



**Carnegie
Mellon
University**

*The Carnegie Mellon University Cloud Lab:
Automating and Democratizing Science
for the Future*

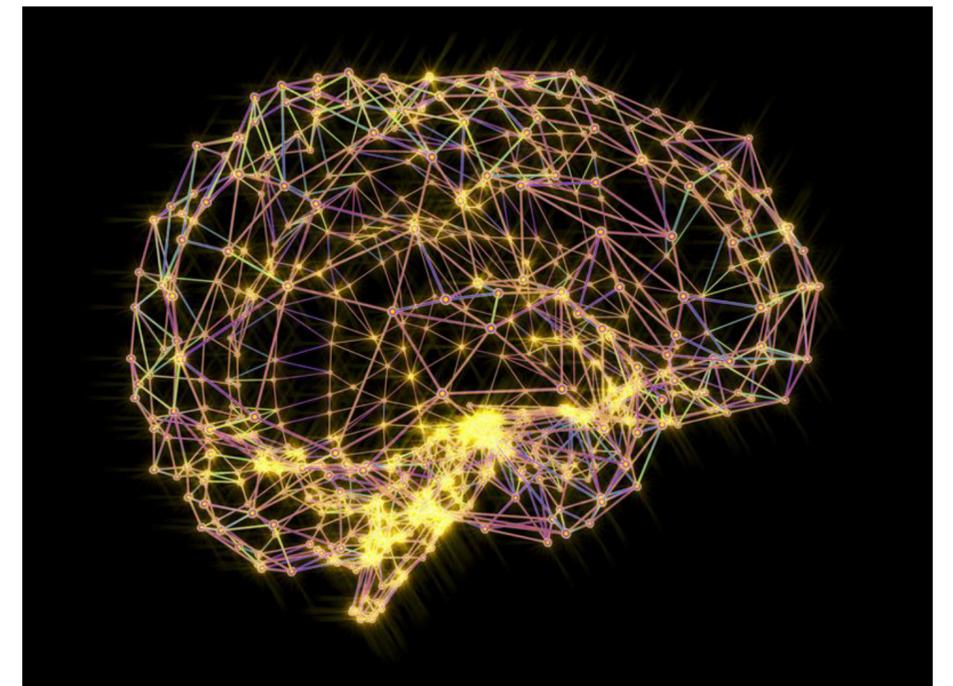
THE FUTURE OF SCIENCE IS AUTOMATED SCIENCE

Rebecca Doerge, Glen de Vries Dean of the Mellon College of Science

An Unprecedented Era of Science

The future of science is being driven by:

- New tools that are collecting **massive amounts of data** — about the natural world, the built environment, human biology and behaviors, and more
- Powerful **AI technologies** that are essential to analysis, connection and application of massive data streams
- People working and training together **across disciplinary boundaries** using flexible workspaces and shared resources



Much of life sciences research is currently limited by speed and scale of data that can be acquired

CMU Cloud Lab

- Remote controlled laboratory
- Automated instrumentation + technicians
- Everything traceable → reproducible science
- On site synthesis and/or samples sent to lab
- Based on **Emerald Cloud Lab** facility/partnership

<https://www.emeraldcloudlab.com/>

CMU Cloud Lab

- Instrument working group
- Undergrad + grad training
- Faculty research
- Business planning
- Grant applications
- Publications



User-designed scripts allow complex workflows



Example Workflow 1

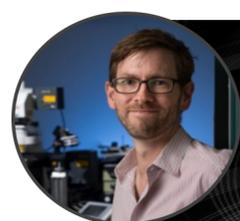
1. Solid Phase Peptide Synthesis
2. Reverse-phase prep HPLC purification
3. QC by MALDI mass spectrometry & analytical HPLC re-injection
4. Bioconjugation chemistry on a liquid handler
5. Fluorescence kinetics binding assay in a cell-free environment

Example Workflow 2

1. Precipitation conditions screening on a liquid handler
2. Validate by sample imaging
3. Isolate powder by filtration and wash
4. Dry on spin vacuum
5. Further characterization on power XRD

Example Workflow 3

1. DNA oligomer synthesis
2. Purification by ion exchange HPLC with fraction collection
3. MALDI all fractions and pool desired aliquots
4. Freeze and lyophilize fractions
5. UV melting experiments



Proof of Concept in Progress

Marcel Bruchez (Molecular Biosensors & Imaging) is actively conducting experiments on intracellular signaling at ECI



Why an Academic Cloud Lab?

Research productivity

- Access to array of instrumentation, instead of just what is in individual labs

Increased productivity

- Estimates from ECL customers – up to 7x increase in research productivity
- Conservative estimate – 2x of research papers in same amount of time

Reproducibility

- Results are fully traceable

Open science

- Improved collaborations
- Knowledge sharing
- Open source

Educating next generation of scientists

Democratization of science

- Accessibility for anyone with internet

Next gen robotics

Active learning

- Automated instrumentation + machine learning

Pilot Project – Proof of Concept 2019

Dima recreated research that focused on creating novel compounds (synthetic DNA)

- synthetic DNA targeted generation of singlet oxygen using light activation of a genetically encoded fluorogen-activating proteins

In a traditional lab, synthesis of three compounds typically takes one week.

Using the Emerald Cloud Lab, he was able to synthesize hundreds of compounds in one week.

Kolodiezneyi now works at Emerald Cloud Lab.



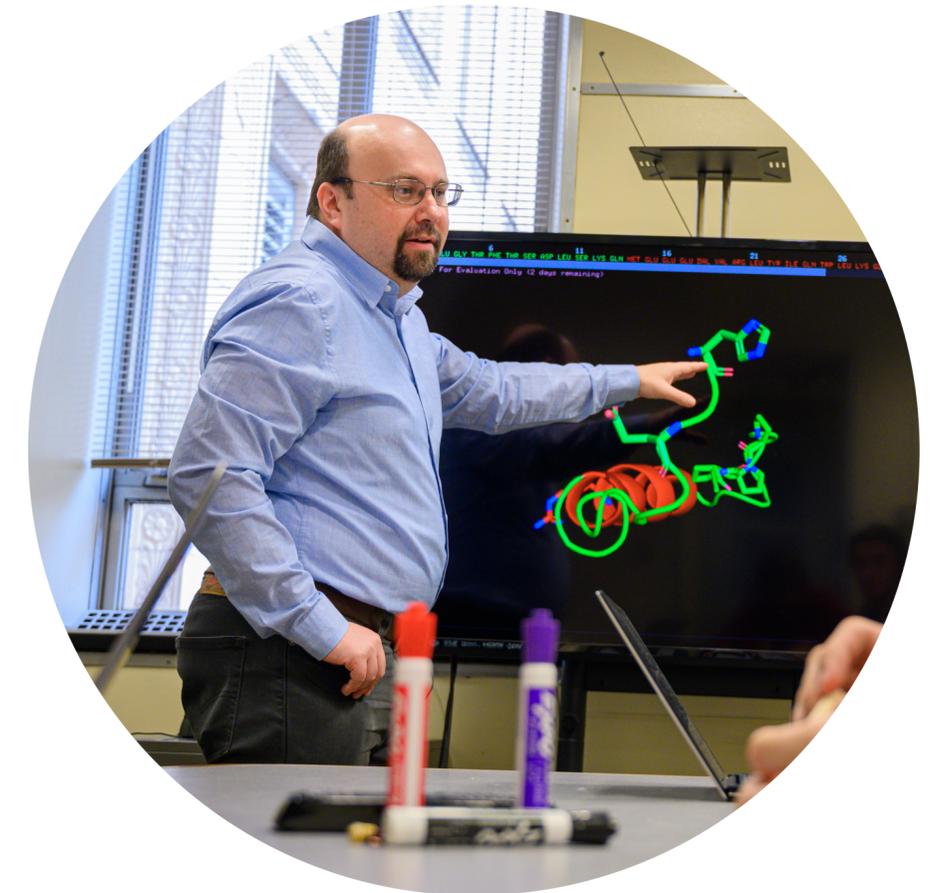
Dmytro
Kolodieznyi,
former MCS,
Chemistry Ph.D.
student

Proof of Concept: Automation & AI for Polymer Synthesis 2021

In a traditional lab, finding ideal candidates for an MRI imaging agent would require scientists to create and test 50,000 monomer compositions, which could take years — and likely would be impossible.

Using artificial intelligence, active learning, and automation Oles and colleagues found the ideal candidates by testing less than 400 polymers in a week of lab time.

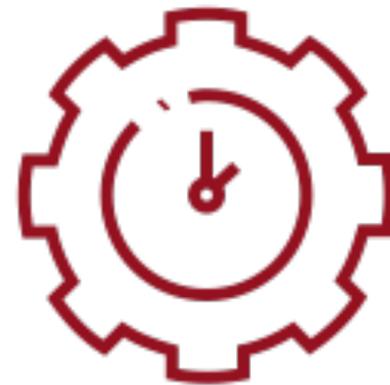
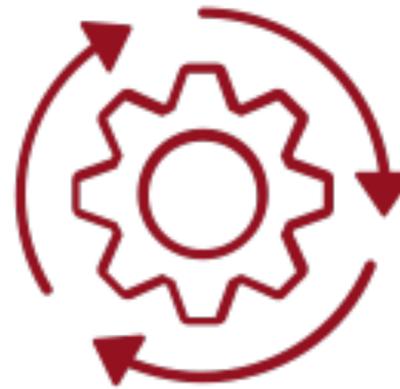
<https://pubs.acs.org/doi/10.1021/jacs.1c08181>



Olexandr(Oles) Isayev, MCS, Chemistry

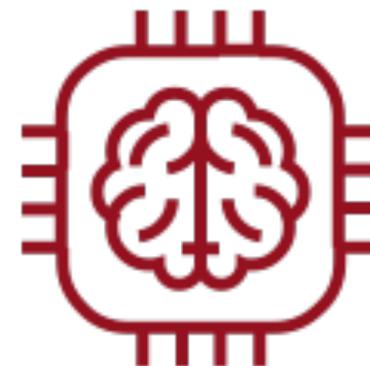
How Will the Cloud Lab Transform Science?

Reproducibility



Productivity

Open Science



Active Learning

How Will the Cloud Lab Transform Society?

Address Socially
Relevant
Challenges



Flexibility

Workforce
Development

Democratization
of Science

Timeline for CMU Academic Cloud Lab

Winter 2020	Financial planning	Complete
Spring 2021	University approval	Complete
Summer 2021	Hire architect and engineer	Complete
Fall – Winter 2021	Design + Renovations	In process
Fall – Winter 2021	Equipment negotiations and purchases	In process
Early 2022	Equipment receipt	
Summer 2022	Academic Cloud Lab build out and initial testing	
Fall 2022	Academic Cloud Lab fully operational	

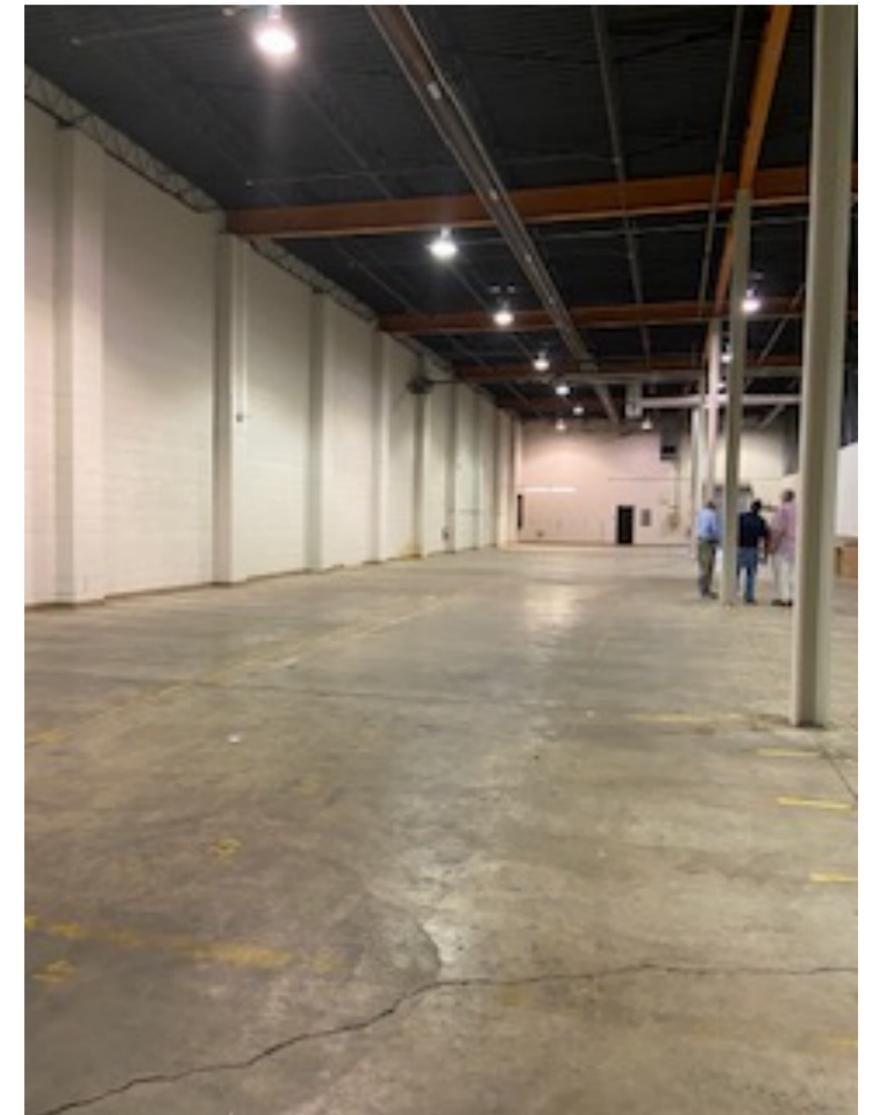
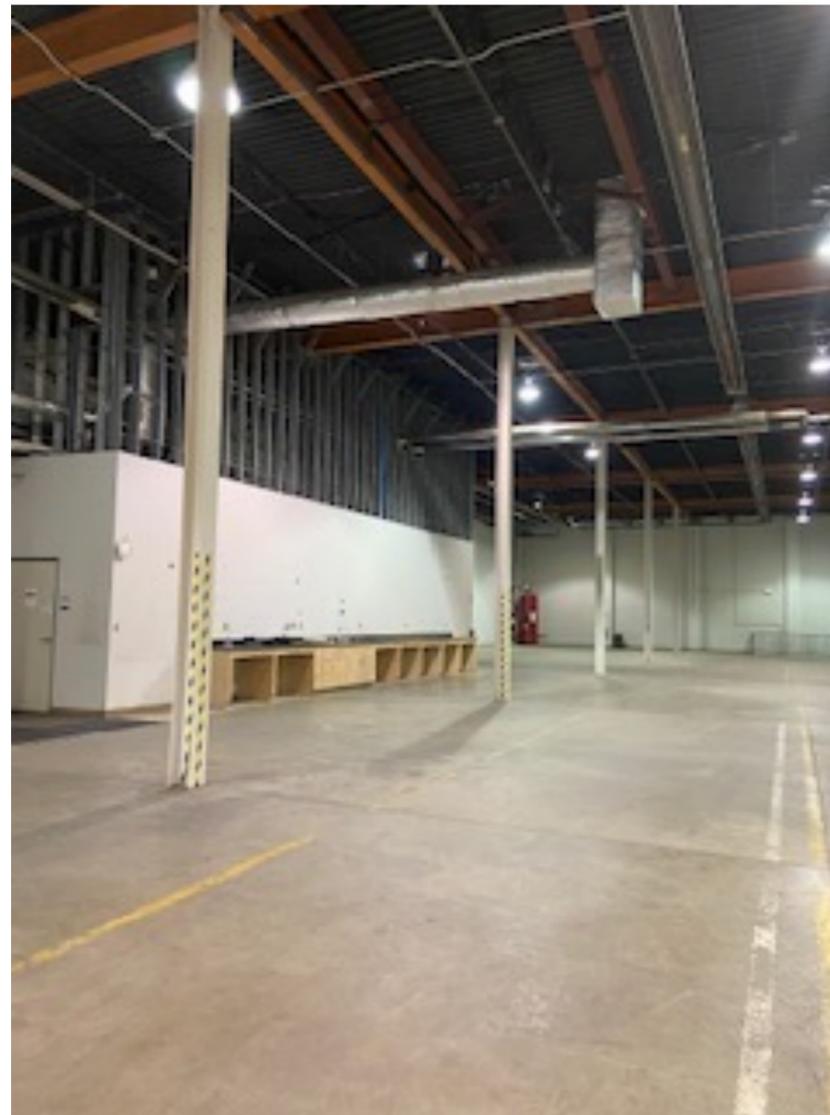
Future Home of the CMU Cloud Lab

- CMU owned building
- Near Bakery Square
- 10-min drive to campus



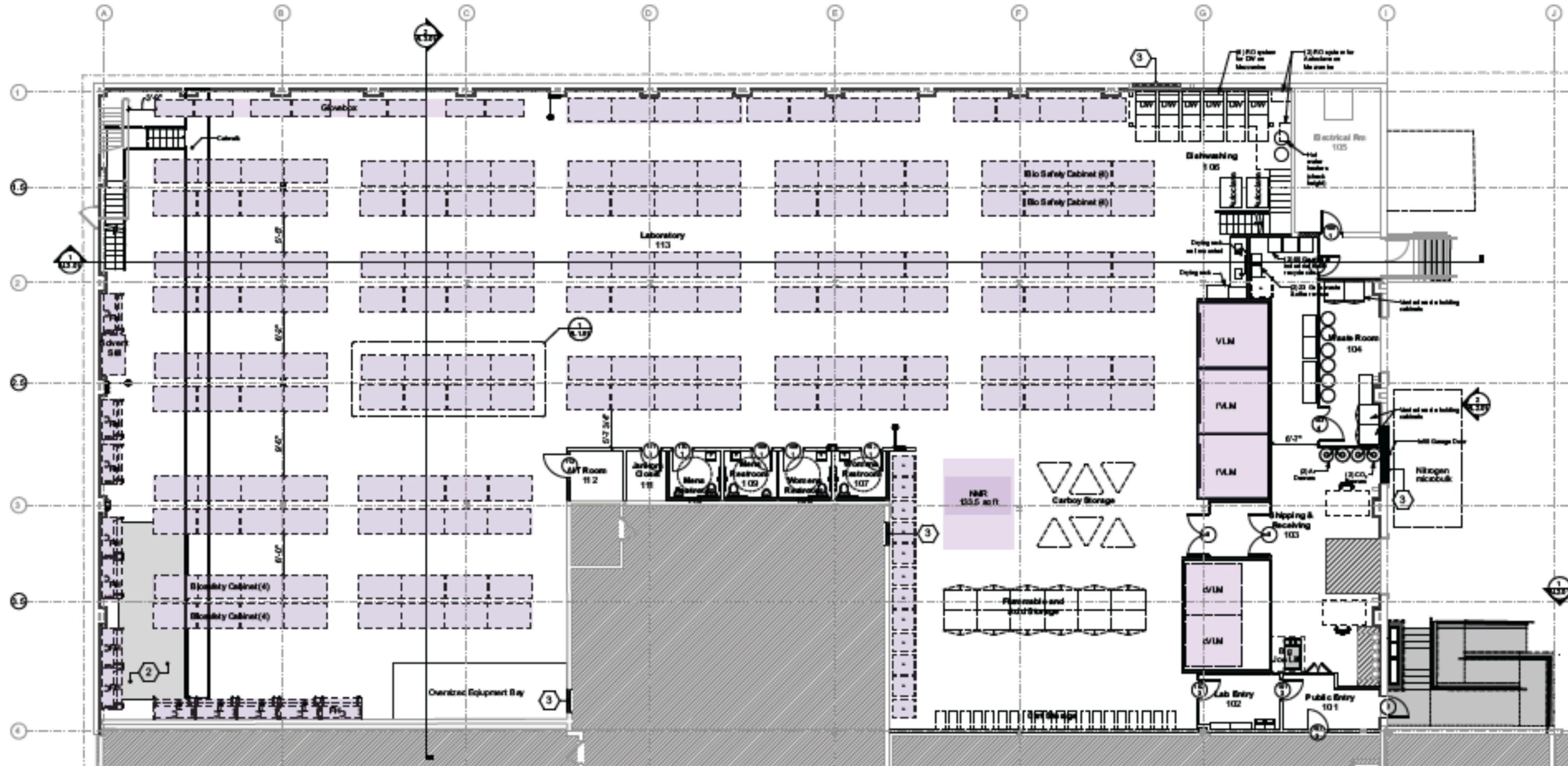
Construction Update – Nov 2021

- Architect and construction management teams hired
- Space vacated
- Core testing done (site is ok for heavy equipment)



Cloud Lab Layout

- ~ 16,000 sf
- ~200 unique instruments





What does future look like

- Open and collaborative
- Train next generation
- Expand platform capabilities – instrumentation, robotics and AI/ML
- Creation of new disciplines
- Streamline flow of technology and knowledge

The logo for Carnegie Mellon University, featuring a dark blue background with a grid of intersecting lines in red, green, and yellow. The text "Carnegie Mellon University" is written in a white, serif font, stacked in three lines.

**Carnegie
Mellon
University**

Thank you

14 DECEMBER 2021

Cloud labs, open science and the data deluge

Keith Webster

Dean of University Libraries

Helen and Henry Posner Jr. Dean's Chair

Director of Emerging and Integrative Media Initiatives

Carnegie Mellon University



@cmkeithw



CMU CLOUD LAB





1. Command

Ship your samples to an ECL facility and design your experiments in the ECL Command Center application.



2. Run

Emerald remotely conducts your experiments in a highly automated ECL facility exactly to your specifications.



3. Explore

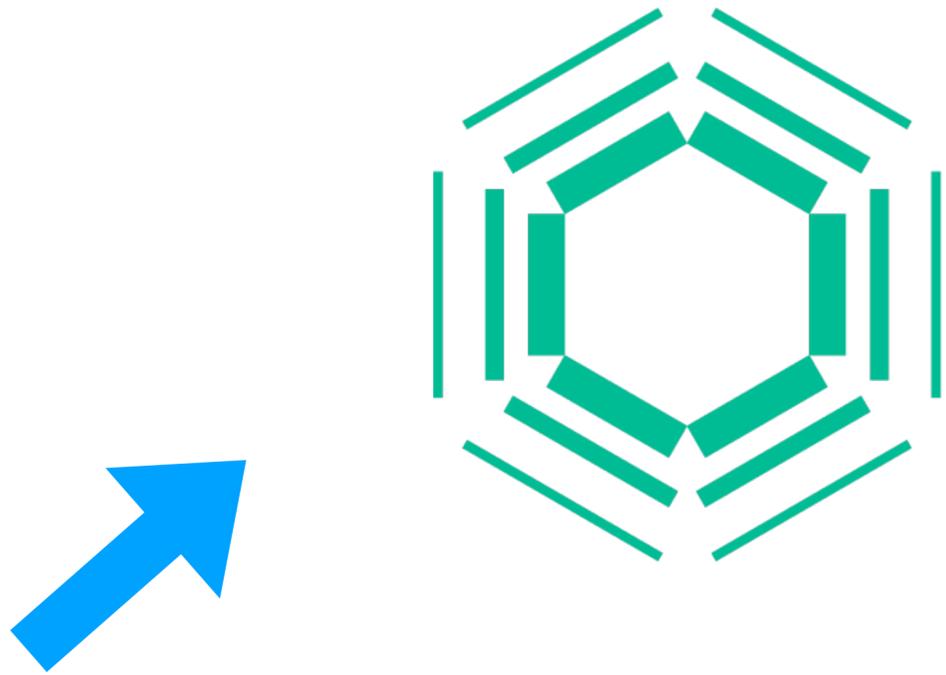
ECL Constellation organizes your data into a powerful knowledge graph, growing automatically over time as you conduct more experiments.



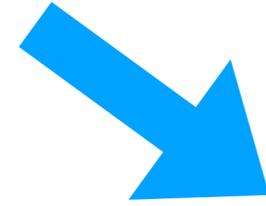
4. Analyze

ECL Command Center provides an extensive suite of tools to plot, analyze, and visualize your results.





ECCL





INDIVIDUAL EXPERIENCE

UNIVERSITY COMMUNITY

SOCIAL IMPACT

Search

CREATING A 21ST CENTURY LIBRARY

INDIVIDUAL EXPERIENCE

UNIVERSITY COMMUNITY

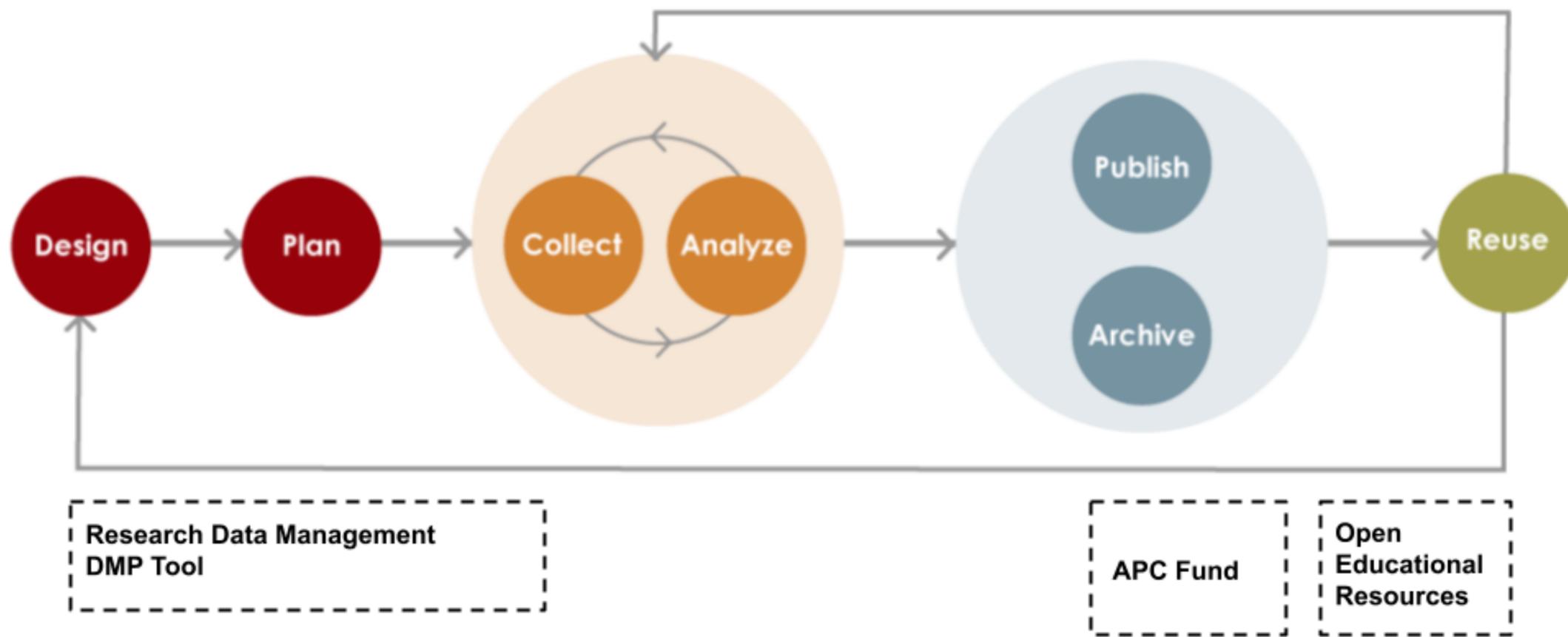
SOCIAL IMPACT

STRATEGIC RECOMMENDATION

Create a 21st century library that serves as a cornerstone of world-class research and scholarship.

In a world where digital is becoming the default format for information, the library will remain a vital presence on campus, sustaining serious scholarship and providing opportunities for interactive research and study environments. To support this important work for students, faculty, and staff, and to create 21st century library spaces for 21st century learners, the library will:

- Develop information specialists as partners in research, teaching, and learning.
- Collaborate with peer institutions to provide coordinated access to a global collection of information resources.
- Steward the evolving scholarly record, and champion new forms of scholarly communication.
- Be recognized globally as a leader in the development of the scholarly information ecosystem.



LabArchives

R
Python
OpenRefine
Citizen Science

protocols.io
Open Science Framework

KiltHub

Open Science Symposium
AIDR
Collaborative Bioinformatics Hackathon
Data CoLab

Final NIH Policy for Data Management and Sharing

Notice Number:

NOT-OD-21-013

Key Dates

Release Date:

October 29, 2020

Effective Date:

January 25, 2023

Related Announcements

[NOT-OD-21-014](#) – Supplemental Information to the NIH Policy for Data Management and Sharing: Elements of an NIH Data Management and Sharing Plan

[NOT-OD-21-015](#) – Supplemental Information to the NIH Policy for Data Management and Sharing: Allowable Costs for Data Management and Sharing

[NOT-OD-21-016](#) – Supplemental Information to the NIH Policy for Data Management and Sharing: Selecting a Repository for Data Resulting from NIH-Supported Research

[NOT-OD-20-013](#) - Request for Public Comments on a DRAFT NIH Policy for Data Management and Sharing and Supplemental DRAFT Guidance

Issued by

Office of The Director, National Institutes of Health ([OD](#))

Purpose

Summary

The National Institutes of Health (NIH) is issuing this final NIH Policy for Data Management and Sharing (DMS Policy) to promote the management and sharing of scientific data generated from NIH-funded or conducted research. This Policy establishes the requirements of submission of Data Management and Sharing Plans (hereinafter Plans) and compliance with NIH Institute, Center, or Office (ICO)-approved Plans. It also emphasizes the importance of good data management practices and establishes the expectation for maximizing the appropriate sharing of scientific data generated from NIH-funded or conducted research, with justified limitations or exceptions. This Policy applies to research funded or conducted by NIH that results in the generation of scientific data.



OPEN SCIENCE

PROCEEDINGS OF A WORKSHOP

Developing a Toolkit for Fostering OPEN SCIENCE PRACTICES



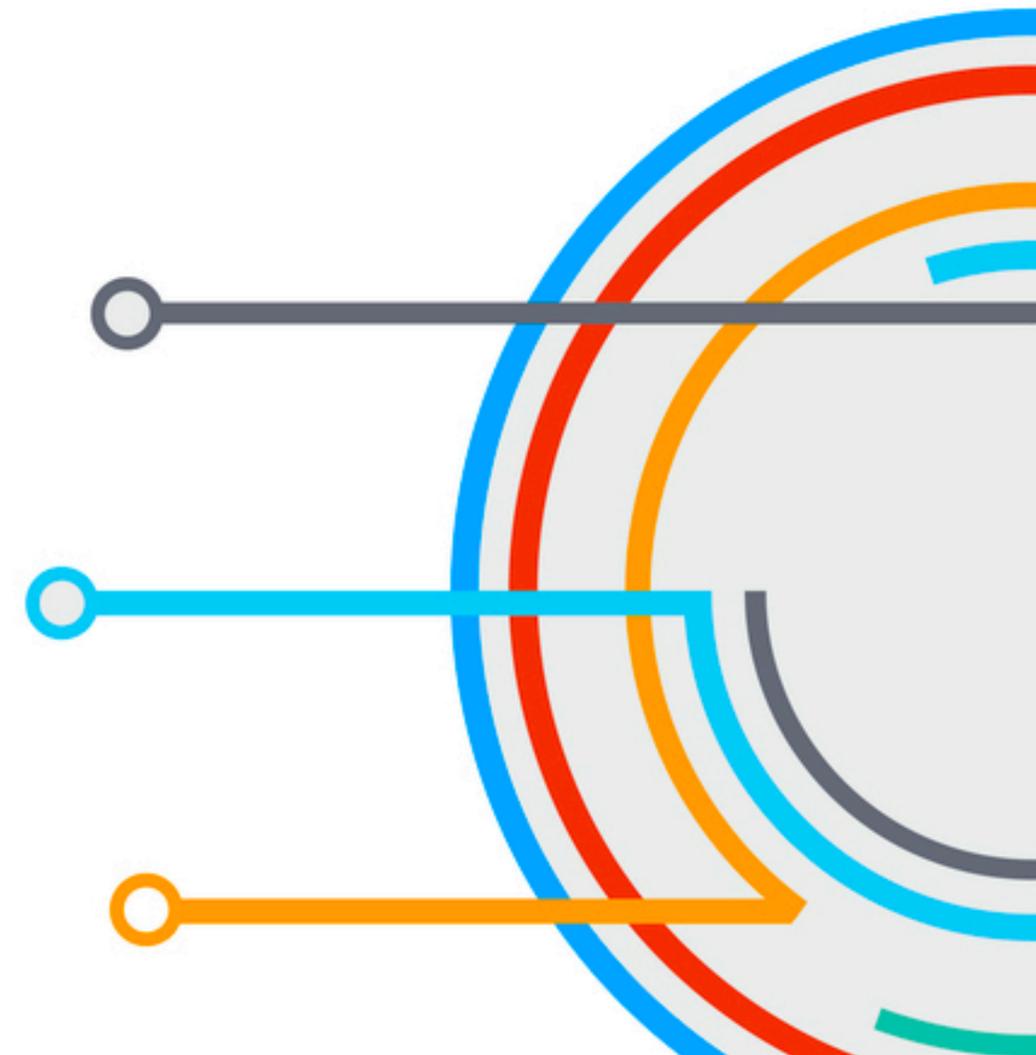
The National Academies of
SCIENCES • ENGINEERING • MEDICINE

The National Academies of
SCIENCES • ENGINEERING • MEDICINE

CONSENSUS STUDY REPORT

OPEN SCIENCE BY DESIGN

Realizing a Vision for 21st Century Research



The COVID-19 pandemic raises our awareness of the importance of science, both in research and international cooperation. The present crisis also demonstrates the urgency of stepping up information sharing through open science. The time has come for us to commit all together.

~ UNESCO

Open Science aims to make scientific knowledge, methods, data and evidence freely available and accessible for everyone, increase scientific collaborations and sharing of information for the benefits of science and society, and open the process of scientific knowledge creation and circulation to societal actors beyond the institutionalized scientific community



UNESCO Recommendation on Open Science

<https://en.unesco.org/science-sustainable-future/open-science/recommendation>

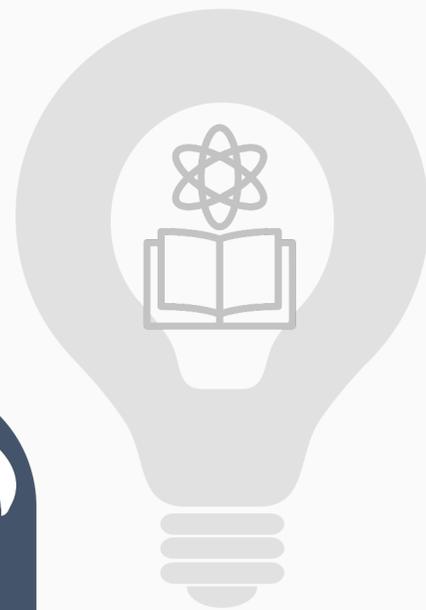


**CMU LIBRARIES
& OPEN SCIENCE**

CMU Libraries and Open Science

Our role is grouped into five themes

Tools
KiltHub
LabArchives
OSF
protocols.io



Carnegie Mellon University

A Project Management Tool for the CMU Community: [Get Help at CMU](#) | [About OSF](#) | [OSF Support](#) | [Terms of Use](#)

All Projects >

Filter displayed projects ✕

Collections

All Projects

All Registrations

Contributors

< 1/17 >

Hannah C. Gunderman

Melanie Gainey

Huajin Wang

Katie Behrman

Tags

< 1/17 >

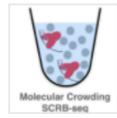
Carnegie Mellon University

Active Learning

Database Instruction

Library Instruction

Name ^ v	Contributors	Modified ^ v Fetching more projects
 Speed Databasing	Chisnell, MacGregor	4 days ago
 TOFU: An artificial Mandarin-like tonal language	Wiener	9 days ago
 Lifestyle and Mental Health Disruptions During Covid-19	Saccardo, Sadoff + 2	4 months ago
 Motor Vehicle Technology Forecasting Survey	Savage	14 days ago
 Empathy, Theory of Mind, Emotion Recognition, and Alcohol; Meta-Analysis and Syste...	Kumar, Creswell	21 days ago
 SAT-M	Carvalho, Goldstone	24 days ago
 The effect of pluralization on valence biases in judgments of group membership	Sloman	a month ago
 2021 Bringing Genomics Data to the Clinic Hackathon	Al Khleifat, Smith + 22	a month ago
 Effects of perceptual abilities and lexical knowledge on the phonetic categorization of s...	Wiener, Liu	a month ago
 Reflections on "Best Practices" from a Neurodivergent and Anxious RDM Librarian	Gunderman	2 months ago
 TrackIt Task	Fisher	2 months ago
 Scoping Review of Pre-clinical and Translational Studies on Macrophage Polarization in...	Colette Bilynsky, Han + 11	2 months ago
 Dynamic Decision Making Laboratory	Gonzalez, Flagg	2 months ago
 Five-Item Guilt Proneness Scale (GP-5)	Cohen, Kim + 1	2 years ago



mcSCRB-seq protocol V.2

Nature Communications

Johannes JWB Bagnoli¹, Christoph Ziegenhain¹, Aleksandar Janjic¹, Lucas Esteban LEW Wange¹, Beate Vieth¹, Swati Parekh¹, Johanna Geuder¹, Ines Hellmann¹, Wolfgang Enard¹

¹Ludwig-Maximilians-Universität München

Version 2

May 22, 2018

11 Works for me

Share

dx.doi.org/10.17504/protocols.io.p9kdr4w

Bookmark

Human Cell Atlas Method Development Community

Run

Copy / Fork

Aleksandar Janjic
Ludwig-Maximilians-Universität München

View all 43 comments

Search in comments

Hira Mani Jul 22, 2020

To digest lamda DNA with different restriction endonuclease and analyse the difference an restriction digestion pattern of lamda DNA ?

REPLY

Steps Guidelines Materials Forks Metadata Metrics

ABSTRACT

Single-cell RNA sequencing (scRNA-seq) has emerged as a central genome-wide method to characterize cellular identities and processes. Consequently, improving its sensitivity, flexibility and cost-efficiency can advance many research questions. Among the flexible plate-based methods, "Single-Cell RNA-Barcoding and Sequencing" (SCR-seq) is one of the most sensitive and efficient ones. Here, we systematically evaluated experimental conditions of this protocol and find that adding polyethylene glycol considerably increases sensitivity by enhancing cDNA synthesis. Furthermore, using Terra polymerase increases efficiency due to a more even cDNA amplification that requires less sequencing of libraries. We combined these and other improvements to a new scRNA-seq library protocol we call "molecular crowding SCR-seq" (mcSCRB-seq), which we show to be the most sensitive and one of the most efficient and flexible scRNA-seq methods to date.

ATTACHMENTS



mcSCRBseq_oligodT.txt

BEFORE STARTING

Wipe bench surfaces with RNase Away and keep working environment clean.



labarchives Free Edition

Search Notebook

Stephanie Lehman

Notebook Navigator

- Stephanie's Lab Notebook at UMB
 - Protocols
 - Lipofectamine
 - SOPs
 - MSDS
 - Templates
 - Experimental Data
 - Ideas
 - Lab Meeting Notes
 - Paper Outlines
 - Manuscripts
 - Published data
 - Grants
 - Presentation
 - References
 - Project 1
 - New Page
 - New Folder
 - Transfect FLAG-LacZ into HeLa

Transfect FLAG-LacZ into HeLa

Add Entry Attachment Office document Heading More

Page Tools Revisions Sign DOI

- Mix lipofectamine gently before use, dilute appropriate amount (0.5ul) of lipofectamine into 25ul of OptiMEM1 per reaction. Mix gently, incubate 5 min, RT. Add lipofectamine to DNA within 30 min.
- After 5m in incubation, combine diluted DNA with diluted lipofectamine (total volume is 50ul). Mix gently, incubate 20 min, RT.

(**values in red are appropriate for a T75 flask, as per kit instructions.) (kb uses 10ug for T75 if the protein expresses well)

- Dilute sample and control DNA (10ug) in 1.5 mL of Opti-MEM1 Reduced serum medium without serum. Mix gently.
- Mix lipofectamine gently before use, dilute appropriate amount (60ul) of lipofectamine into 1.5 mL of OptiMEM1 per reaction. Mix gently, incubate 5 min, RT. Add lipofectamine to DNA within 30 min.
- After 5m in incubation, combine diluted DNA with diluted lipofectamine (total volume is 3 mL). Mix gently, incubate 20 min, RT.

Transfection

- Add 100ul of Opti-Mem to the 50ul DNA-lipofectamine complexes. Remove spent media on cells in SLIDE Add all 150ul to each well. Mix gently by rocking plate back and forth.
- Add 15 mL of Opti-Mem to the 3 mL DNA-lipofectamine complexes. Remove spent media on cells in FLASK Add all 18 mL to the flask. Mix gently by rocking plate back and forth.
- Incubate cells at 37C, 5%CO2, 24-48hrs. Observe via microscopy.

Stephanie Lehman - Jan 26, 2016 @10:17 PM ES

©2016 LabArchives, LLC. All rights reserved.

KiltHub

Discover research from Carnegie Mellon University

+ Follow

20,309 posts | 6,243,151 views | 9,863,302 downloads | more stats...

ALL CATEGORIES GROUPS SEARCH Q

sort by: Posted date



How What Goes on in Others' Minds Affects our Choices and We...
Thesis posted on 07.05.2021
Andras Molnar



Delta Oscillations: Prevalence, Propagation, and...
Thesis posted on 07.05.2021
Timothy Whalen



Higher Inductive Types and Internal Parametricity for...
Thesis posted on 07.05.2021
Evan Cavallo



Future Tenant
a space for art
In the Company of Men
Media posted on 07.05.2021
Future Tenant Gallery

CMU open access agreements



ELSEVIER

First institutional agreement for Elsevier - for all CMU corresponding authors

1 JAN 2020



ACM

Agreement via ACM Open - aims to move all journals to OA

1 JAN 2020



PLOS

Our largest open access only publisher

1 JULY 2020



CAMBRIDGE

Our first agreement with a university press

1 JAN 2021



WILEY

Agreement with a commercial publisher with substantial society portfolio

1 JAN 2022

CMU Libraries and Open Science

Our role is grouped into five themes

Training
Carpentries
Other workshops



Tools
KiltHub
LabArchives
OSF
protocols.io





**We teach foundational coding
and data science skills to
researchers worldwide.**



What we do

The Carpentries teaches foundational coding and data science skills to researchers worldwide. Software Carpentry, Data Carpentry, and Library Carpentry workshops are based on our lessons. Workshop hosts, Instructors, and learners must be prepared to follow our [Code of Conduct](#).

[More >](#)



Who we are

Our diverse, global community includes [Instructors](#), helpers, [Trainers](#), [Maintainers](#), [Mentors](#), community champions, [member organisations](#), supporters, workshop organisers, [staff](#) and a whole lot [more](#).

[More >](#)



Get involved

See all the [ways you can engage](#) with The Carpentries. Get information about upcoming events such as workshops, meetups, and discussions from our [community calendar](#), or from our twice-monthly [newsletter](#), *Carpentry Clippings*. Follow us on [Twitter](#), [Facebook](#), and [Slack](#).

[More >](#)

May 2021						
Su	Mo	Tu	We	Th	Fr	Sa
25	26	27	28	29	30	1
2	3	4	5	6	7	8
9	10	11	12	13	14	15
16	17	18	19	20	21	22
23	24	25	26	27	28	29
30	31	1	2	3	4	5

Search for event...

Calendar*
Workshops

Filter by
 Category
 Library
 Onsite And Online

Workshops - Upcoming Events

Time Zone: Eastern Time - US & Canada ([change](#))



MAY 25 Tue, 1:00pm - 4:00pm Online
Online Workshop

1 SEAT LEFT

Introduction to R

This event is a virtual example-driven workshop on basic skills for working with data in R. Short tutorials alternate with hands-on practical exercises, and participants are encouraged both to help on... [More](#)

Open Science Working With Data

JUN 7 Mon, 12:00pm - 5:00pm Online
Online Workshop

Intro to Python, Shell, and Git

This event is a virtual example-driven workshop on basic computing skills that runs over 3 half days (15 hours total). Short tutorials alternate with hands-on practical exercises, and participants are... [More](#)

Open Science Working With Data

JUN 8 Tue, 12:00pm - 5:00pm Online
Online Workshop

Intro to Python, Shell, and Git

This event is a virtual example-driven workshop on basic computing skills that runs over 3 half days (15 hours total). Short tutorials alternate with hands-on practical exercises, and participants are... [More](#)

Open Science Working With Data

JUN 9 Wed, 12:00pm - 5:00pm Online
Online Workshop

Intro to Python, Shell, and Git

This event is a virtual example-driven workshop on basic computing skills that runs over 3 half days (15 hours total). Short tutorials alternate with hands-on practical exercises, and participants are... [More](#)

Open Science Working With Data

JUN 24 Thu, 1:00pm - 3:00pm Online
Online Workshop

9 SEATS LEFT

Data Visualization with R

This event is a virtual example-driven workshop on basic skills for creating data visualizations in R. Short tutorials alternate with hands-on practical exercises, and participants are encouraged both... [More](#)

Open Science Working With Data

OSDC MiniSeries: Reproducible Research

All the goodness of our popular Carpentries workshops packed into shorter, bite-size portions!

These courses were created from curriculum originally developed by [The Carpentries](#).

The aim of the OSDC MiniSeries workshops is to teach researchers basic concepts, skills, and tools for conducting research in a reproducible manner. [Reproducible research](#) is the idea that any result obtained by an experiment or observational study should be documented by making all data and code available in such a way that the computations can be executed again with identical results. Click [here](#) to learn more about research reproducibility at CMU.

Mini courses are designed to be completed within 2 to 4 hours. This abbreviated format is intended to complement the more intensive 2-3 day format of our [Carpentries workshops](#) and the 1-hour [Library workshop series](#). Minis provide more modular presentations of the Carpentries curriculum and are therefore ideal for participants with more limited availability.

OSDC workshops are designed for participants with no programming experience.

Courses in this MiniSeries

Below is a list of the courses contained in this MiniSeries. Check out the [CMU Libraries' Events Calendar](#) to learn about upcoming workshops and to access the registration page, or follow the links provided under each course description.

Cleaning Untidy Data with OpenRefine

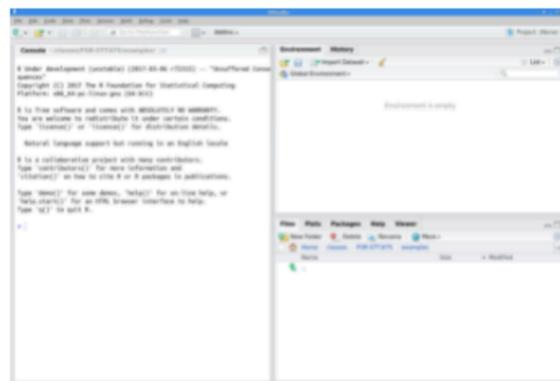


OpenRefine is a free, open source tool to help you prepare your data for analysis. Quickly and easily transform data, split and merge columns, remove whitespace, and perform many more common data cleaning tasks.

Duration: 2 hours.

[View course curriculum](#). Date of next workshop: Coming soon!

Introduction to R



This 3-hour course was designed for those interested in working with data in R. It begins with basic information about R syntax and the RStudio interface, and moves through how to import CSV files, the structure of data frames, how to deal with factors, how to add/remove rows and columns, and how to calculate summary statistics from a data frame.

Duration: 3 hours.

Date of next workshop: 05/25/2021

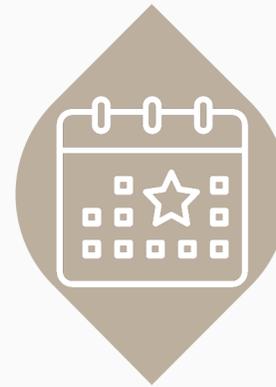
[View course curriculum](#). Register [here](#).

CMU Libraries and Open Science

Our role is grouped into five themes

Events

AIDR & OSS



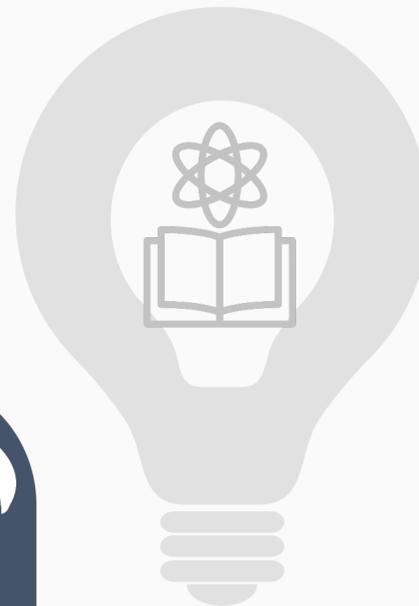
Training

Carpentries
Other workshops



Tools

KiltHub
LabArchives
OSF
protocols.io



The logo for AIDR 2020 features the word "AIDR" in a large, bold, white, sans-serif font. To the right of "AIDR", the year "2020" is displayed in a smaller, white, sans-serif font, with the "20" stacked above the "20". A horizontal white line with a pixelated end is positioned below the "2020". The entire logo is set against a red background that is part of a larger graphic design.

October 19, 2020

<https://events.library.cmu.edu/aidr2020/>



October 20, 2020

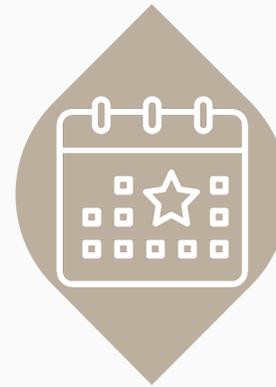
<https://events.library.cmu.edu/oss2020/>

CMU Libraries and Open Science

Our role is grouped into five themes

Events

AIDR & OSS



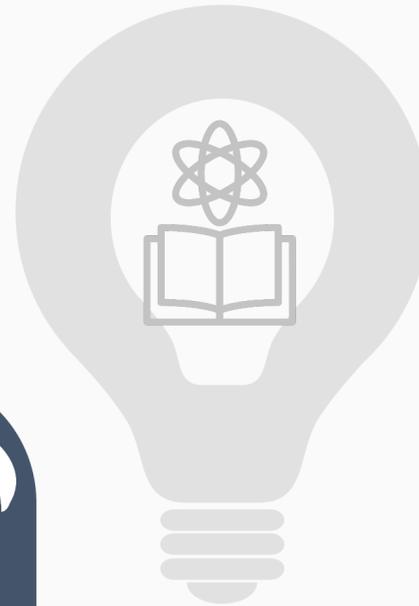
Training
Carpentries
Other workshops



Collaboration
dataCoLAB
Bioinformatics Hackathon
Citizen Science



Tools
KiltHub
LabArchives
OSF
protocols.io



Data Collaborations Lab

Helping data producers and data scientists connect and collaborate.

The Data Collaborations Lab (dataCoLAB) at Carnegie Mellon University Libraries connects the research community across disciplinary borders, and facilitates collaborations between data producers and data scientists. The program connects researchers who want more from their datasets with individuals who have data and computer science skills, creating opportunities for people with different technical and disciplinary backgrounds to work together. The ultimate goal of the dataCoLAB program is to help build a strong community and a healthy data ecosystem.

Anyone from the CMU or Pittsburgh community can participate. Perhaps you have an existing dataset and want help analyzing, organizing, or visualizing it. Or maybe you have data science skills and want to gain experience consulting on interesting real-world data problems. [Learn more about whether dataCoLAB is for you!](#)

Information professionals from the CMU University Libraries will help participants connect with collaborators and get started. Participants get support on research data management, project documentation, and other research methodologies. Your library support will help you find and use tools for collaborating and documenting your project, and provide guidance on best practices for making your project publically available and citable.

src > components > sidebar > sidebarComponent.svelte > ...

You, 2 months ago | 1 author (You)

```
1 <script>
2   import sidebarController from './sidebarController.js'
3   import mementoes from '../store.js'
4   import './sidebar.scss'
5 </script>
6
7 <div class="sidebar">
8   <button class="add-memento" on:click=sidebarController.addMemento>
9
10  <ul class="mementoes">
11    {#each $mementoes as memento}
12      <li class="memento-item" class:active=memento.isActive
13        on:click={() => sidebarController.selectMemento(memento)}
14        {memento.title}
15      </li>
16    {/each}
17  </ul>
18 </div>
```

Coming soon: Coding Collab

CMU Libraries and Open Science

Our role is grouped into five themes



Developing a Logic Model

INPUTS

Project team
Library leadership
Other library staff
Volunteers
Library space
Digital tools and databases

Time
Effort
Money

ACTIVITIES

KiltHub
Open Science Framework
Protocols.io
LabArchives
Emerald Cloud Lab integration

Carpentries workshops
Reproducibility MiniSeries
Other workshops

AIDR
Open Science Symposium

Data Collaborations Lab
Hackathons
Citizen Science

Newsletter
Website
Advisory board

OUTPUTS

437 KiltHub deposits
370 research projects (non-KiltHub)
> 450 users total
~10 training sessions per year

~20 workshops per year
> 270 participants

2 day-long events per year
> 230 participants

Weekly office hours
1 hackathon per year
26 collaborative projects to date

> 500 newsletter subscribers
8 external board members

OUTCOMES

Short Term

Learn about open workflow
Learn about open science tools
Begin to learn how to code
Learn about reproducibility

Medium Term

Use open workflow in daily work
Integrate open science tools
Practice coding
Adopt reproducible practices

Long Term

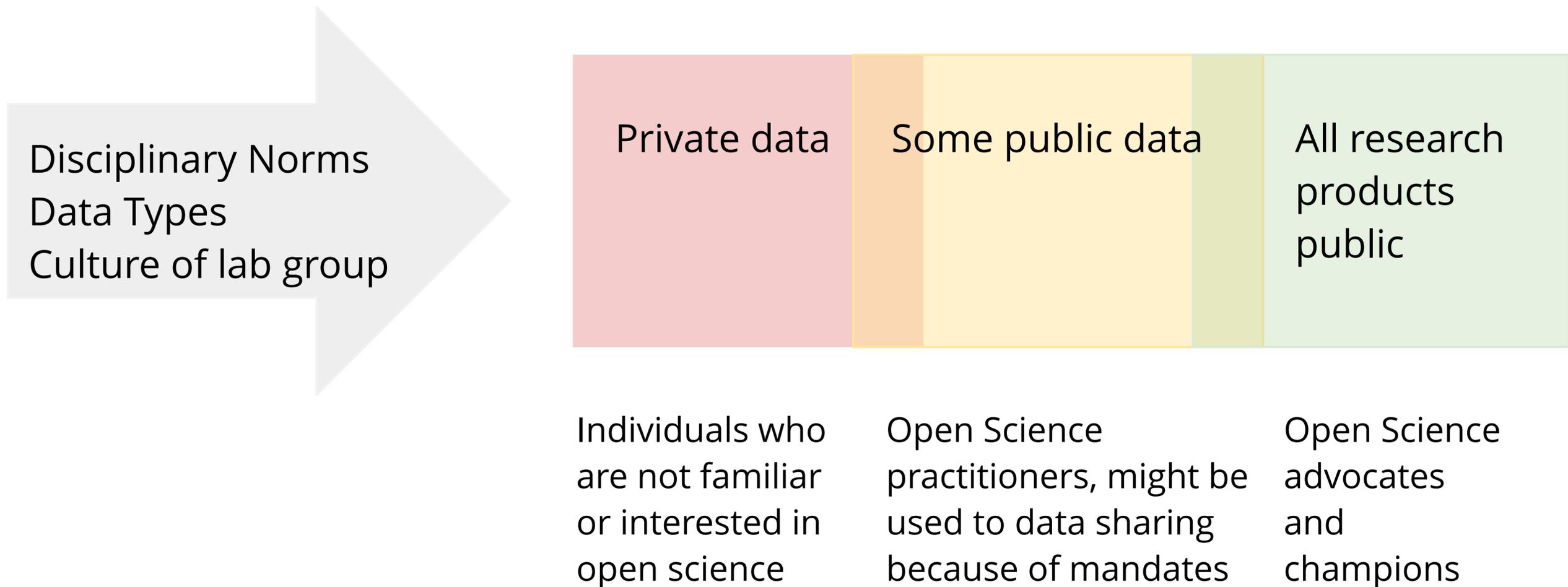
Cultural shift towards open and reproducible work across disciplines

PARTNERS

University and college leadership
Mellon College of Science
Digital platform providers
Open science advocates

Program Outreach

Presenting open science as a gradient of practices



Program Outreach

Example of outreach for protocols.io



Private Protocols

- Improved documentation
- Version Control
- Reproducibility for your later self

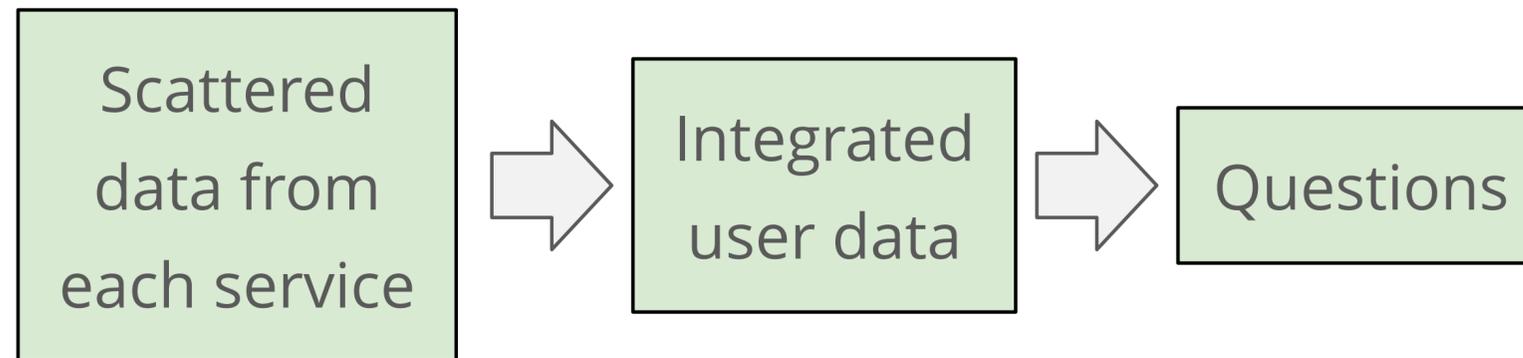
Protocols shared with research group

- Reproducibility over time within a lab
- Publishing complete methods

Public protocols

- Transparency
- Discoverability
- Importance for fixing the reproducibility crisis

Developing meaningful metrics



Who uses our tools and participates our activities?

Who are our top users?

Which **disciplines** are the most engaged?

How do people use our tools or activities?

Why do people use our tools or activities?

What **impact** are we making?

Developing meaningful metrics

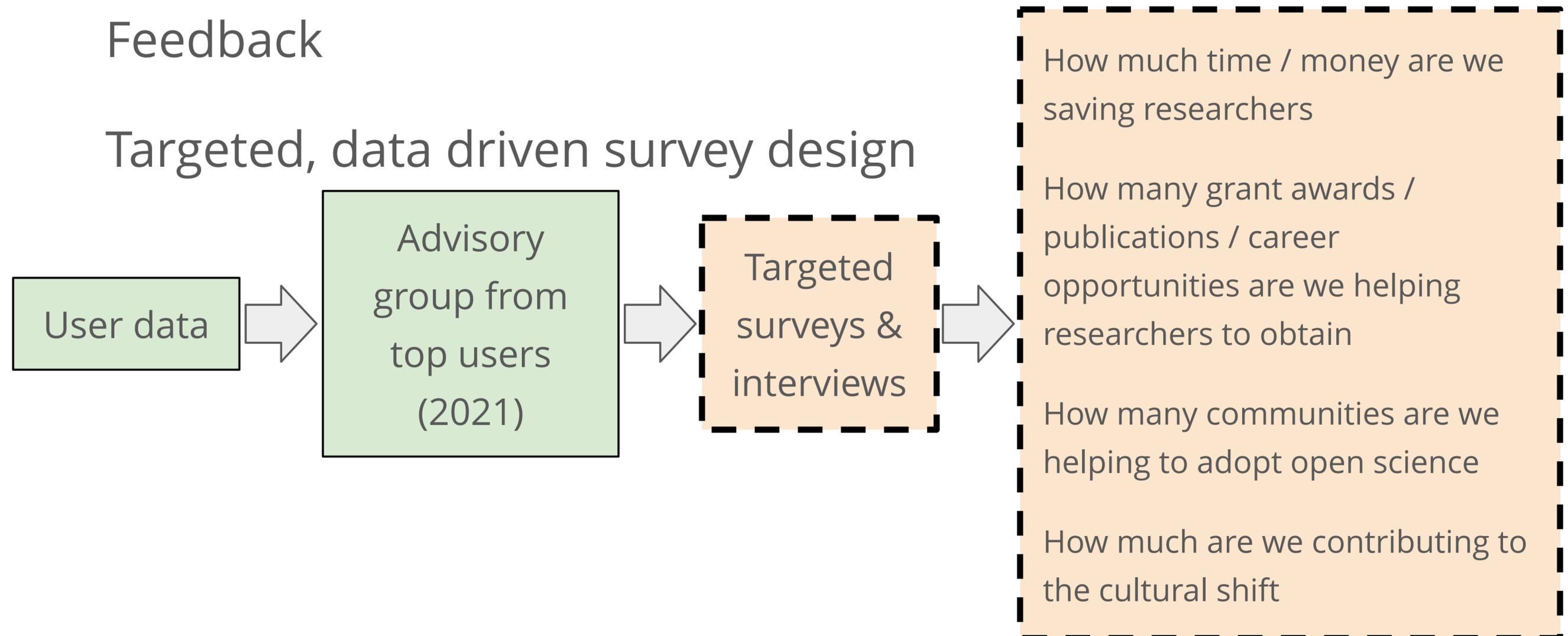
	Metric	Variable(s)	Source of Data
Who	User affiliation	Institution, Department	Dashboard
	Stage of career	User type (faculty, postdoc, etc.)	Derived
	Superusers	Counts, Number of projects and registrations (all tools/events)	Derived
What	Number of users per tool	User (T/F) - all tools/events	Dashboard, Vendor
	Number of tools/events used per user	User (T/F) - all tools/events	Derived
	Number of registrations per event	Count (all events/workshops)	Dashboard
	Number of attendances per event	Count (all events/workshops)	Dashboard
	Number of event/workshop registrations per user	Counts (all events/workshops)	Derived
	Departmental breakdown of users per tool/event	User (T/F), Institution, Department	Derived
	Career stage breakdown of users per tool/event	User (T/F), Career Stage	Derived

When	Growth rate (growth over time)	Number of users plus time/date field	Derived
Why	User satisfaction (qualitative and quantitative)	User comments / feedback	Advisory Board, Surveys
	Financial metrics (for users)	Cost savings	Vendors
How	Output (number of products, tasks completed, etc.)	Number of projects and registrations (OSF), Number of notebooks (LabArchives), Number of activities (LabArchives), Number of protocols (protocols.io), Count of events of each type attended (workshops, Carpentries, DataCoLAB, AIDR_OSS), Count_KiltHub (KiltHub)	Dashboard, Vendors
	Activity over time	Output plus date/time fields	Derived
	Reach	Open rate, Click rate (newsletter)	Dashboard

Why: what do people get out of our services

Feedback

Targeted, data driven survey design



OPEN SCIENCE NEWSLETTER

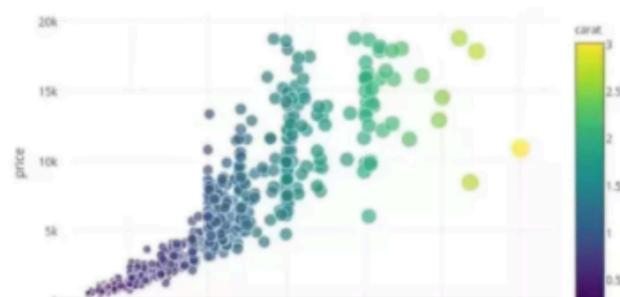
With cooler weather on the way, it's the perfect time to curl up on the couch with your favorite hot beverage (pumpkin spice latte, perhaps?) and enjoy one of this month's many virtual open science workshops and events, or take advantage of our new asynchronous courses in R, fresh off the press and ready whenever you are on the CMU Libraries YouTube channel (link below)! Round things off with a quick browse through this month's recommendations in open science news and then reserve your spot for a live demo of the "Just One Giant Lab" platform for open science and innovation. Hope to see you there!

Please contact us at openscience@andrew.cmu.edu if you have any questions and follow us on social media at #CMUOpenScience.

[Subscribe to our newsletter.](#)

Our Reproducible Research MiniSeries continues with *DATA VISUALIZATION WITH R!*

Last month we brought back our [Reproducible Research MiniSeries](#) with an encore presentation of our Introduction to R workshop. OSDC Minis are designed to teach



MELANIE GAINEY



HUAJIN WANG

Subscribe to our newsletter

Join our events

Look out for next AIDR and OSS conferences

Gainey & Wang presentation to CNI Fall 2021 (slides 48-53)

 openscience@andrew.cmu.edu

 [#CMUOpenScience](https://twitter.com/CMUOpenScience)

<https://www.library.cmu.edu/services/open-science>