

CaRCC Researcher-Facing Call, 2022-06-09

Supporting researchers with the NIH Data Management and Sharing Plan Requirement

Speakers

- Lisa Federer (Data Science and Open Science Librarian, National Library of Medicine): Introduction to NIH Data Management and Sharing Policy
- Kat Milligan-McClellan (Assistant Professor of Molecular and Cell Biology / Microbiology, UConn): Researcher challenges
- Tobin Magle (Lead Data Management Specialist / Northwestern University): How can researcher-facing staff provide support?

Please Note:

- **We will record this call and post shortly thereafter on CaRCC's YouTube channel.**
- **We expect all persons on the call to adhere to [CaRCC's Code of Conduct](#).**

[Connection Details](#)

[Agenda](#)

[Announcements](#)

[Sign-In \(Name / Affiliation /Email\)](#)

[Notes from the call](#)

[Slides](#)

[Notes](#)

[Questions](#)

[Chat Comments](#)

Connection Details

<https://utah.zoom.us/my/carcc?pwd=TjFuR3VVM2d5eE5zWnEvWWxDTFBCUT09>

Meeting ID: 824 051 8198

Password: 31415926

One tap mobile

+13462487799,,8240518198#,,#31415926# US (Houston)

+16699006833,,8240518198#,,#31415926# US (San Jose)

Dial by [your location](#)

+1 346 248 7799 US (Houston)

+1 669 900 6833 US (San Jose)

Join by Skype for Business: <https://utah.zoom.us/j/8240518198>

Agenda

Time	Item
12:55- 1:05 pm EST	Welcome + Introduction to CaRCC • Researcher-Facing Track description • CaRCC is your community for research computing and data professionals. Please see this brief overview (our OneSlide); more information on activities on our Groups web pages. • If you have questions about CaRCC or are interested in becoming more involved, please contact: ◦ rf-coordinators@carcc.org for R-F-related activities, or ◦ getstarted@carcc.org or getinvolved@carcc.org for other CaRCC work • We expect all persons on the call to adhere to CaRCC's Code of Conduct.
1:05 - 1:35 pm EST	Topic: Supporting Researchers with the NIH Data management and sharing plan requirement Panelists: • Lisa Federer (Data Science and Open Science Librarian, National Library of Medicine): Introduction to NIH Data Management and Sharing Policy • Kat Milligan-McClellan (Assistant Professor of Molecular and Cell Biology / Microbiology, UConn): Researcher challenges • Tobin Magle: How can researcher-facing staff provide support?
1:35 - 1:55 pm EST	Q&A Brainstorming Jam Board: https://jamboard.google.com/d/1OznFZiyGv0fqimklxiWfA8xNZI4ghFdZr4z5Dr3ykB0/viewer?f=0
1:55 - 2:00 pm EST	Session close

Announcements

Sign-In (Name / Affiliation /Email)

Note: please follow the suggested sign-in format so our evolving data science intelligence routines won't trip up and forget to enter you in our \$1m sweepstakes.

1. Jason Wells / Harvard University SEAS / jrwells@seas.harvard.edu
2. Lisa Federer / National Library of Medicine / lisa.federer@nih.gov
3. Kat Milligan-McClellan / UConn / kat.milligan-mcclellan@uconn.edu
4. Justin Booth / Michigan State / boothj@msu.edu
5. Alicia Hofelich Mohr / University of Minnesota / hofelich@umn.edu
6. Marla Hertz / University of Alabama at Birmingham / mihertz@uab.edu
7. Amy Cole / University of North Carolina School of Medicine
8. Anita Orendt / University of Utah / anita.orendt@utah.edu
9. Shannon Sheridan / Pacific Northwest Nat'l Laboratory / shannon.sheridan@pnnl.gov
10. William Rodriguez / Harvard Medical School / william_rodriguez@hms.harvard.edu
11. Eleanor Koehl / UCLA / ekoehl@oarc.ucla.edu
12. Janna Nugent / Northwestern / janna.nugent@northwestern.edu
13. Leonard Apanasevich / University of Illinois at Chicago / apana@uic.edu
14. Sergiu Sanielevici /Pittsburgh Supercomputing Center / sergiu@psc.edu
15. Cyd Burrows-Schilling / UC San Diego / cburrows@ucsd.edu
16. Daureen Nesdill/ Univ of Utah / daureen.nesdill@utah.edu
17. Andrea Denton / Univ of Virginia / ash6b@virginia.edu
18. Dave Dai / Michigan State University
19. Jim Leous / Penn State Office of the Associate CIO for Research / leous@psu.edu
20. Alper Kinaci / Northwestern University / akinaci@northwestern.edu
21. Christine Nieman / Univ. of MD, Baltimore (NNLM R1) / cnieman@hshsl.umaryland.edu
22. Brendan Kohn / University of Washington DLMP / kohrnb@uw.edu
23. Christina Divoll / Tufts University / christina.divoll@tufts.edu
24. Michael Weiner / Georgia Tech / mweiner3@gatech.edu
25. Randal Melen / Stanford University / randy@stanford.edu
26. Mike Edmonds / UC Santa Cruz / medmonds@ucsc.edu
27. Jen Simms / Indiana U Bloomington / simms29@iu.edu
28. Brian Westra / University of Iowa / brian-westra@uiowa.edu
29. John McGee / UNC-CH / mcgee@unc.edu
30. Torin White / University of Illinois Chicago / torinw@uic.edu
31. Paul van der Mark / Florida State University / pvandermark@fsu.edu
32. Amy Nurnberger / Massachusetts Institute of Technology / nurnberg@mit.edu
33. Wendy Kozlowski / Cornell University Library / wak57@cornell.edu
34. Wendy Mann / George Mason University / wmann@gmu.edu
35. Tom Phimmasen / Tufts University / tom.phimmasen@tufts.edu
36. Heather Ganshorn / University of Calgary / heather.ganshorn@ucalgary.ca
37. Lori Sloane/ University of New Mexico / lsloane@salud.unm.edu
38. Kirk M. Anne / Rochester Institute of Technology / kirk.m.anne@rit.edu
39. Jessica Pierce / Harvard Medical School / jessica_pierce@hms.harvard.edu

40. Joey Netterville / University of Central Florida / joey.netterville@ucf.edu
41. Lynnee Argabright / University of North Carolina-Wilmington / argabrightl@uncw.edu
42. Kimberly Thomas / UC San Diego / kkt008@ucsd.edu
43. Summer Wang / Ohio Supercomputer Center / xwang@osc.edu
44. Hakizumwami Birali Runesha / The University of Chicago / runesha@uchicago.edu
45. Sara Gonzales / Northwestern University / sara.gonzales2@northwestern.edu
46. Sandi Caldron / University of Illinois Urbana Champaign / caldron2@illinois.edu
47. Chris Maher / UNC Charlotte / cmaher9@uncc.edu
48. Thea Atwood / Indiana University Bloomington / tpatwood@iu.edu
49. Elizabeth Kinney/ University of British Columbia/ em.kinney@ubc.ca
50. Qian Zhang / Digital Research Alliance of Canada / qian.zhang@alliancecan.ca
51. Jennifer Huck / University of Virginia / jah2ax@virginia.edu
52. Moira Downey / NC State / mcdowney@ncsu.edu
53. Jacobus Kats / UC Riverside / bart.kats@ucr.edu
54. Jennifer Abel / University of Calgary / jennifer.abel@ucalgary.ca
55. Kat Koziar / University of California, Riverside / katherine.koziar@ucr.edu
56. Randall White / Stanford Research Computing Center / rcwhite@stanford.edu
57. Biru Zhou / McGill University / biru.zhou@mcgill.ca
58. Carlie Oakenshield / Southwestern Oklahoma State University / oakenshieldc@student.swosu.edu
59. Alastair Neil/ George Mason University/ aneil2@gmu.edu
60. Maithilee Mitra/Johns Hopkins
61. Michael Peterson / Duke University / michael.peterson@duke.edu
62. Jeanine Finn / The Claremont Colleges Library / jeanine.finn@claremont.edu
63. Randall Love / Elsevier / r.love@elsevier.com
64. Vivek J. Khandwala/ University of Cincinnati/ khandwvj@ucmail.uc.edu
65. Himanshu Sharma/ University of Illinois at Chicago / himanshu@uic.edu
66. Mark Kever / Vanderbilt University / mark.keever@vanderbilt.edu
67. Jeff Dusenberry / UMass Boston / jeff.dusenberry@umb.edu
68. Daria Orlowska / Western Michigan University / daria.orldowska@wmich.edu
69. Mike Hutcheson / Baylor University / mike_hutcheson@baylor.edu
70. Lindsay Gypin / University of North Carolina Greensboro / lagypin@uncg.edu
71. Ho Jung Yoo / UC San Diego / hjsyoo@ucsd.edu
72. Jessica Davila/University of Oklahoma Libraries/jdavila@ou.edu
73. Nina Exner / Virginia Commonwealth University/nexner@vcu.edu
74. Steve Sears / Johns Hopkins Medicine and University / ssears1@jh.edu
75. Angela Atkins / Wake Forest University / atkinsae@wfu.edu
76. Paul Levett / George Washington University prlevett@gwu.edu
77. Lucy Carr Jones / UVA Health Sciences Library / lqc3t@virginia.edu
78. Paula Sanematsu / Harvard University FAS RC / paula_sanematsu@g.harvard.edu
79. Susan Chobanoff / Penn State University / smc179@psu.edu
80. Gladys Andino / University of Virginia / gka6a@virginia.edu
81. Bob Freeman / Harvard Business School / rfreeman@hbs.edu
82. Bala Desinghu/Rutgers University/bala.desinghu@rutgers.edu
83. Andy Evans / UVM / aevans7@uvm.edu

84. Claire Mizumoto / University of California San Diego / claire@ucsd.edu
85. Mary Murphy/ University of Wisconsin, Madison RCI /murphy22@wisc.edu
86. Brad Anderson/University of Kansas Medical Center/banderson4@kumc.edu
87. Beth Gern/Penn State University/beh11@psu.edu
88. Mara Sedlins / Colorado State University / mara.sedlins@colostate.edu
89. Stacy Winchester/ University of South Carolina/ winches2@mailbox.sc.edu
90. Luca Belletti / Elsevier / l.belletti@elsevier.com

Max attendee count - 116 in zoom - 90 registered above

Notes from the call

Slides

[Lisa's slides](#)

[Kat's slides](#)

[Tobin's slides](#)

Notes

Lisa Federer (Data Science and Open Science Librarian, National Library of Medicine): Introduction to NIH Data Management and Sharing Policy

New policy covers all NIH funded research

Does not apply funding that doesn't generate data

Policy requires a Data Management and Sharing Plan (DMSP)

Not all data is appropriate to be shared, DMSP should identify those as well

Expectation that DMSP will actually be followed

Plan no more than 2 pages long

Elements of a DMSP: <https://grants.nih.gov/grants/guide/notice-files/NOT-OD-21-014.html>

PI is responsible for DMSP

Requires adequate data to replicate study findings, including null findings and non-published data

Don't need to share data that isn't final research data.

Human research data requires informed consent or deidentification - an understanding there are privacy/ethical/legal reasons not to share data sometimes

Phases: Submit/ Assess/Compliance

Compliance is monitored and may factor into future funding decisions

Guidance is very generic because it covers all of NIH, additional requirements may apply depending on the award

Allowable costs: <https://grants.nih.gov/grants/guide/notice-files/NOT-OD-21-015.html>

Selecting a repo: https://www.nlm.nih.gov/NIHbmic/nih_data_sharing_repositories.html

Biomedical data researcher guide:

https://nap.nationalacademies.org/resource/25639/Researcher_User_Guide_Biomedical_Data.pdf

Additional considerations for human data:

<https://grants.nih.gov/grants/guide/notice-files/NOT-OD-21-016.html>

Sharing at NIH: <https://sharing.nih.gov/>

Kat Milligan-McClellan (Assistant Professor of Molecular and Cell Biology / Microbiology, UConn): Researcher challenges

Host-microbe interactions - microbiome research

How do the complex communities within the host affect the development of the host - eco-toxocology

Microbe-microbe and host-microbe interactions

3-spine stickleback, found in Alaskan lakes as a research model

Each Alaskan lake is distinct and can be compared to identify variations in stickleback gut microbiota

Isolate gut microbiota from the fish, perform bioinformatic analysis, identify the microbiotic communities - use Illumina machines to get FastQ files - 22/24GB per run, 2 for each fish sample

Data generated:

FastQ file (should be included in the plan)

Trimmed and sorted table

Primers and protocols (should be included in the plan)--data to go to Protocols.io

Metadata files (.csv) (should be included in the plan)

Read me

Pipeline .txt file

Analysis output

Images of fish (should be included in the plan)

Behavioral data with videos (600Kb per video, many videos)

Stored on a google drive, researcher isn't sure this is the right destination

How will the NIH policy affect Kat?

www.openscapes.org

Tobin Magle (Lead Data Management Specialist / Northwestern University): How can researcher-facing staff provide support?

Data management required in the entire data lifecycle - places we can help throughout

Stay calm

Objections can have responses, for example planning ahead is good for research, etc.

Refer - who else can help on campus?

You don't need to be the expert in all of this, but know who is on campus - data librarians

Know the data storage options on campus, esp. For regulated data

Not everyone has access to the same resources - schools may be different

<https://dmptool.org>

Finding repos
Cloud resources for research data archiving (STRIDES discounts may apply)
Learn and sneak it into your own practices
Write a data management plan
Integrate these topics into other more popular workshop topics
Collaborate
Train together
Identify gaps
Advocate for resources to fill gaps

Questions

Question for Lisa: after the end of the funding period, non-compliance with the NIH ICO-approved DMSP might impact future funding decisions. Does this have an impact for the recipient's institution or for the individual researcher for future funding opportunities?

Is the requirement for data sharing on the side of data deposit in a reputable repository with PID or is it more about available data access? Sharing software codes might get quite challenging.

Question for Kat (answered already): Why feel ashamed about using Google Drive?

- Answer: Google Drive is not a repository for long term preservation from data curation perspective, and doesn't address needs like migration. Though for data storage during the project, there is Google for education which provides security
-

Does the NIH DMS Policy mandate sharing all de-identified data this lab generates, measurements + analytical data? Or just the analysis?

Do people have workflows where funding requests and grant applications encumbered by such requirement need to be cleared by the research computing group to make sure appropriate funding requests for storage have been made?

Can repos for NIH-funded data also include deposits of non-NIH funded research data?

Chat Comments

How will the slides from Lisa Federer be distributed?? Pls share. Thank you.

Please use if you find helpful, lots of links! Guidance: We've written both a "Keypoints" and a "Resources" document to help our community key in to what they need to pay attention to, and where to go for help.

- Keypoints:
https://www.dropbox.com/s/qlfbh9f3vt753uo/Handout_nih2020dmisp_summary.pdf?dl=0

- Resources:
https://www.dropbox.com/s/tk8btffuyul98sy/Handout_NIH_DMSP_MITresources.pdf?dl=0