WEBINAR

How to Make Your Little Data Big by Being FAIR

Thursday, September 2, 2021
11:00 AM—1:00 PM Eastern U.S.

Speakers:
• Dr Susan Gregurick (NIH)
• Dr Sean Watford (US EPA)
• Dr Michelle Heacock (NIEHS)
• Dr Christine Kirkpatrick (SDSC)

Moderator: Dr George Woodall (US EPA, SLC)

Abstract
Connecting and tagging your digital information can make it reusable to others. Your research can have a greater impact and you might find new partnerships to further the science. How to do this? By implementing FAIR—a set of guiding principles to make data Findable, Accessible, Interoperable and Reusable. Building upon the SLC webinar *Big Data: What Is It and What Does It Mean to Me?* (April 2020), four speakers have been recruited to explain and translate how to use and apply FAIR principles across disciplines. An interactive discussion session is planned to close out the webinar to address your questions. The intent is to continue the conversation in a Breakout Session at the *Future Tox V* meeting scheduled for November 8 and 9 2021.

Program
• Dr Susan Gregurick (NIH) will provide an overview of leading-edge concepts in Big Data and FAIR Principles (30 minutes)
• Case Studies demonstrating application of FAIR data principles to advance applied public health (20 minutes each)
  o Dr Michelle Heacock (NIEHS)
    “Data Integration Activities in the Superfund Research Program”
    This presentation will provide an overview of data sharing activities within the Superfund Research Program (SRP). It will highlight lessons learned and recommendations from data sharing use cases that were designed to expand and enhance data interoperability and reuse to advance a research question.
  o Dr Sean Watford (US EPA)
“Interoperability of Health Effects Data: Toxicity Reference Database (ToxRefDB)”
In this presentation, I’ll introduce topics related to interoperability including controlled vocabularies, cross-referencing to other resources, and data exchange standards. I then provide a deeper look into how we redeveloped a health effects database, Toxicity Reference Database (ToxRefDB), making the resource more interoperable and amenable for use in research.

- Dr Christine Kirkpatrick (GO FAIR and SDSC) will promote applying big data principles and incorporation of disparate data across broad disciplines (30 minutes)
  “Why FAIR Data Matters”
  Scientific data is produced at an increasing pace and is recognized as an asset to be protected. We count on being able to mine vast amounts of data as well as integrate between new types of data not combined previously. Much of the data produced is not ready for such transformation. Scientists spend inordinate amounts of time searching for and then cleaning data. The FAIR data movement gives a path forward for preparing data in ways that allow us to realize its full value.

- Discussion Session will be moderated by Dr George Woodall (US EPA, SLC) (20 minutes)

This free, live webinar is open to members of societies in the Scientific Liaison Coalition.
Advance registration is required.

Registration Link

Speaker Bios

Susan K. Gregurick, PhD is Associate Director for Data Science and Director of the Office of Data Science Strategy (ODSS) at the National Institutes of Health (NIH). Under her leadership, the ODSS leads the implementation of the NIH Strategic Plan for Data Science through scientific, technical, and operational collaboration with the institutes, centers, and offices that comprise NIH. Dr Gregurick received the 2020 Leadership in Biological Sciences Award from the Washington Academy of Sciences for her work in this role. She was instrumental in the creation of the ODSS in 2018 and served as a senior advisor to the office until being named to her current position.

Michelle Heacock, PhD is a Program Officer with the National Institute of Environmental Health Sciences. In this role she oversees grants that support research in DNA repair and telomere biology, and for the Superfund Research Program (SRP), multi-project grants that span biomedical science with environmental science and engineering. She also is the lead for the data science and sharing for the SRP, where she helps to foster activities that focus on
increasing the findability, accessibility, interoperability, and reusability (FAIR) of SRP datasets, emphasizing data interoperability and reusability.

**Christine Kirkpatrick, PhD** oversees the San Diego Supercomputer Center’s (SDSC) Research Data Services division, which manages infrastructure, networking, and services for research projects of regional and national scope. Dr Kirkpatrick’s expertise is in the implementation of research computing services and operational cyberinfrastructure (CI) at scale. She founded and hosts the US GO FAIR Office at SDSC, is on the leadership of the West Big Data Innovation Hub, the Open Storage Network, and is PI of the EarthCube Office (ECO). She serves on the National Academies of Sciences’ US National Committee for the Committee on Data (CODATA).

**Sean Watford, MSPH, PhD** currently works as an environmental health data scientist at US EPA/ORD Center for Public Health and Environmental Assessment. His work is focused on making EPA data and systems work better together (i.e., interoperability) both internally and externally for increased transparency, easier access to information, and availability of resources for use in research.