Library Strategic Planning

Develop campus partnerships to collect, manage, share, and preserve Georgia Tech digital research data

1. Needs Assessment
   - *Research data survey & interviews*
   - *NSF DMP content analysis*
2. Outreach, Training & Consultation Services
   - *Outreach for advocacy and classes*
   - *Data management planning services*
3. Partnerships
   - *New library teams/partnerships*
   - *Campus partnerships through IT governance*
4. Repository Infrastructure & Applications
   - *Research & discovery phase*
   - *Partnerships beyond library & campus boundaries*

Georgia Tech Library

Campus Strategic Planning

Library Proposed Goal

Develop an institute-wide framework for research data stewardship:

- institutional policies
- technical infrastructure
- complementary data services

"Creating a Data Management Framework"
Australian National Data Service
http://ands.org.au/guides/dmframework/data-management-framework.pdf

Georgia Tech Goal

Goal 2—Sustain and enhance excellence in scholarship and research.

Strategy 4: Demonstrate relevance and vitality by Investing in faculty and infrastructure

"Designing the Future: A Strategic Vision and Plan"
http://www.gatech.edu/vision/sites/gatech.edu.vision/files/Georgia_Tech_Strategic_Plan.pdf

Georgia Tech Library

1. Research Data Needs Assessment

Conduct IRB-approved research to determine gaps in data curation services provided to researchers.

Digital Asset Framework-Based Survey

Our goals were to discover:

- what data assets are created & held
- how that data are managed, stored, shared & reused
- researchers' attitudes regarding data creation, sharing & preservation

DAF Implementation Guide: http://www.data-audit.eu/docs/DAF_Implementation_Guide.pdf

DAF Methodology: http://www.data-audit.eu/docs/DAF_Implementation_Guide.pdf

Supporting tomorrow's research: Assessing faculty data curation needs at Georgia Tech
Susan Wells Parham, Jon Bodnar, and Sara Fuchs
http://crln.acrl.org/content/73/1/10.full

Georgia Tech Library

Research Data Assessment

- About the Assessment

NSF Data Plan Requirement

Data Curation Bibliography

About the RD Project

Contact Us

- GT Logout

By *research data* we mean digital information structured by formal methodology for the purpose of creating new research or scholarship. For examples, read about the data assessment.

GEORGIA TECH RESEARCH DATA ASSESSMENT

View | Edit | Webform | Results

Project Information (section 1 of 5)

Please report on the data generated from one research project throughout the entire survey. Choose a project of significant importance to your work (either in process or complete), regardless of your intent to share or preserve the data.

1. Your name:

2. Project name:

3. Briefly describe the project you are using to answer this survey:

4. Briefly describe the data you are using to answer this survey:

5. Indicate your role in this project (check all that apply):

- ☐ **Principal Investigator or Co-PI**
- ☐ **Research/Academic Faculty**

Survey Results

Choose the file formats which best describe your data:	1) Text 2) Spreadsheet 3) Image
Identify how long you plan on keeping the data associated with this project:	1) 1-5 years 2) Indefinitely 3) 5-10 years
Identify which services might be useful in regards to the management of research data:	1) Data storage & preservation (73%) 2) Tools for sharing data (67%) 3) Information regarding data management best practices (52%)
Designate if you have a data management plan or policy:	1) No (60%) 2) Yes (25%) 3) Don't know (15%)

Georgia Tech Library

Researcher Interviews

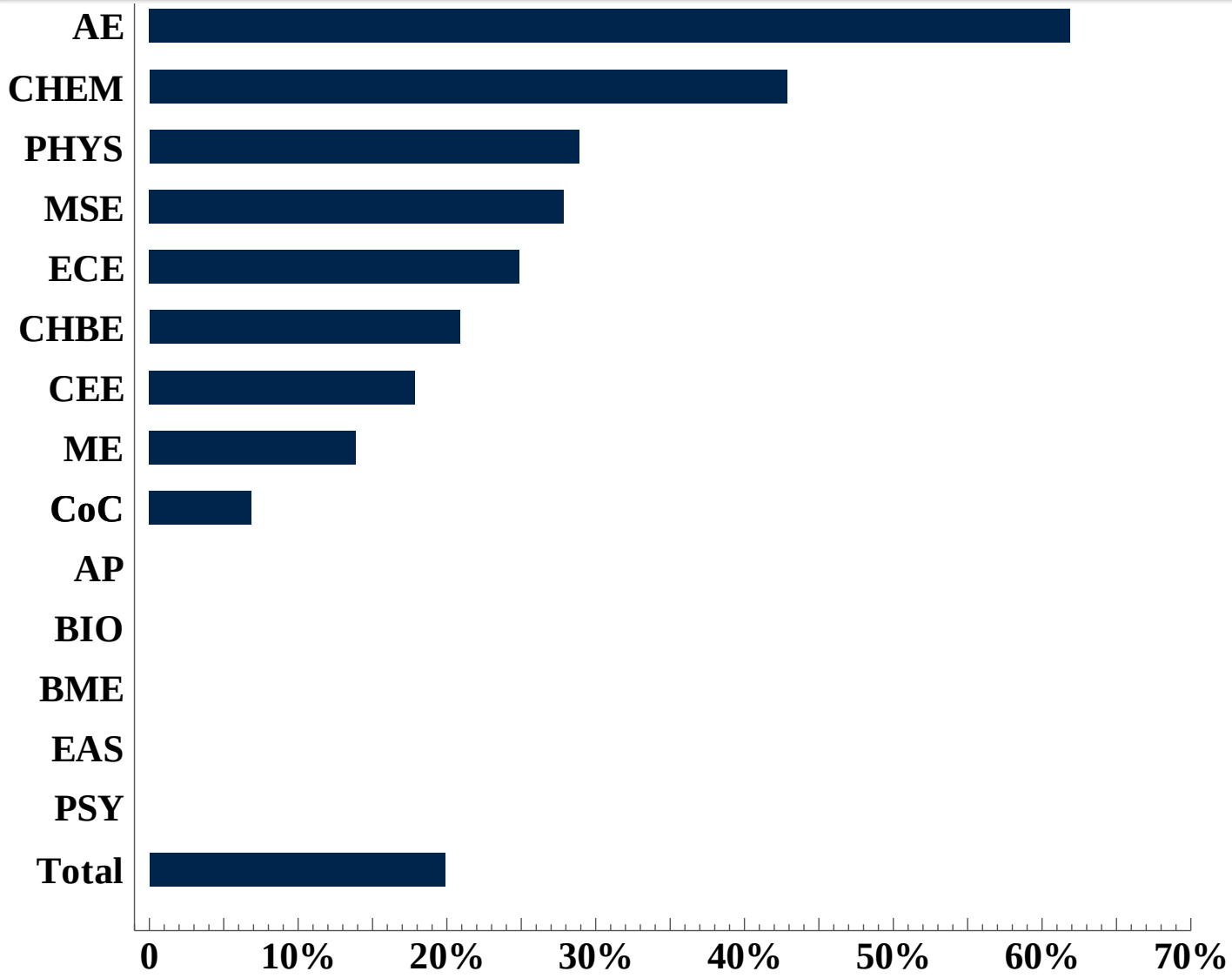
Trends:

- General lack of understanding regarding metadata
- No file naming or organizational standards
- Usually graduate students manage the research data
- Lack of understanding regarding data governance (ownership)
- Often share data using Dropbox or Google Docs
- Unaware of SMARTech (IR) and discipline-based repository services
- Everyone has lost data
- Everyone wants better versioning/revision control for data
- Would like more guidance & services for data sharing and formulating data management plans

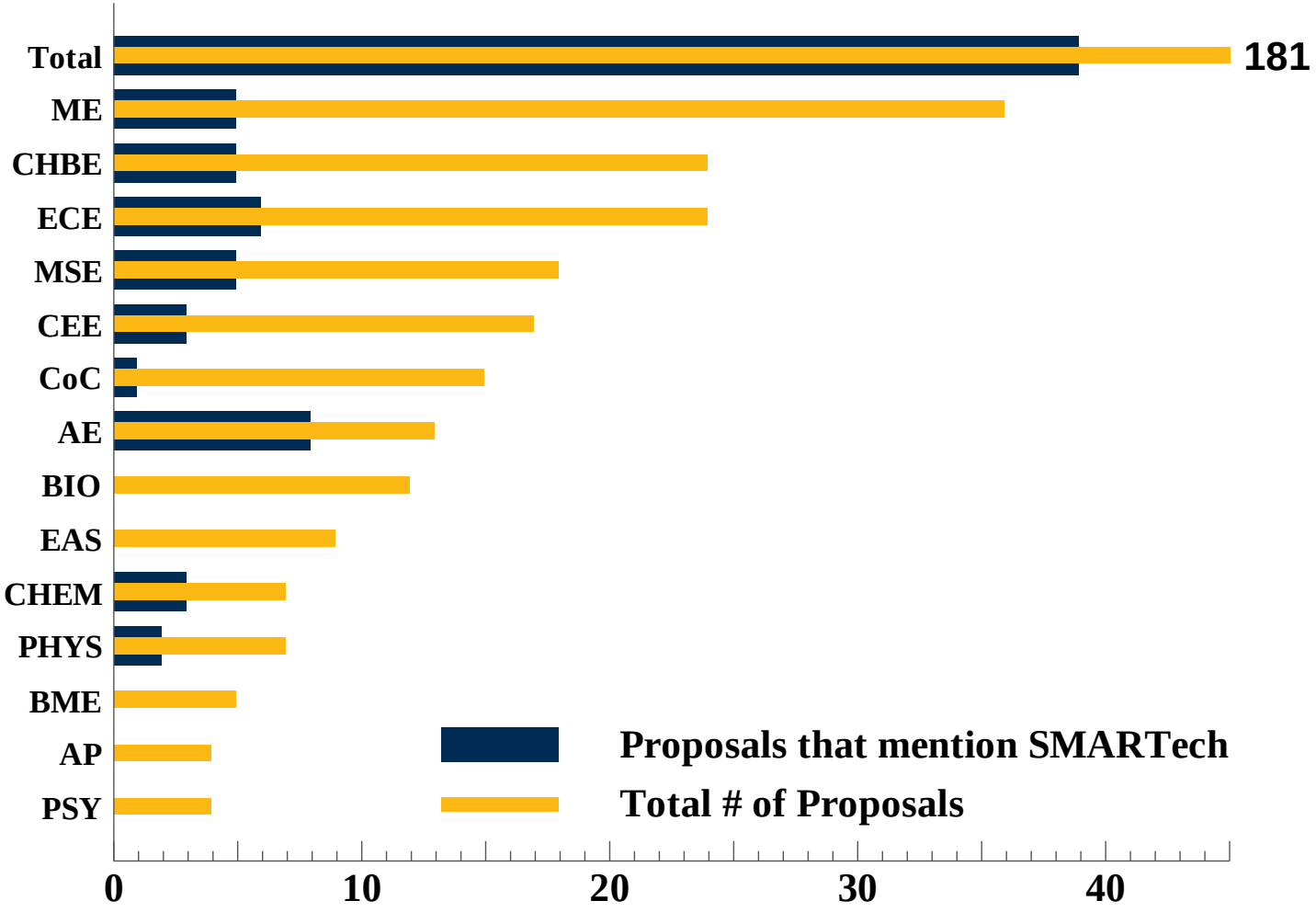
"If there were a magic wand that would make your work with research data easier, what would it allow you to do?"

Georgia Tech Library

NSF DMP Content Analysis



NSF DMP Content Analysis



2. Outreach, Training & Consultation

Improve and develop new resources & services to assist researchers with data management.

Georgia Tech Library

Outreach, Training & Advocacy

Classes

- Ethical Data Management (RCR)
- Data Stewardship (RCR)
- Writing a Data Management Plan

Presentations

- Georgia Tech Library Research Data Project
- Data Curation Strategic Planning

Articles

- Faculty Newsletter
- Research News; Research Admin Buzz
- The Whistle (GT News); *What Does GT Think?*

Data Management Web Guides & Print Materials

Georgia Tech Library

NSF DMP Consultation Services

- Research Data Management Guide
- NSF DMP Guidelines & Checklist
- Workshops - How to Write an NSF Data Management Plan
- Consultation Services - provided to researchers writing data management plans
- Contributing institution to the Data Management Plan Tool provided by UC3 and partners https://dmp.cdlib.org/
- Planning stages – internal training to develop team of DMP consultants

Georgia Tech Library

3. Partnerships

Build partnerships within the library, across the Institute, and beyond our campus boundaries to develop a comprehensive response to research data management and preservation needs.

Georgia Tech Library

Internal Partnerships

Research Data Project Team (members from multiple departments)	• Research data assessment • Data management guide & best practices
Repository Team (members from multiple departments)	• Repository policies for research data • DMP text for repository services
Scholarly Communication Collaborative (multiple departments)	• Scholarly communication advocacy, including OA week and web content
Subject Liaisons	• Data assessment & feedback on NSF DMP guidelines & workshops
User Engagement Dept	• Outreach text, graphics & emails

Georgia Tech Library

Campus Partnerships

Office of Information Technology	• Enterprise services • High Performance Computing • Data storage (not preservation)
College IT	• Data storage & backup during research
Office of Executive Vice Pres For Research	• Central administration for all research
Office of Sponsored Programs	• Services for research administration
Academic & research units on campus	• Research, research data
Legal, Faculty Senate Committees, Communications & Marketing, etc.	• Policies & services

Georgia Tech Library

IT Governance – Data Stewardship

STIC

Strategic Technology Investment Collaboration:

- review & prioritize IT programs at Institute level
- allocate resources
- increase stakeholder engagement

Data Stewardship

- maintain data integrity & reliability
- preserve confidential information
- support the validity of research
- increase visibility & re-use of data
- support compliance

Georgia Tech Library

Stewardship of Research Data Proposal

Research Data Stewardship		
Policies	IT Infrastructure	Services
▪ Data management ▪ Data ownership ▪ Data security ▪ Data retention & disposal ▪ Data sharing & reuse ▪ Intellectual property ▪ Ethical standards & relevant laws	▪ Data & metadata storage and backup ▪ Network connectivity ▪ Identity management, authentication & authorization ▪ Software development & support ▪ Data visualization ▪ Collaborative research environments ▪ High performance computing	▪ Data management planning ▪ Identification of datasets ▪ Metadata services ▪ Data & metadata formats and standards ▪ Data discovery & access ▪ Data sharing & publishing ▪ Data curation ▪ Data preservation

Georgia Tech Library

Stewardship of Research Data, proposal to GT STIC http://hdl.handle.net/1853/42091

4. Data Repository Infrastructure

Develop proof-of-concept for a library hosted research data access and preservation repository, with associated applications. Recruit campus research partners and appropriate subject librarians to develop pilot project.

Georgia Tech Library

Proof of Concept Technologies

Infrastructure Technologies

- Repository Layer - Fedora
- Web Frameworks
  - Drupal + Islandora
  - Django + EULfedora
  - Ruby on Rails + Active Fedora
- Storage Architecture
  - Local installation (Library)
  - Peachnet (Georgia Tech + UGA)
  - DuraCloud (Amazon S3)

Georgia Tech Library

Data Curation Roles

Thanks to Gail Steinhart, Cornell University Library

Susan Wells Parham
Head, Scholarly Communication
& Digital Curation Services
Georgia Tech Library
(404)894-4522
susan.parham@gatech.edu