NIAMS IRP DATA DISCOVERY AND ANALYSIS (DnA)

National Institutes of Arthritis and Musculoskeletal and Skin Diseases (NIAMS)

Christine Winchester — IRP Liaison

CIVIC DIGITAL FELLOWSHIP

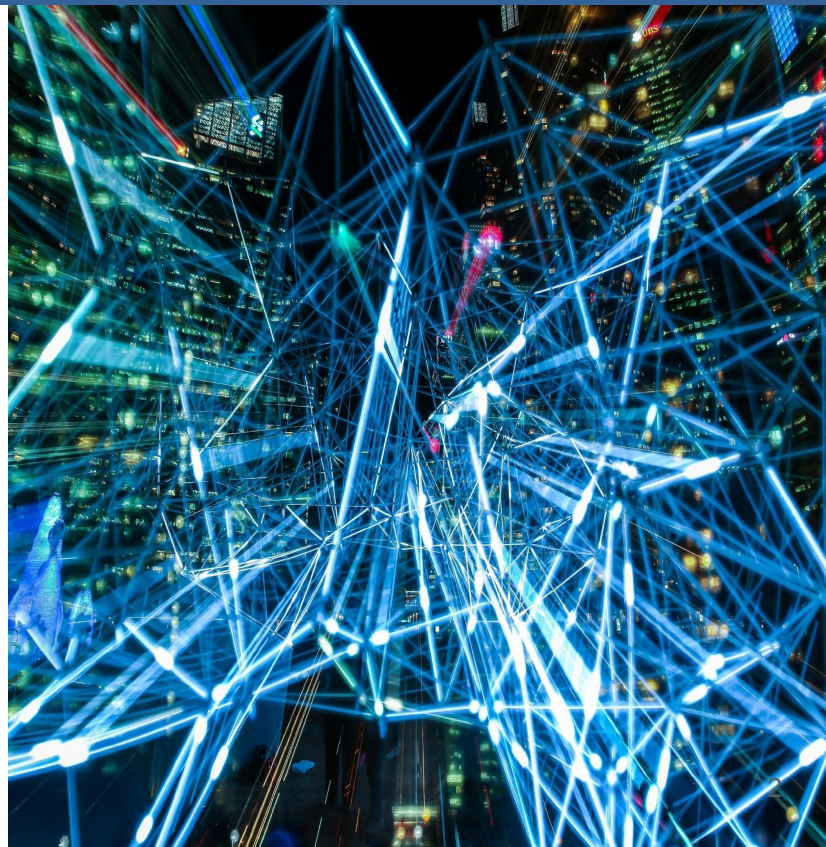
NIH

ALANA BURRELL
Spelman College
CS + Computer Engr.

ELEANE YE
Duke University
Public Policy + CS

BACKGROUND

- NIAMS IRP Scientific research in genomics, flow cytometry, microscopy, and other experiments generate a large amount of data
    - 10 TB of data generated monthly
    - Over 1 PB of data currently stored on-premise
- There is currently no disciplined approach for managing data as an asset



CIVIC DIGITAL FELLOWSHIP

OUR CHALLENGE

Establish an approach for managing NIAMS data that:

- Can be implemented incrementally
- Incorporates the different goals/needs of various stakeholders
- Will support the NIAMS mission and evolving research needs
- Encourages researchers to realize the value and benefits of good data management

CIVIC DIGITAL FELLOWSHIP

TIMELINE

- **Weeks 1-2:** Understanding existing policies, practices, and procedures and how they shape the current data environment

- **Weeks 3-6:** Data collection, surveys & interviews

- **Weeks 7-9:** Analysis and review data/research findings, preparing presentations to management/leadership

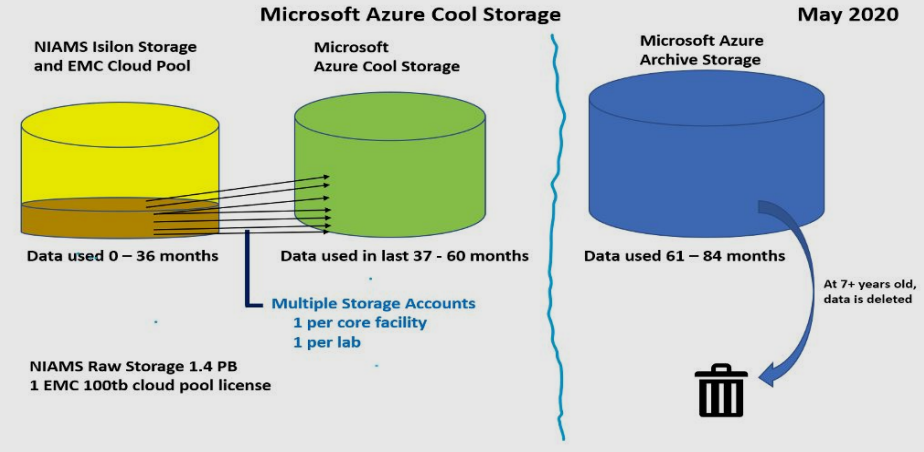
- **Week 10:** Wrapping up, next steps!

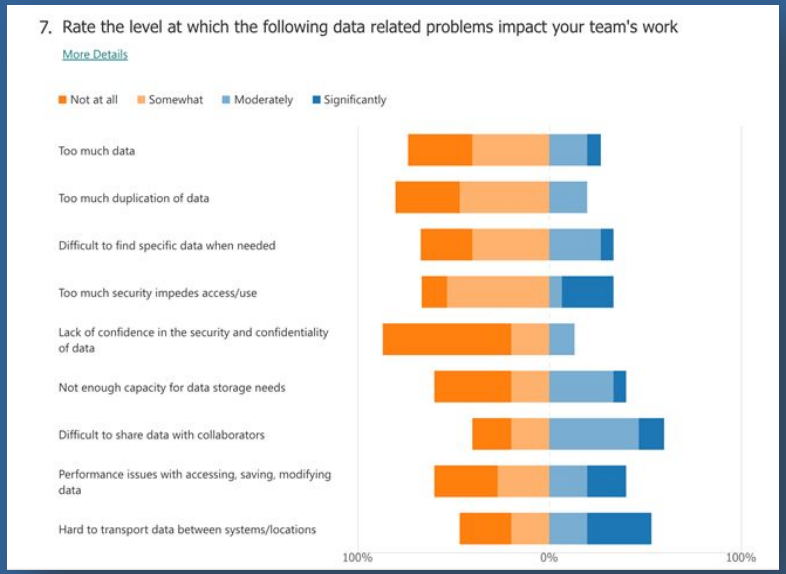
DISCOVERY

NIH/NIAMS POLICY

- Conduct of Research in the IRP at NIH
- NEW NIH Policy on Data Management and Sharing (effective January 25, 2023)

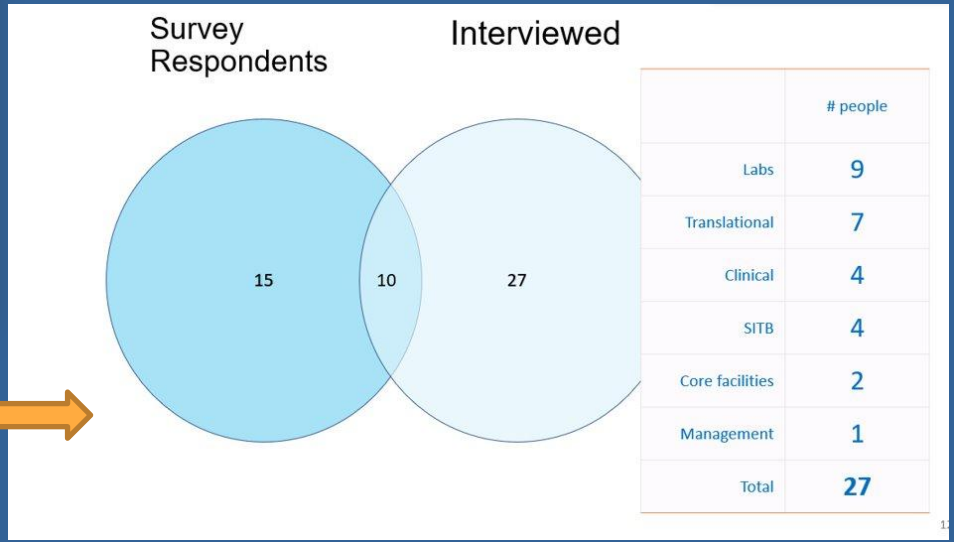
IT PROJECTS & INFRASTRUCTURE





Sample Question From Survey

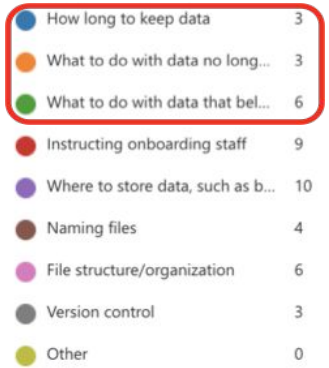
Stakeholder Interview Composition

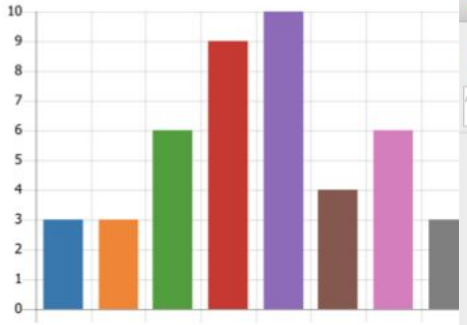


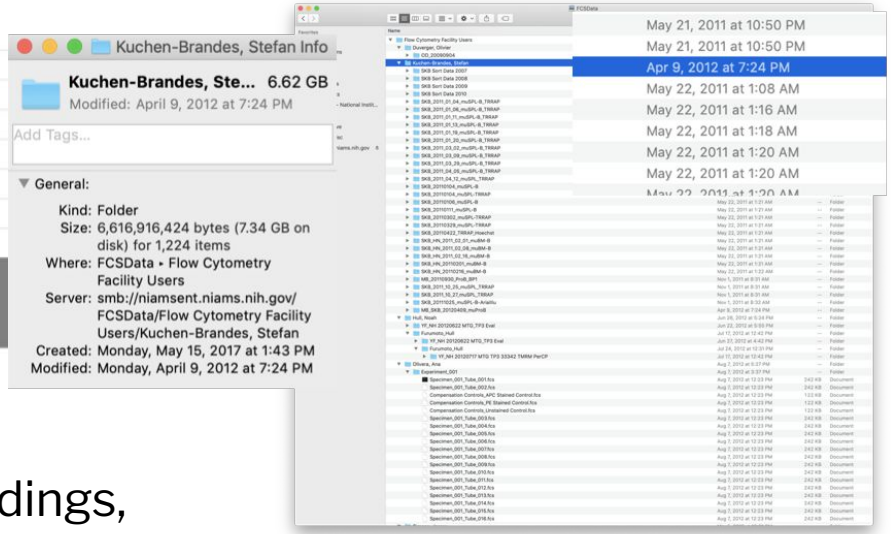
	# people
Labs	9
Translational	7
Clinical	4
SITB	4
Core facilities	2
Management	1
Total	27

CIVIC DIGITAL FELLOWSHIP

ANALYSIS

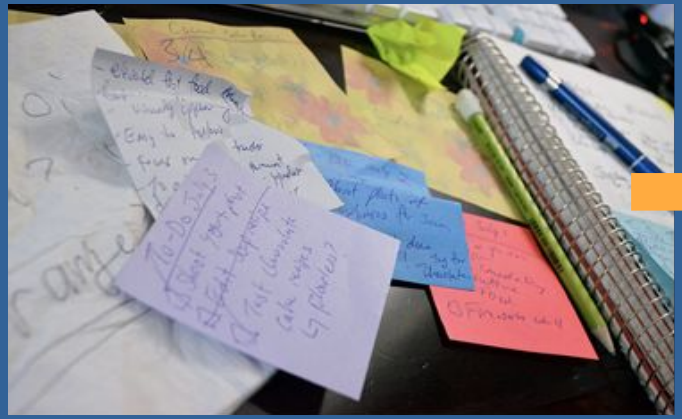


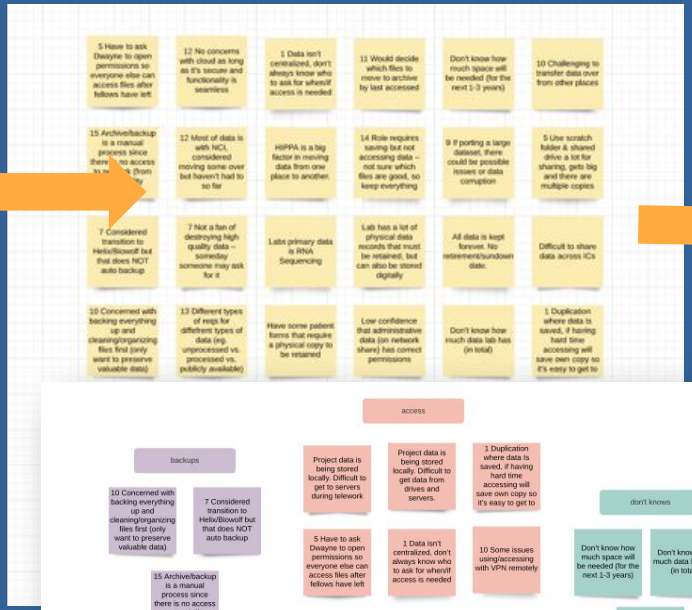


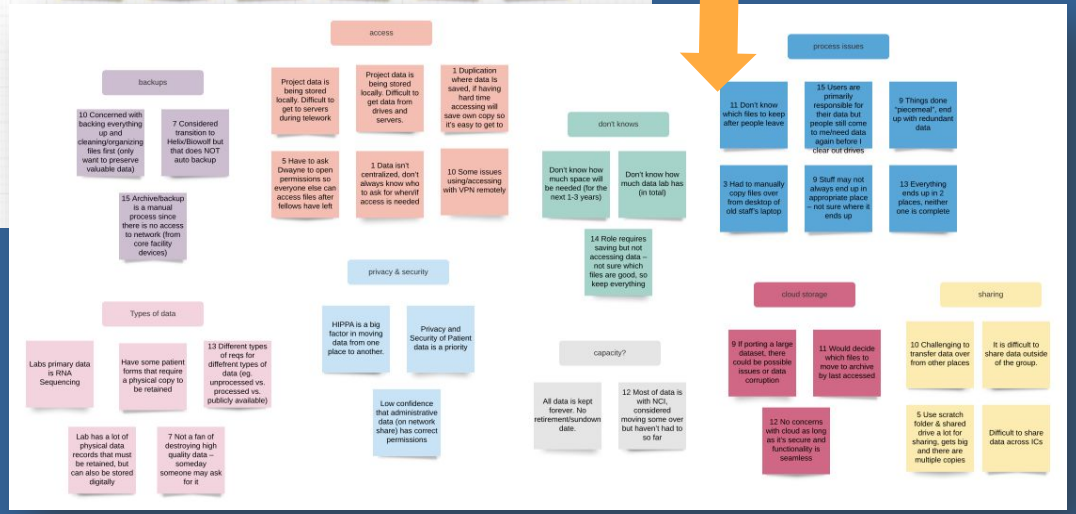


Reviewing our survey and interview findings, comparing qualitative and quantitative data

CIVIC DIGITAL FELLOWSHIP







Affinity mapping helped diagnose complicated problems by organizing qualitative data to reveal themes associated with the problems.

Our affinity mapping process

CIVIC DIGITAL FELLOWSHIP

RESULTS

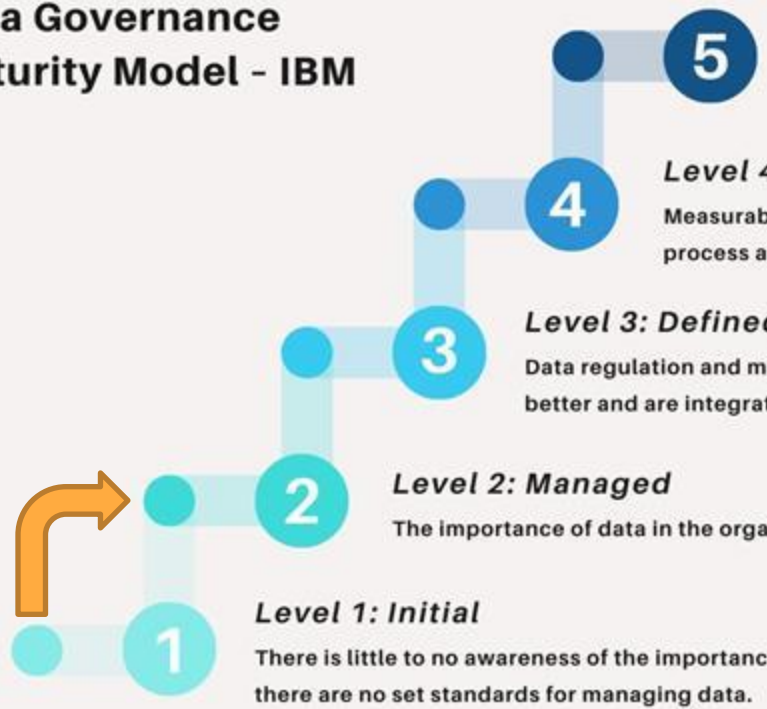
- The majority of groups don’t delete or archive data
  - “It’s impossible to know - things come back at an unexpected angle. [High quality data], someday someone may ask for it.”
- Some decisions made for convenience, rather than process
  - “...If we have a hard time getting the data, we will save [our] own copy so it's easy to get to"

RECOMMENDATIONS

- **Immediate (Within 3 months)**
  - Data Cleanup Campaign
- **Short-term (3-6 months)**
  - Cloud Archive
  - Educational modules, newsletters
- **Long-term (6 months & beyond)**
  - Feedback loop between SITB and IRP partners

CIVIC DIGITAL FELLOWSHIP

Data Governance Maturity Model - IBM



Level 5: Optimizing

Data governance becomes an enterprise-wide effort that improves productivity and efficacy.

Level 4: Quantitatively managed

Measurable quality goals are set for each project, data process and maintenance.

Level 3: Defined

Data regulation and management guidelines are defined better and are integrated with the company processes.

Level 2: Managed

The importance of data in the organization is realized.

Level 1: Initial

There is little to no awareness of the importance of data and there are no set standards for managing data.

A BIG PICTURE TAKEAWAY

"Supposing is good,
but finding out is better"

CIVIC DIGITAL FELLOWSHIP

THANK YOU

- **Christine Winchester, La'Tanya Burton, Diana Mungai**, and SITB
- **Jessica Mazerik** and the NIH team
- **Rachel Dodell, Chris Kuang, Ariana Soto** from Coding it Forward
- Mentors **Will Slack** and **Erica Pincus** for their advice/expertise
- Everyone who participated in our survey and interviews
- All the other **Civic Digital Fellows!**

CIVIC DIGITAL FELLOWSHIP