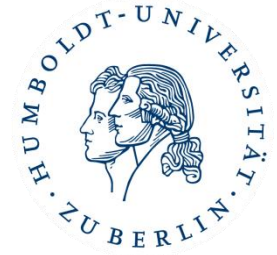


HUMBOLDT-UNIVERSITÄT ZU BERLIN



An Integral Approach to Support Research Data Management at the Humboldt-Universität zu Berlin

Malte Dreyer
Director IT Center

Humboldt-Universität zu Berlin



Founded in 1810

Faculties: 9

Students: 35,000

Degree Programs: 190

Professorships: 419(483)

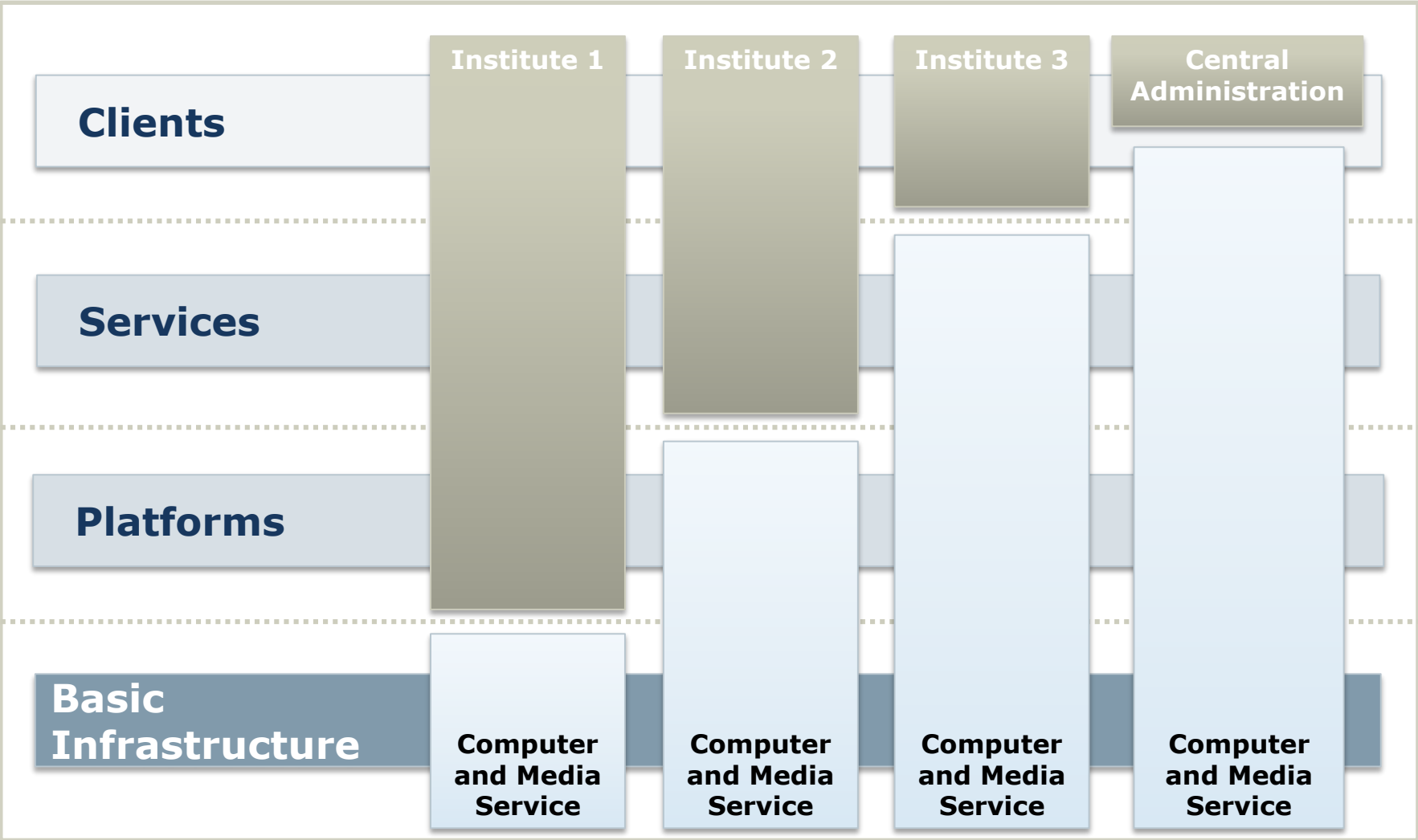
Academic Staff: 2,000

Non-Academic: 1,500

3 Campuses, ~180 Buildings



Central IT (Computer and Media Service)



Research Data Management

REWARD £1,000
cash

My laptop and external hard drive
were removed from my office [REDACTED]
[REDACTED], on August 29.

If you have found the external
hard drive, PLEASE CALL ME
([REDACTED])
[REDACTED]

The drive contains many years of my
teaching and research files, so I need
it desperately.

When opened, you will see that there is very little disk space left. I have had this USB for 4 years. It has all of my academic work at Ryerson (which I am still working towards completing for graduation this fall), plus personal photographs and professional resumes etc.

This USB may not be of significance to you, but it is extremely important to me. If you have found it, please kindly contact me to

**\$\$\$LOST: MY
ENTIRE THESIS**

ONE BLACK RISD LOGO
BINDER FILLED WITH
NEGATIVES

TWO LACIE HARD DRIVES:
ONE ORANGE, ONE BLACK

REWARD: \$\$\$\$\$, ALL THE
ALCOHOL YOU CAN DRINK
ATHOMAS01@G.RISD.EDU

National Science Foundation (NSF) Biographical Sketch(es)

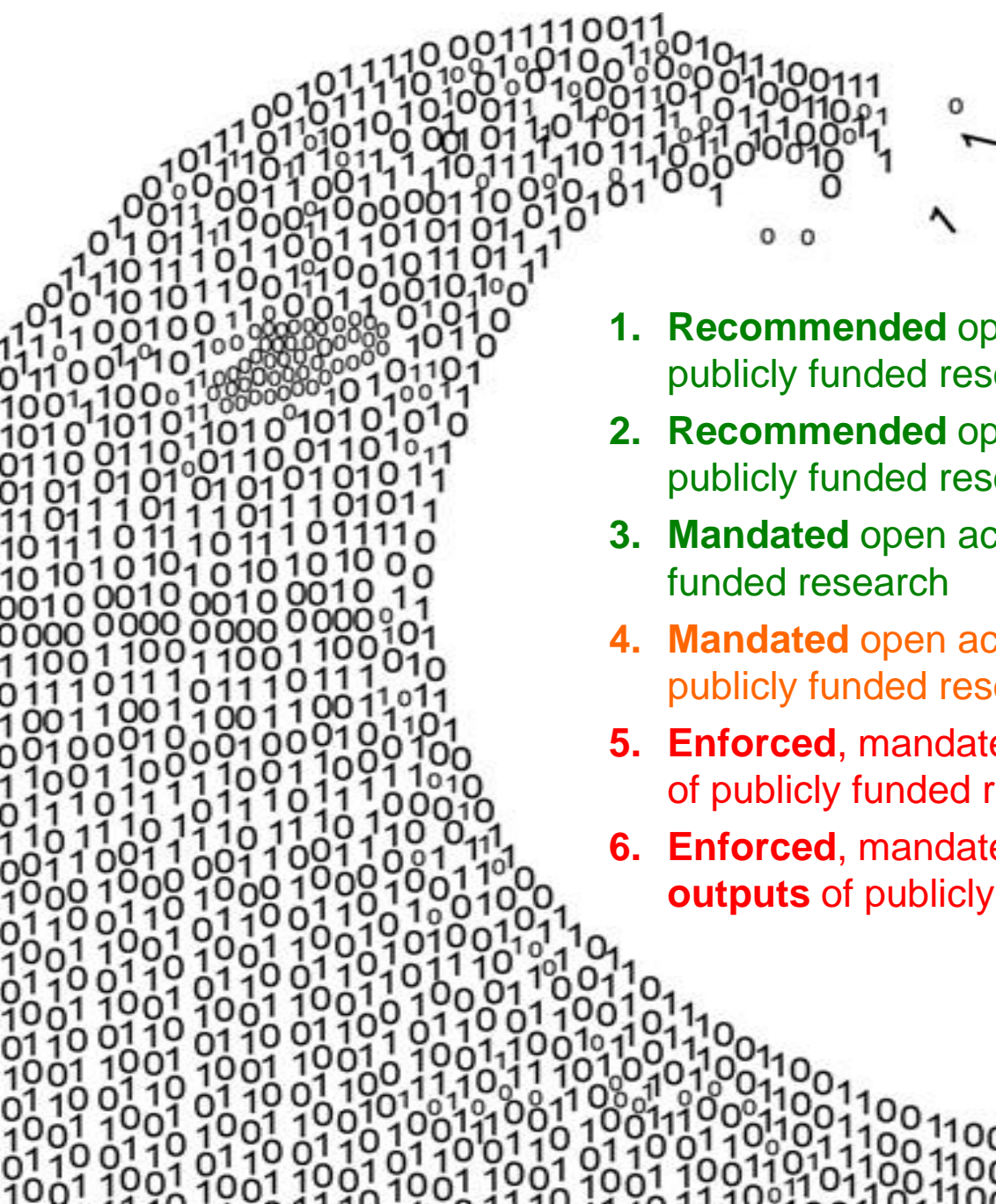
Biographical Sketch(es), has been revised to rename the “Publications” section to “Products” and amend terminology and instructions accordingly.

(c) **Products**

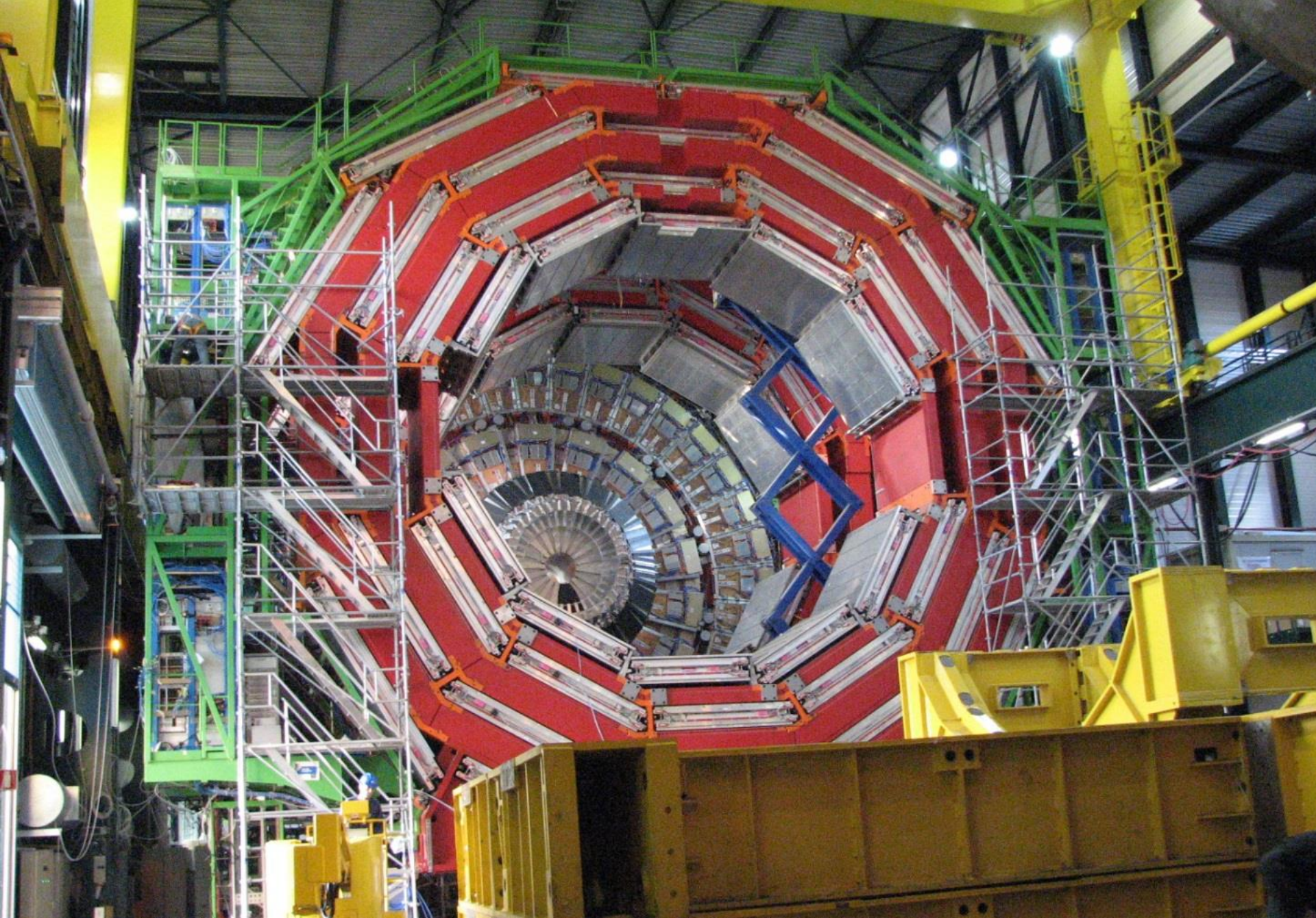
A list of: (i) up to five products most closely related to the proposed project; and (ii) up to five other significant products, whether or not related to the proposed project. **Acceptable products must be citable and accessible** including but not limited to **publications, data sets, software, patents, and copyrights**. Unacceptable products are unpublished documents not yet submitted for publication, invited lectures, and additional lists of products. Only the list of 10 will be used in the review of the proposal.

Each product must include full citation information including (where applicable and practicable) names of all authors, date of publication or release, title, title of enclosing work such as journal or book, volume, issue, pages, website and Uniform Resource Locator (URL) or other Persistent Identifier.





1. **Recommended** open access to **scholarly papers** of publicly funded research
2. **Recommended** open access to all **digital outputs** of publicly funded research
3. **Mandated** open access to **scholarly papers** of publicly funded research
4. **Mandated** open access to all **digital outputs** of publicly funded research
5. **Enforced**, mandated open access to **scholarly papers** of publicly funded research
6. **Enforced**, mandated open access to all **digital outputs** of publicly funded research



By Muriel (originally posted to Flickr as Muon Barrel) [CC-BY-SA-2.0 (<http://creativecommons.org/licenses/by-sa/2.0>)], via Wikimedia Commons



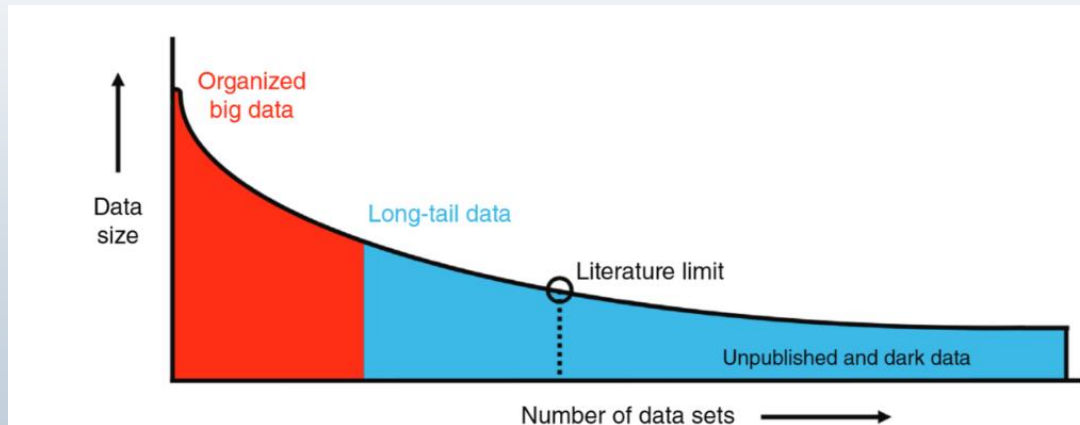
By Hannes Grobe (Hgrobe 06:16, 26 April 2006 (UTC)), Alfred Wegener Institute for Polar and Marine Research (Own work) [CC-BY-SA-2.5 (<http://creativecommons.org/licenses/by-sa/2.5>)], via Wikimedia Commons

Small Data

„Long Tail of Research Data“

The 2011 survey by *Science*, found that 48.3% of respondents were working with datasets that were less than 1GB in size and over half of those polled store their data only in their laboratories.

Science 11 February 2011: Vol. 331 no. 6018 pp. 692-693 DOI: 10.1126/science.331.6018.692



Source: Ferguson et al. (2014): Big data from small data: data-sharing in the 'long tail' of neuroscience. DOI: 10.1038/nn.3838

Why the Data Center Should Care?

Google



Instagram

„Google Graveyard“

Name	Days of Operation	Name	Days of Operation
iGoogle	3.106	Google Friend Connect	1.400
Reader	2.824	Knol	1.377
Google Video	2.769	Google Code Search API	1.370
Google Desktop	2.539	Google Health	1.340
Google Bookmarks Lists	2.252	Places Directory app (Android)	1.308
Picasa for Linux	2.162	Google Maps API for Flash	1.206
Google Notebook	2.006	Wave	1.095
Google Sync for BlackBerry	1.858	Buzz	980
Image Labeller	1.827	Classic Plus	884
Google Pack	1.706	Google Listen	883
Google Answers	1.702	Google Sidewiki	803
Google Spreadsheet			
Gadgets	1.645	Fast Flip	730
EAS sync for Gmail	1.603	Google Search Timeline	640
Google Talk Chatback	1.589	Google News Badges	458
Jaiku	1.583	+1 Reporting for webmaster tools	458
Google Video for Business	1.521	Google One Pass	367
Google Adsense for Feeds	1.492	Google Aardvark	365
Google Gears	1.468	Google Lively	175
Google Apps for Teams	1.461		

Google Terms of Service (Apr. 2015)

These conditions have been present since March 2012.

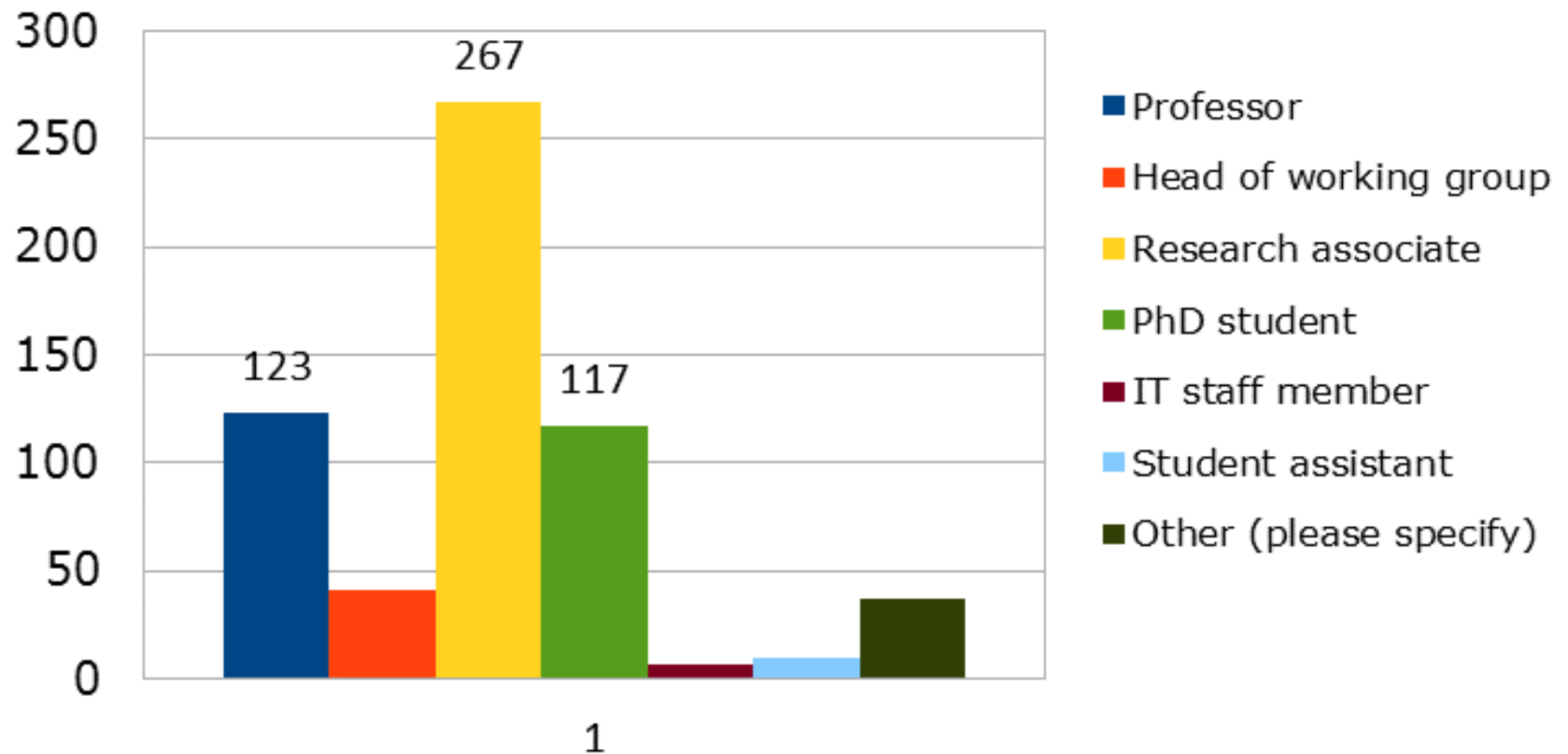
When you upload, submit, store, send or receive content to or through our Services, you give Google (and those we work with) a worldwide license to use, host, store, reproduce, modify, create derivative works (such as those resulting from translations, adaptations or other changes we make so that your content works better with our Services), communicate, publish, publicly perform, publicly display and distribute such content. The rights you grant in this license are for the limited purpose of operating, promoting, and improving our Services, and to develop new ones. This license continues even if you stop using our Services (for example, for a business listing you have added to Google Maps). Some Services may offer you ways to access and remove content that has been provided to that Service. Also, in some of our Services, here are terms or settings that narrow the scope of our use of the content submitted in those Services. Make sure you have the necessary rights to grant us this license for any content that you submit to our Services.

In short, you retain IP over the content, but grant Google and those they work with the rights to use your content to develop and promote Google services.

What To Do?

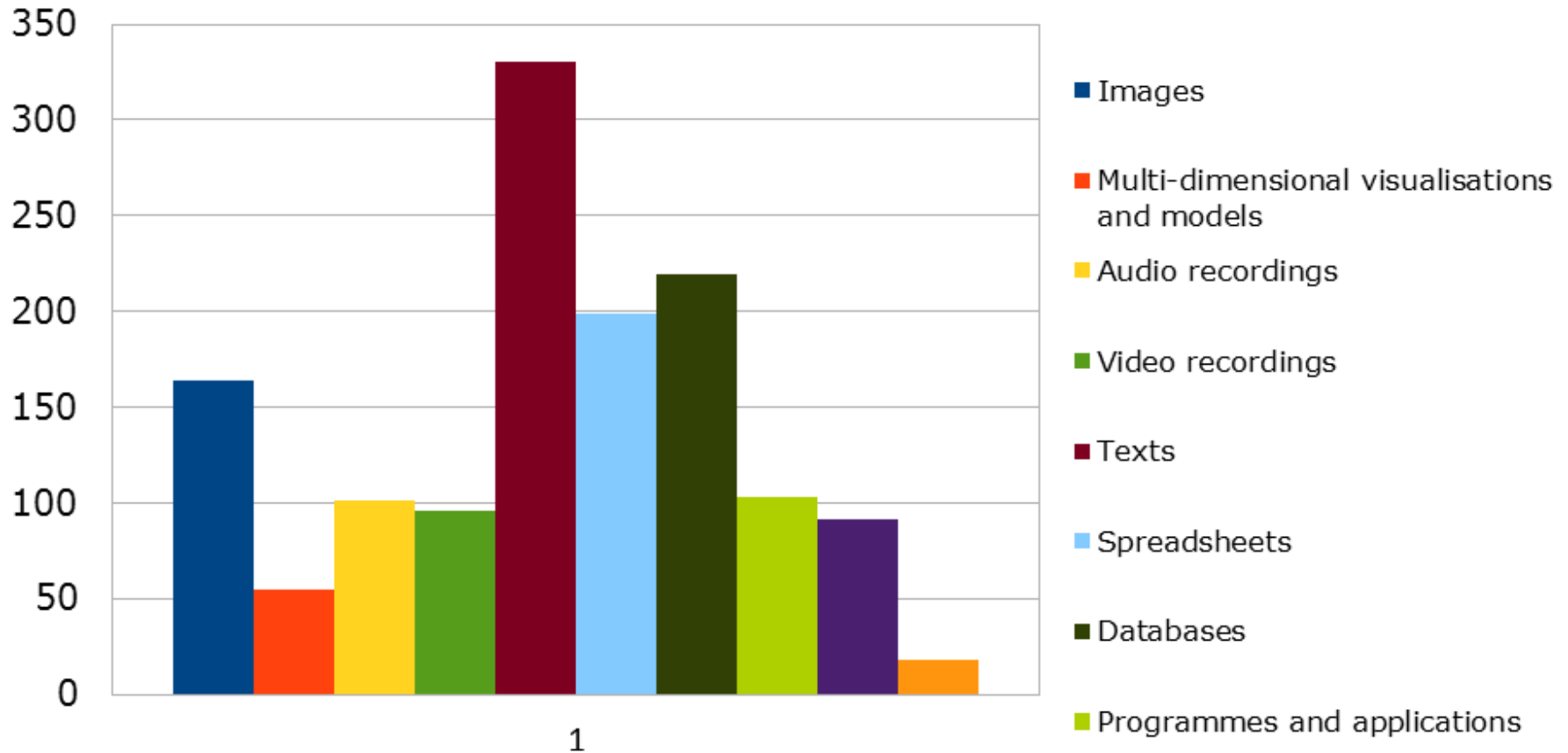
... We Conducted a Survey in 2013

Respondents groups



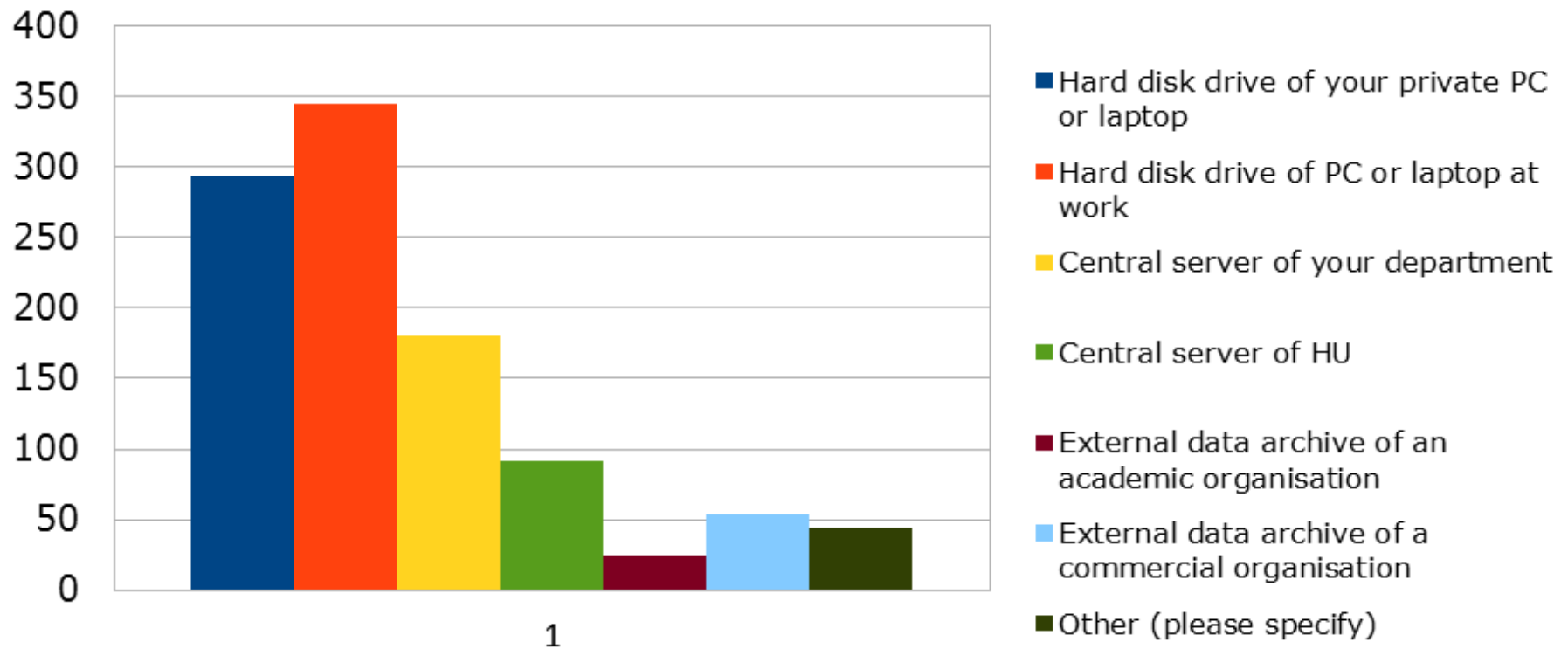
Please indicate data types more specifically.

Research data types



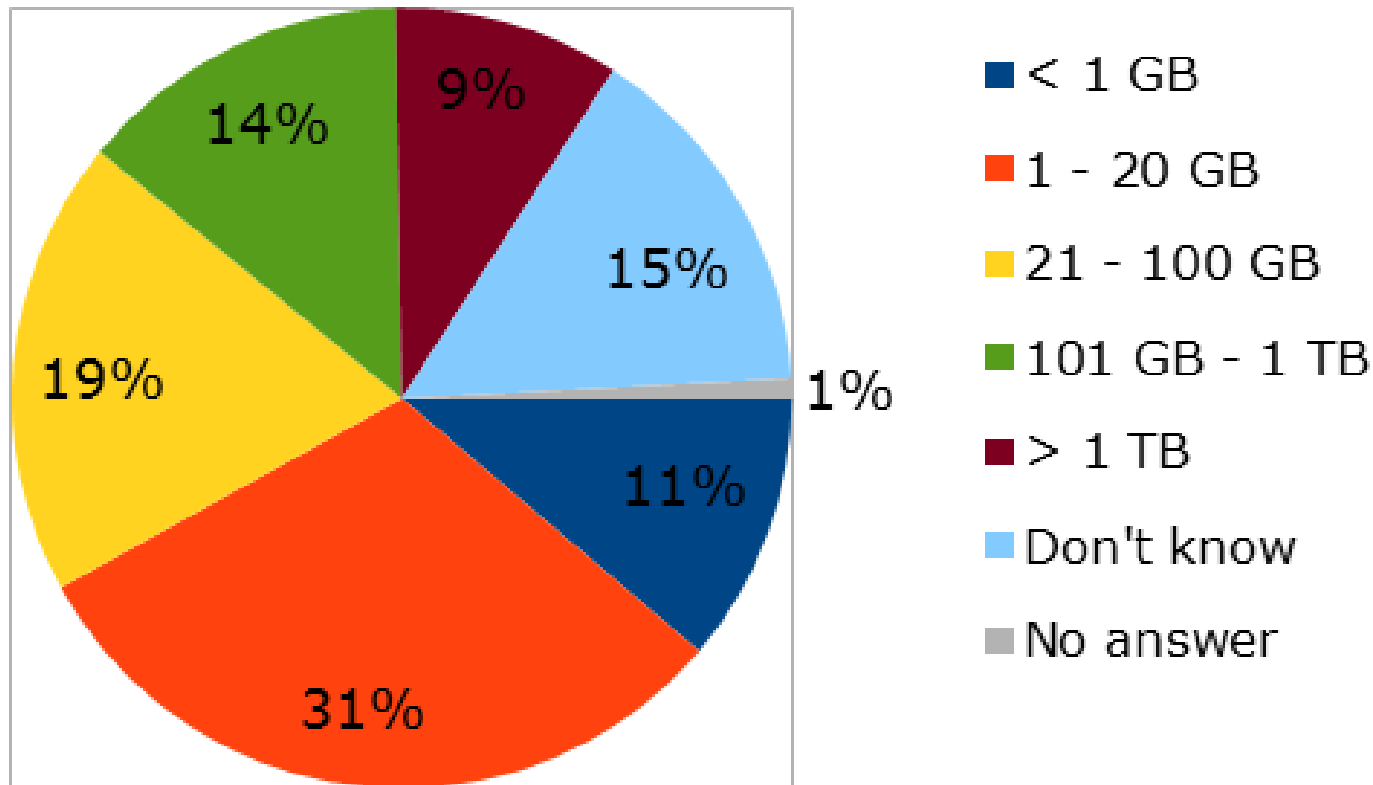
Where do you store your research data?

Storage places



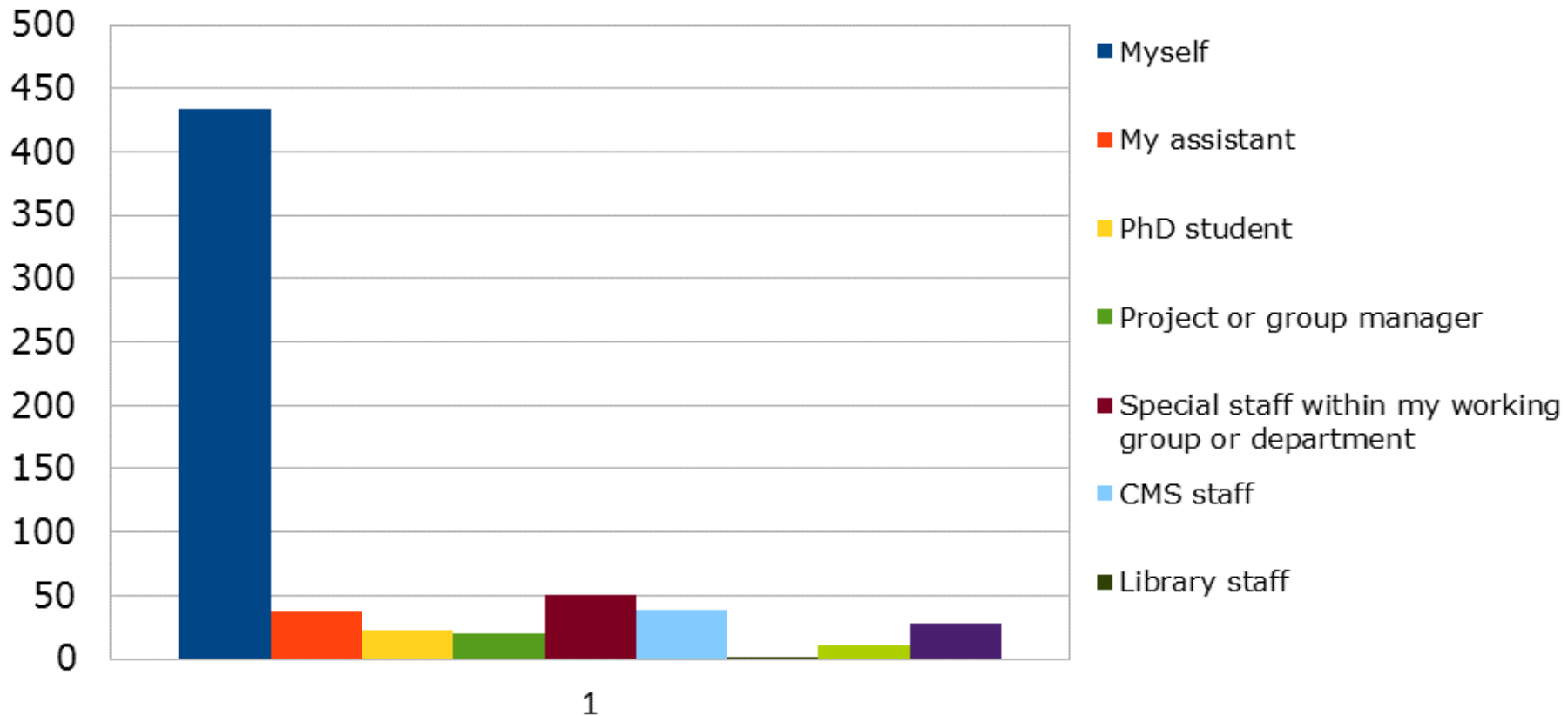
Please estimate how much research data you currently hold in terms of disk space.

Overall size of research data



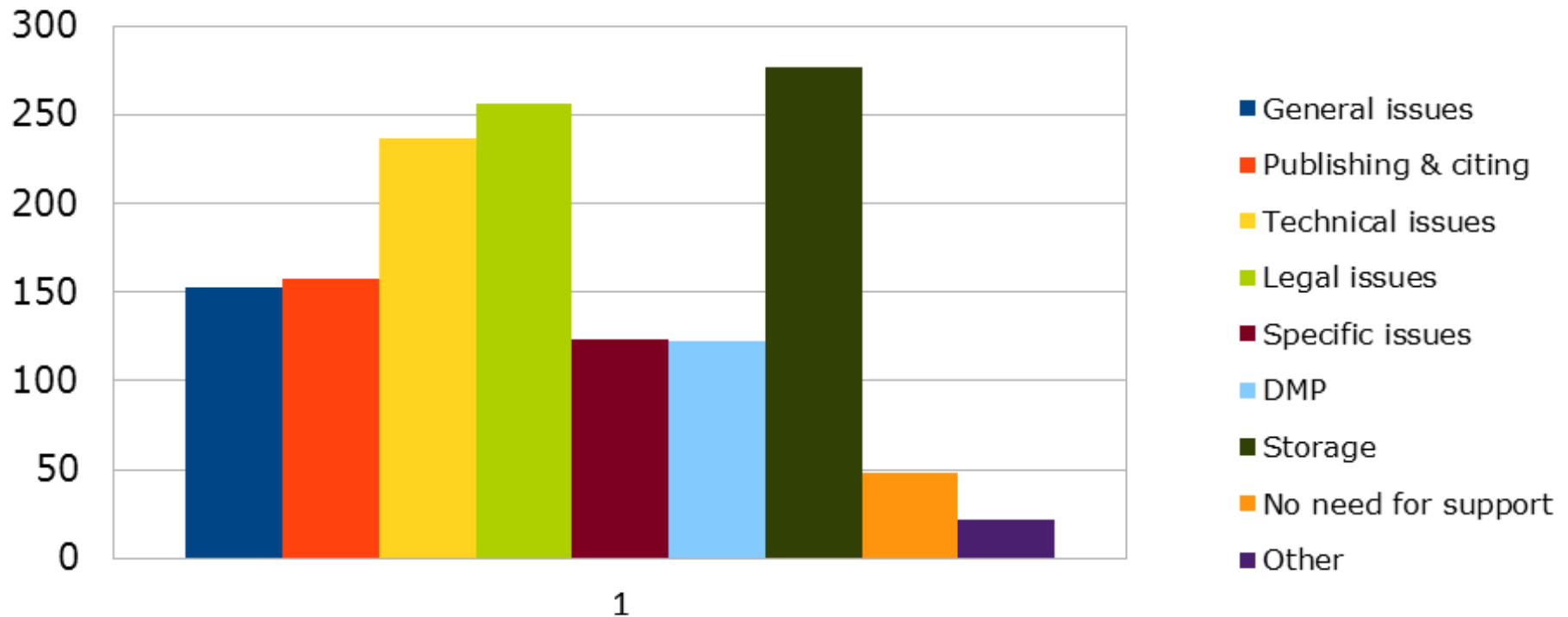
Who is currently responsible for storing, back up or archiving your research data?

Responsible for research data

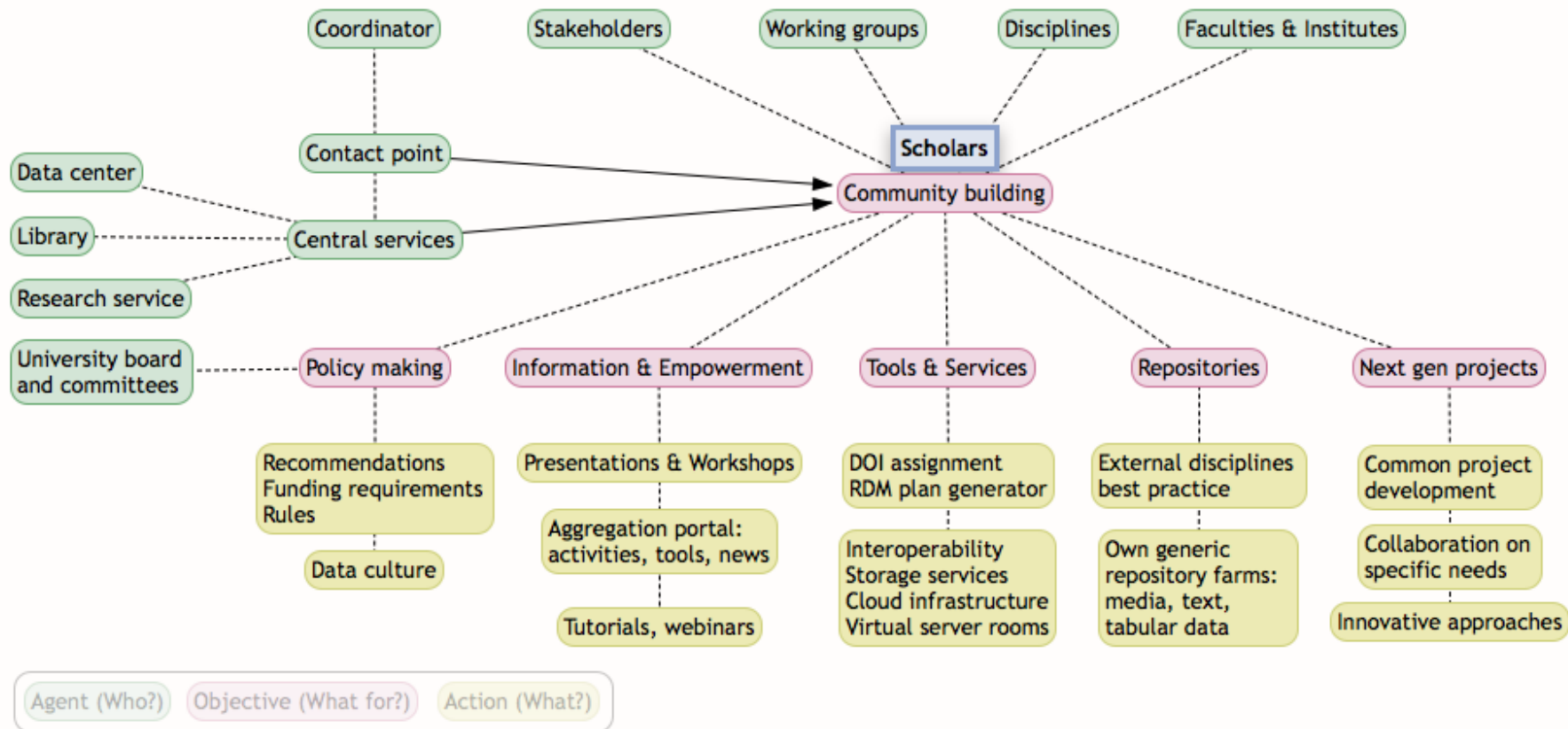


What support or services do you wish to have at HU?

Support & services needed



Conceptual Map of our Research Data Management Approach



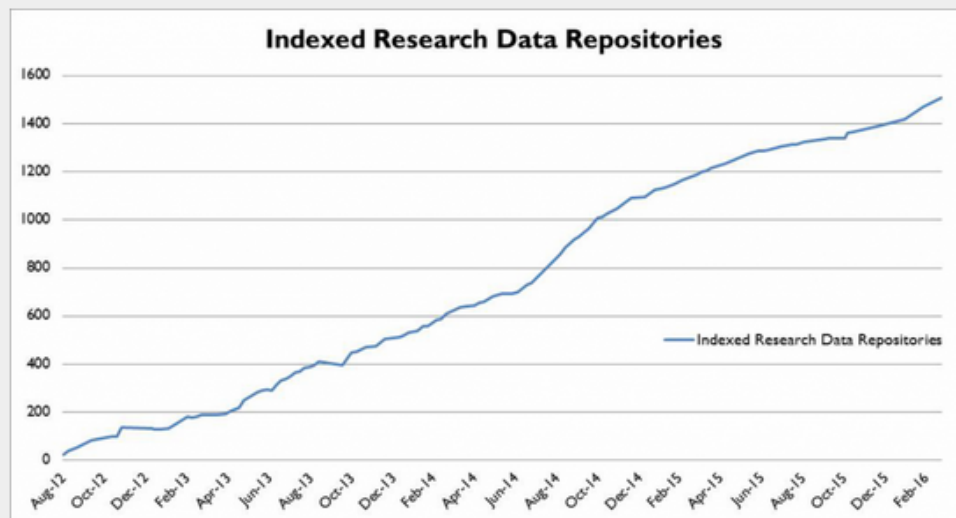
Building Blocks

- Building tools and services
- Communication and community building
- Developing strategic projects
- One central coordinator
- Online tutorials
- Presentations and workshops
- Repository services

re3data.org Reaches a Milestone & Begins Offering Badges

Posted on [April 13, 2016](#) by [re3data.org team](#)

re3data.org has reached a milestone of identifying and listing 1,500 research data repositories, making it the largest and most comprehensive registry of data repositories available on the web. It has grown steadily since its launch four years ago to cover a wide range of disciplines from around the world.



SEARCH

A SERVICE BY



PARTNERS





Computer and Media Service
Research data management

Welcome

Work with research data

Share research data

Information

Support

Malfunction information

Students | Staff

Humboldt-Universität zu Berlin | Computer and Media Service | Research data management

Welcome

Welcome to the research data management website of Humboldt-Universität zu Berlin - a joint initiative of the central units [Computer and Media Service \(CMS\)](#), [Research Service Centre](#) as well as [University Library](#) and [Vice President for Research](#) of the university.

Guidance for creating a datamanagement plan in Horizon 2020 available, more at [Create datamanagement plan](#).

Work with research data

[Find research data](#)

[Cite research data](#)

[Create datamanagement plan](#)

[Secure research data](#)

Information

[Relevance of research data](#)

[Principles of the university](#)

[Discipline-specific policies](#)

[Further reading](#)

Share research data

[Find repository](#)

[Documentation and metadata](#)

[Select license](#)

[Persistent identification](#)

Support

[Training](#)

[Material](#)

[Contact](#)

DE EN

News

[5th DINI-nesstor Workshop Tools for Research Data](#)

May 18, 2016

[ResourceSpace Community Meeting](#)

May 18, 2016

[Workshop: Research Data Management for Agricultural Scientists and Biologists](#)

Apr 18, 2016

[More news...](#)

Contact

Coordinator
research data management

Kerstin Helbig

kerstin.helbig@cms.hu-berlin.de

+49(0)30 2093-70072

Helpdesk

Pamela Aust

pamela.aust.1@ub.hu-berlin.de

+49(0)30 2093-99706

Technical information

Dennis Zielke

dennis.zielke@cms.hu-berlin.de

+49(0)30 2093-70004

Open access publications

[Sandra Golda](#)



Computer and Media Service
Research data management

Welcome

Work with research data

Share research data

Information

Support

Malfunction information

Students | Staff

Humboldt-Universität zu Berlin | Computer and Media Service | Research data management

Welcome

Welcome to the research data management website of Humboldt-Universität zu Berlin - a joint initiative of the central units Computer and Media Service (CMS), Research Center Centre as well as University Library

DE EN

News

5th DINI-nestor Workshop Tools for

Work with research data

Find research data

Cite research data

Create datamanagement plan

Secure research data

Share research data

Find repository

Documentation and metadata

Select license

Persistent identification

Information

Relevance of research data

Principles of the university

Discipline-specific policies

Further reading

Support

Training

Material

Contact

Research Data Management

Trainer: Kerstin Helbig/ Ulrike Schenk/ Pamela Aust

Date: 09.12.2016, 4 p.m. – 6 p.m.

Target Group: Candidates in all years of PhD studies (open to all doctoral candidates)

OBJECTIVES

Research data management is an integral part of good scientific practice. It supports later data sharing, which can lead to increased citation and reuse. In addition, more and more funding organizations expect a data management plan during the application process. This workshop informs about data management planning and services at Humboldt-Universität zu Berlin. Technical, legal and financial aspects of research data management are covered. In separate discipline specific groups metadata as well as appropriate research data repositories and archives are presented and discussed.

CONTENT

- What is research data?
- Why research data management? What are the requirements of funding organizations such as DFG, BMBF and EU?
- How does successful research data management look like? How do I write a data management plan? Which aspects are important?
- Research data management at Humboldt-Universität zu Berlin (research data policy, services)
- Discipline specific metadata

Tools & Services: Cloud Storage

Files



Mine



Shared



Groups

Tools



Starred



Activities



Devices

Share Admin



Libraries



Folders



Links

Mine / FoDaShare / MyData /

Upload

New Folder

New File

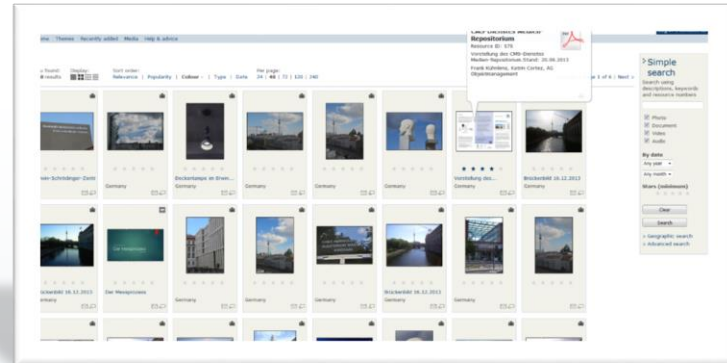
Share



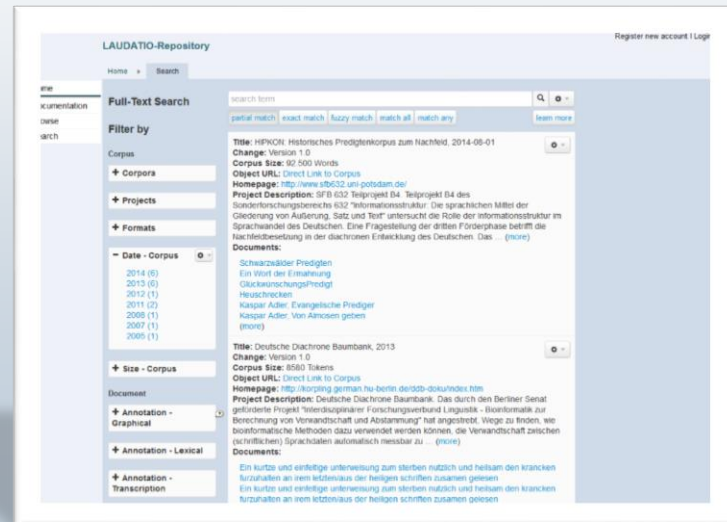
<input type="checkbox"/>	Name ▼	Size	Last Update
<input type="checkbox"/>	shapefiles		2014-05-02
<input type="checkbox"/>	raster		2013-03-31
<input type="checkbox"/>	projects		2014-05-02
<input type="checkbox"/>	grassdata		2009-08-27
<input type="checkbox"/>	gps		2009-11-30
<input type="checkbox"/>	gml		2008-10-08
<input type="checkbox"/>	csv		2008-10-09
<input type="checkbox"/>	climate		2009-04-30
<input type="checkbox"/>	README.html	3.5 KB	2013-04-03

Tools & Services: Repositories

Semi-Tailored



Tailored





Kimberlit
Deutschland
20



Kupfererz
Deutschland
19



Porphyrit
Deutschland
18



Porphyritbrekzie
Deutschland
17



Porphyritbrekzie
16



Porphyritbrekzie
15



Porphyritbrekzie
Deutschland
13



Porphyritbrekzie
12



Porphyritbrekzie
11



Porphyrit
10



Porphyrit
9



Kupferschiefer mit...
Deutschland
7



Freigoldführender Quarz
3



Schwefel in Gips
2



Schwefel in Gips
1



Freigoldführender Quarz
5



Freigoldführender Quarz
4



Steinsalz
21



Featured Collections

Under this heading one can find the main themes under which the digitized data is organized.



Simple search

Search using descriptions, keywords and resource numbers

- Photo
- Document
- Video



Welcome to

The project on Universitat zu B and Ms. Klaudia from the Ditson knowledge and in the form of Germany and S the publication into the life of importance to u archives. We se published book

The digital data various mission include: the Ber Library Archive well as field res

Original diary of the Missionary Carl Hoffmann. Part 9

[Previous result](#) | [View all results](#) | [Next result](#)



Resource tools		
File information	File size	Options
Original PDF File	68.5 MB	Download
Low resolution print 1708 x 2000 pixels (3.42 MP) 14.5 cm x 16.9 cm @ 300 PPI	879 KB	Download
Screen 683 x 800 pixels (0.55 MP) 5.8 cm x 6.8 cm @ 300 PPI	194 KB	Download
Preview Full screen preview	194 KB	View

> Share

Resource details

Resource ID 489	Access Open	Category Files, UNISA Archive (Files)	Source UNISA Archive Type of Source Diary	Location South Africa DOI 10.18450/hoffmanncollec tion/489	Created by Hoffmann Description Original diary of the Missionary Carl Hoffmann. Part 9	Language of Source German Usage terms Creative Commons Namensnennung 4.0 International Lizenz
--------------------	----------------	--	--	--	---	--

Simple search

Search using descriptions, keywords and resource numbers

- Photo
- Document
- Video
- Audio

By date

Any year

Any month

[Clear](#)

[Search](#)

- > [Geographic search](#)
- > [Advanced search](#)

Log in

Username

Password



Featured Collections

Categories



Search

- All resources
- Manuscript
- Manuscript page

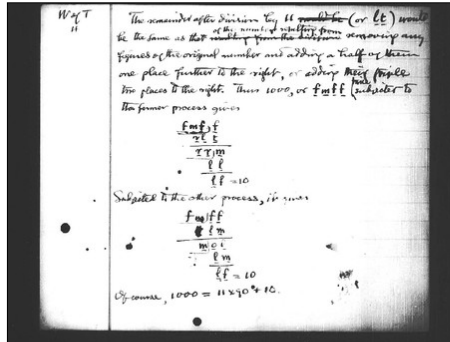


Digital Peirce Archive

The *Peirce Archive* is an image-based archi collaboration between the DFG cluster of e Articulation" funded by the Volkswagen Fou span more than five decades and touch on

On the Ways of Thinking of Mathematics

[Previous result](#) | [View all results](#) | [Next result](#)



Resource tools

File information	File size	Options
Original JPG File 850 × 648 pixels (0.55 MP) 7.2 cm × 5.5 cm @ 300 PPI	212 KB	Download

[> Share](#) [> Transform](#)

Search

- All resources
- Manuscript
- Manuscript page

manuscript number

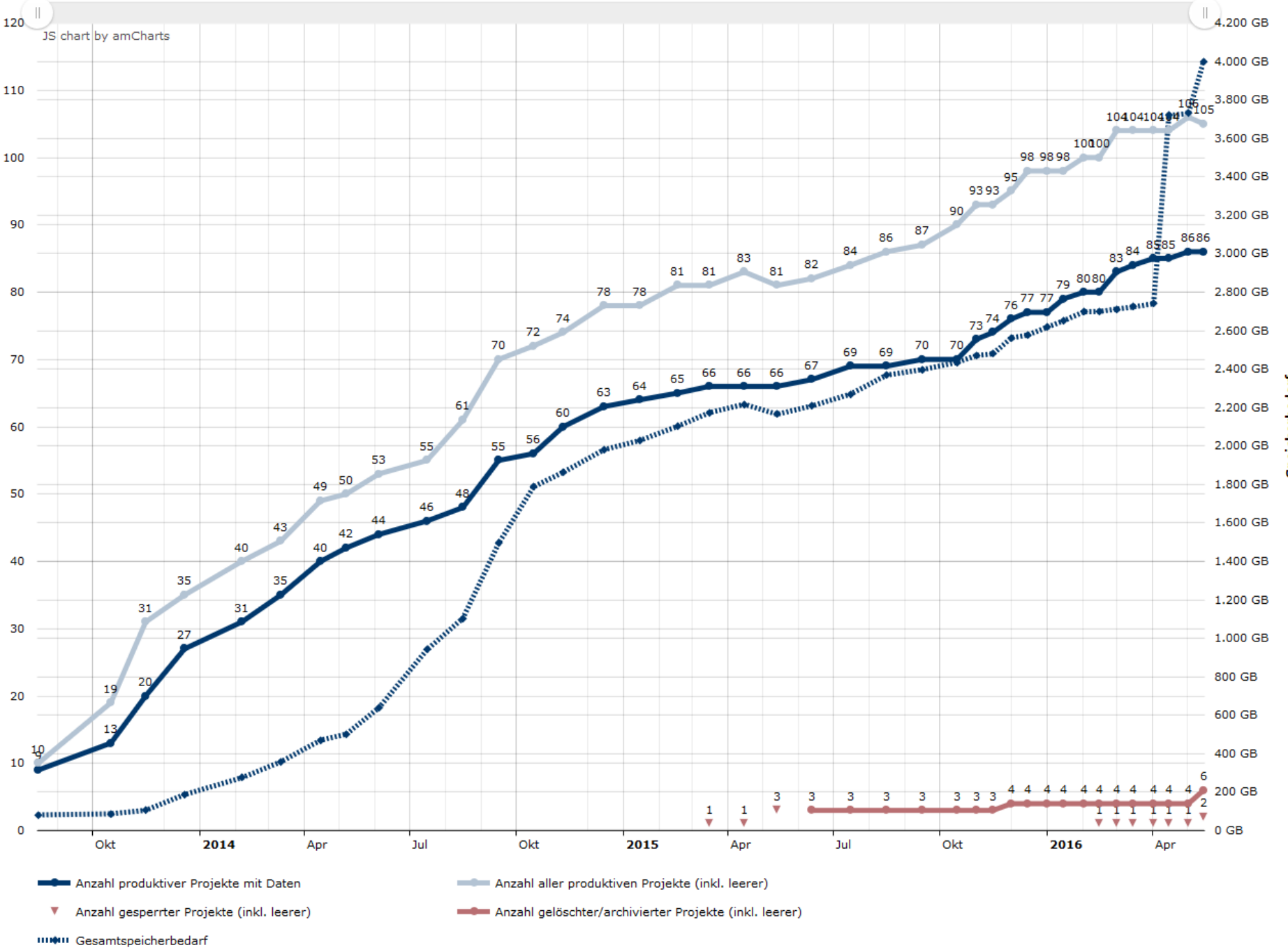
[Clear](#)

[Search](#)

[> Advanced search](#)

Manuscript Metadata

Resource ID	Access	Contributed by	date (Robin)	type of material	description	general index
2795	Open	Frederik Wellmann	c. 1901/1902?	A. MS.	-	Decimal system, Mathematics, numerical Notation, notation and analysis (see also Decimal system Quantity Secundal system Sextal system), Secundal system of enumeration



■ Anzahl produktiver Projekte mit Daten

■ Anzahl aller produktiven Projekte (inkl. leerer)

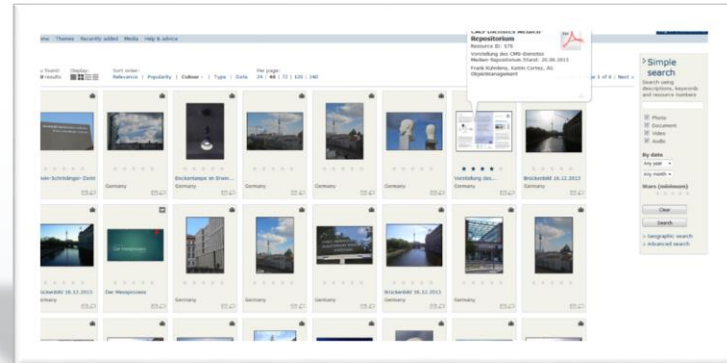
▼ Anzahl gesperrter Projekte (inkl. leerer)

● Anzahl gelöschter/archivierter Projekte (inkl. leerer)

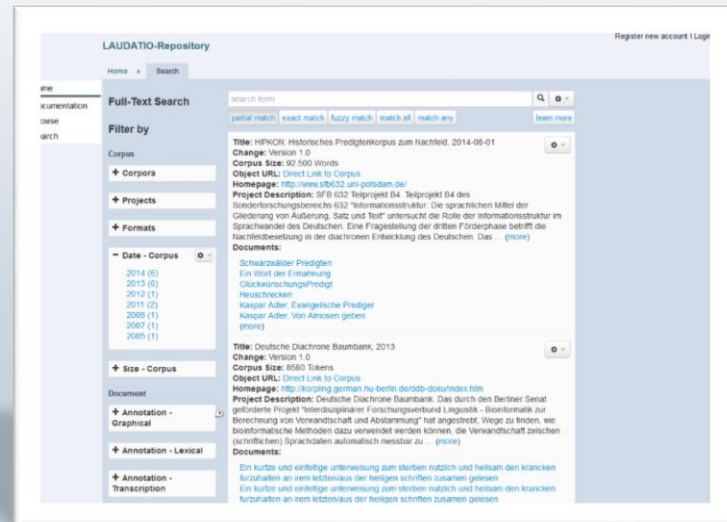
▒ Gesamtspeicherbedarf

Tools & Services: Repositories

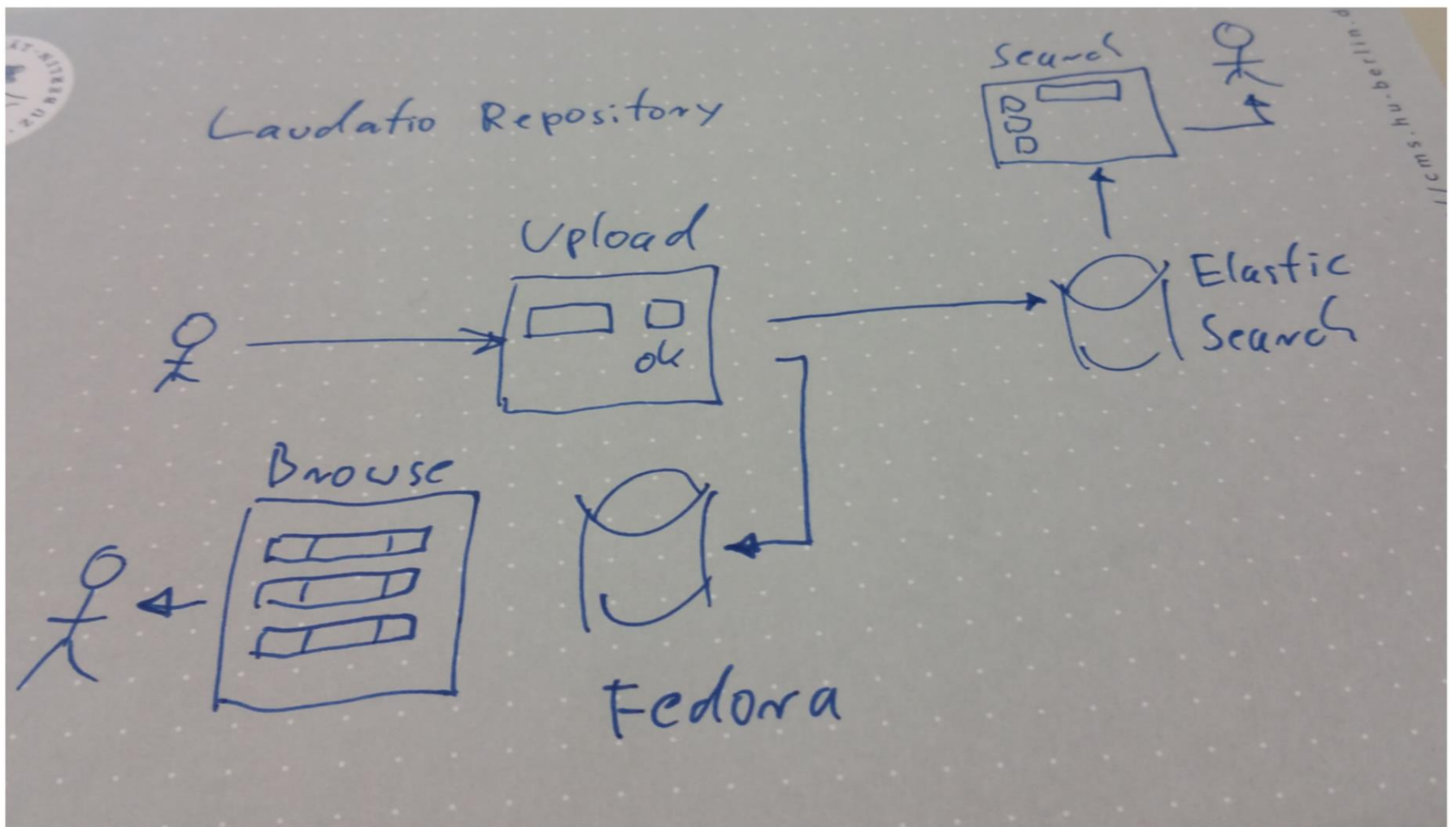
Semi-Tailored



Tailored



Laudatio: Open Access Repository for Historical Linguistic Data



Self-Service



The NIST Definition of Cloud Computing

(US-National Institute of Standards and Technology SP 800-145)

Cloud computing is a model for enabling ubiquitous, **convenient, on-demand network access to** a shared pool of **configurable computing resources** (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released **with minimal** management effort or **service provider interaction**.

Project ^

Compute ^

Overview

Instances

Volumes

Images

Access & Security

Network v

Object Store v

Orchestration v

Database v

Data Processing v

Identity v

Overview

Limit Summary



Instances
Used 1 of 10



VCPUs
Used 2 of 20



RAM
Used 4GB of 50GB



Volumes
Used 1 of 10



Volume Storage
Used 1GB of 1000GB

Usage Summary

Select a period of time to query its usage:

From:

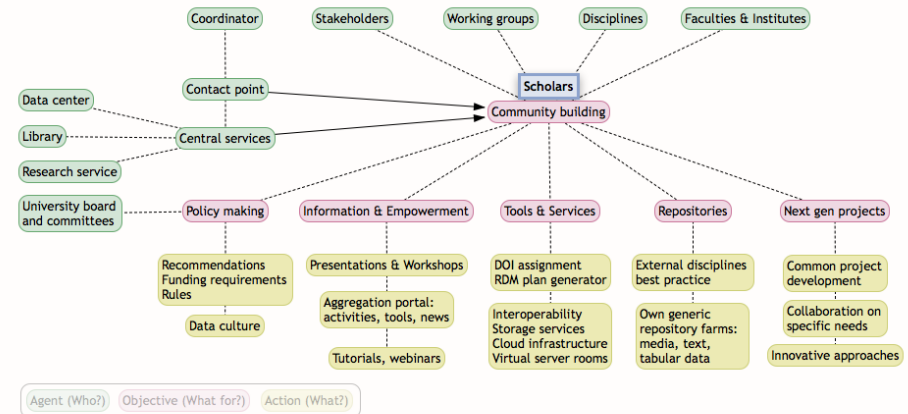
To:

The date should be in YYYY-mm-dd format.

Active Instances: 1 Active RAM: 4GB This Period's VCPU-Hours: 400.30 This Period's GB-Hours: 8005.97 This Period's

Conclusions

- RDM has many facets
- Each discipline has very unique formal and informal practices
- There are many suitable repositories available
- There still is a high demand for local solutions
- RDM is a long-lasting topic
- Communication and community building is key





.....Thanks.....

malte.dreyer@cms.hu-berlin.de



Questions?

malte.dreyer@cms.hu-berlin.de