lesson Plans & Activities

Sample Lesson for Students as Tertiary Consumer:

Evaluating Infographics

https://works.bepress.com/dana-statton/13/

https://www.state.gov/j/prm/about/265946.htm

Evaluating Infographics [Students as Tertiary Consumers]

Prep level: Moderate

Time: About 2 hours

Audience: 1st and 2nd year undergraduates

Materials: Examples of infographics with varying reliability

Objectives: *Student will:*
- Identify data visualizations
- Evaluate the reliability of the visualizations
- Discuss the author's motives and meaning
- Locate the original sources of data

Activity: Discuss examples of misleading data visualizations and have students work in groups to evaluate published infographics using the CRAAP test

For a full course based on infographics and information literacy, visit : https://works.bepress.com/dana-statton/13/

Sample Lesson for Students as Analyzers:

Water Resource Datathon

https://guides.emich.edu/datathon

Water Resource Datathon [Students as Analyzers]

Prep level: High

Time: ~6 hours of contact time

Audience: Students, Community members, etc.

Materials: Computers with internet; local water data Snacks; Prizes

Objectives: *Student will:*
- Construct research questions based on available data
- Present analyses and questions for further research

Activity: Allow groups of 2-4 students to explore datasets and create research questions. Prepare & provide brief demonstrations of technological tools for analysis & visualizations (spreadsheet functions, map software, etc.). Have students present brief 5-10 minute presentations based on their research questions and analyses. Judge based on rubric.

For more information visit :
https://guides.emich.edu/datathon

time	n.risk	n.event	survival	std.err	lower 95% CI	upper 95% CI
5	23	2	0.913	0.0588	0.8049	1
8	21	2	0.8261	0.079	0.6848	0.996
9	19	1	0.7826	0.086	0.631	0.971
12	18	1	0.7391	0.0916	0.5798	0.942
13	17	1	0.6957	0.0959	0.5309	0.912
18	14	1	0.646	0.1011	0.4753	0.878
23	13	2	0.5466	0.1073	0.3721	0.803
27	11	1	0.4969	0.1084	0.324	0.762
30	9	1	0.4417	0.1095	0.2717	0.718
31	8	1	0.3865	0.1089	0.2225	0.671
33	7	1	0.3313	0.1064	0.1765	0.622
34	6	1	0.2761	0.102	0.1338	0.569
43	5	1	0.2208	0.0954	0.0947	0.515
45	4	1	0.1656	0.086	0.0598	0.458
48	2	1	0.0828	0.0727	0.0148	0.462
30	9	1	0.4417	0.1095	0.2717	0.718
31	8	1	0.3865	0.1089	0.2225	0.671
33	7	1	0.3313	0.1064	0.1765	0.622
34	6	1	0.2761	0.102	0.1338	0.569
43	5	1	0.2208	0.0954	0.0947	0.515
45	4	1	0.1656	0.086	0.0598	0.458
48	2	1	0.0828	0.0727	0.0148	0.462

Lesson for Students as Generators:

Organizing Data

Sharing Student Researcher Data [Students as Generators]

Prep level: Low

Time: ~2 hours of lab time

Audience: REU Students (or other student researchers)

Materials: Computer lab & data from student research

Objectives: Student will:
- Identify how their data may be used in future research
- Organize their data
- Begin authoring metadata

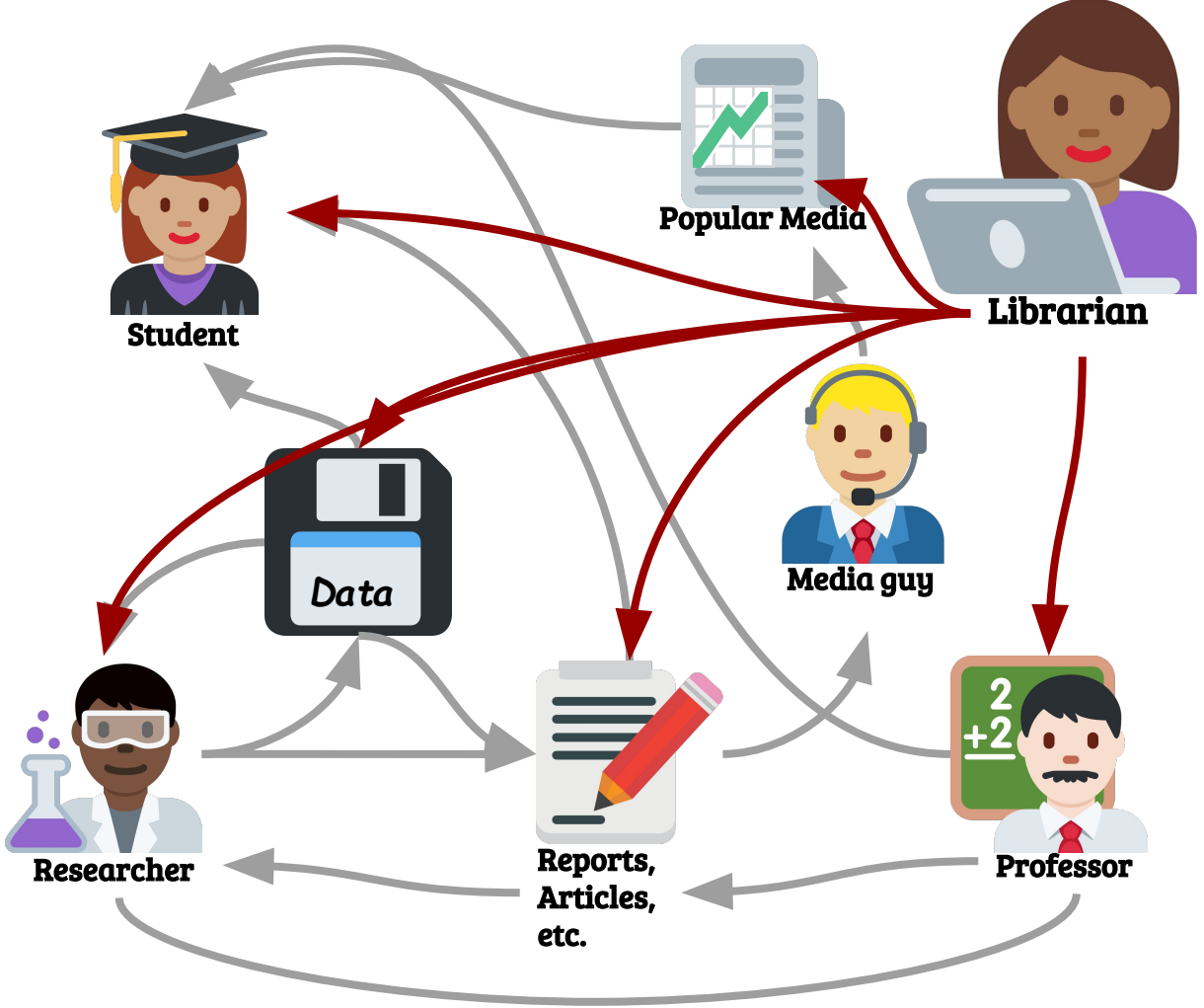
Activity: Hold an open discussion to explore why data sharing is important. Provide an example of a publicly available dataset and metadata. Workshop with students to evaluate how their data tables are set up and the content of their field names. Give students actual forms for submitting data set metadata in an institutional repository.

Pair-Share

- What type of student do you usually work with-- Consumers, Generators, or Analyzers?
- How could you adapt the sample activities to fit with your subject area disciplines?
- How could you adapt the sample lessons to suit additional instructional constraints (time, technology, etc.)?
- What other data literacy lessons do you perform in your classes?

Research Data Ecosystem

Hopefully, this has helped you identify some strategies to figure out where you fit in the research data environment!



Thank you!

Remaining questions or discussion?

Kelly Grossmann Getz

kgrossm3@emich.edu

Further Reading

Rice, R., & Southall, J. (2016). *The Data Librarian's Handbook*. London: Facet Publishing. Retrieved from http://ezproxy.emich.edu/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=nlebk&AN=1454580&site=ehost-live&scope=site

Anderson, D. E. (2016) Top five tips for finding new roles in your library. In Anderson, D. E., & Pun, R. (ed) *Career transitions for librarians: Proven strategies for moving to another type of library.*

Slide citations

Slide 3: Severt, Cinty. (2005) “Discovering a Profession: The accidental data librarian.” Presentation. https://iassistdata.org/conferences/archive/2005?page=5

Slides 4-7:
Faundeen, J. L., ... (2013) “The USGS Science Data Lifecycle Model” *Access Summer 2013.* 16(1). https://www2.usgs.gov/core_science_systems/access/summer_2013/article-12.html

Brodsky, M., Getz, K. (2018) “Diving into the Data Literacy Deep End” *Library Instruction West.* Presentation.

Slide 25-26
Thomson, D. S. “Infographics and visual literacy: Teaching evaluative criteria to increase critical thinking.” *Library Instruction West.* Presentation.

Image attributions

Slide 2: Image of cat with huskies. Used with permission from lilothehusky/Instagram.

Slide 8: Wikimedia Commons. “Chesapeake Waterbird Food Web.” Public Domain.
https://commons.wikimedia.org/wiki/File:Chesapeake_Waterbird_Food_Web.jpg

Slides 9-10:
Wikimedia Commons. “Twemoji2 1f4be.svg” by Ebrahim. CC 4.0. Cropped, resized, word added.
https://commons.wikimedia.org/wiki/File:Twemoji2_1f4be.svg

Wikimedia Commons. “Emojione1 1F4DD.svg” by Ebrahim. CC 4.0. Cropped and resized.
https://commons.wikimedia.org/wiki/File:Emojione1_1F4DD.svg

Wikimedia Commons. “Emojione 1F4C8.svg” by Ebrahim. CC 4.0. Cropped and resized.
https://commons.wikimedia.org/wiki/File:Emojione_1F4C8.svg

Wikimedia Commons. “Twemoji2 1f4f0.svg” by Ebrahim. CC 4.0. Cropped and resized, graph image added.
https://commons.wikimedia.org/wiki/File:Twemoji2_1f4f0.svg

Wikimedia Commons. “Twemoji2 1f468-1f3ff-200d-1f52c.svg” by Ebrahim. CC 4.0. Cropped and resized.
https://commons.wikimedia.org/wiki/File:Twemoji2_1f468-1f3ff-200d-1f52c.svg

Image attributions

Wikimedia Commons. “Twemoji2 1f468-1f3fb-200d-1f3eb.svg” by Ebrahim. CC 4.0. Cropped and resized. https://commons.wikimedia.org/wiki/File:Twemoji2_1f468-1f3fb-200d-1f3eb.svg

Wikimedia Commons. “Twemoji2 1f469-1f3fe-200d-1f4bb.svg” by Ebrahim. CC 4.0. Cropped and resized. https://commons.wikimedia.org/wiki/File:Twemoji2_1f469-1f3fe-200d-1f4bb.svg

Wikimedia Commons. “File:Twemoji2 1f468-1f3fc-200d-1f4bc.svg” by Ebrahim. CC 4.0. Cropped and resized. https://commons.wikimedia.org/wiki/File:Twemoji2_1f468-1f3fc-200d-1f4bc.svg

Slides 13,14: Wikimedia Commons. “A Reporter Watches Secretary Kerry Deliver Remarks on His Cellphone (29864739305).jpg”. Public Domain. https://commons.wikimedia.org/wiki/File:A_Reporter_Watches_Secretary_Kerry_Deliver_Remarks_on_His_Cellphone_(29864739305).jpg

Slides 13, 15: Wikimedia Commons. “Coding Shots Annual Plan high res-5.jpg” by Matthew. CC 3.0. Cropped and resized. https://commons.wikimedia.org/wiki/File:Coding_Shots_Annual_Plan_high_res-5.jpg

Slides 13,16: Wikimedia Commons. “Plant ecologist Carl Roland collects soil temperature and moisture data while performing soil sampling protocol in Denali NP. (8446139444).jpg.” Public domain. https://commons.wikimedia.org/wiki/File:Plant_ecologist_Carl_Roland_collects_soil_temperature_and_moisture_data_while_performing_soil_sampling_protocol_in_Denali_NP._(8446139444).jpg

Image attributions

Slides 17,18: Wikimedia Commons. “Twemoji2 1f469-1f3fd-200d-1f52c.svg” by Ebrahim. CC 4.0. Changes made: Cropped and resized. https://commons.wikimedia.org/wiki/File:Twemoji2_1f469-1f3fd-200d-1f52c.svg

Slides 17,19: Wikimedia Commons. “Twemoji2 1f469-1f3fb-200d-2696-fe0f.svg” by Ebrahim. CC 4.0. Changes made: Cropped and resized. https://commons.wikimedia.org/wiki/File:Twemoji2_1f469-1f3fb-200d-2696-fe0f.svg

Slides 17,20: Wikimedia Commons. “Emoji u1f468 1f3fe 200d 1f4bc.svg” by Ebrahim. CC 4.0. Changes made: Cropped and resized. https://commons.wikimedia.org/wiki/File:Emoji_u1f468_1f3fe_200d_1f4bc.svg

Slide 22: Wikimedia Commons. “Flightless Dung Beetle Circellium Bachuss, Addo Elephant National Park, South Africa.JPG” by Kay-africa. CC share-alike 3.0. Changes made: Cropped and resized. https://commons.wikimedia.org/wiki/File:Flightless_Dung_Beetle_Circellium_Bachuss,_Addo_Elephant_National_Park,_South_Africa.JPG

Slide 23: Wikimedia Commons. “Squirrel eating an ice cream cone” by Einstein2. CC 2.0 Generic. Changes made: Cropped and resized. https://commons.wikimedia.org/wiki/File:Squirrel_eating_an_ice_cream_cone_(3558009610).jpg

Slide 25: U.S. Department of State. “U.S. Department of State, Bureau of Population, Refugees and Migration 2016 Year in Review.” Public domain. Changes made: resized.

Slide 27: Anonymous creator. “Lead.” From the Water Resource Datathon. 2018.
Slide 29: Wikimedia Commons. “Life table for the aml data.png.” by Michaelg2005.. CC Share alike license. Changes made: Cropped and resized https://commons.wikimedia.org/wiki/File:Life_table_for_the_aml_data.png