

# **NIAID Data Science and Big Data**

**Rosemary G. McKaig, MPH, PhD  
Epidemiology Branch  
Basic Sciences Program, Division of AIDS**

[rmckaig@niaid.nih.gov](mailto:rmckaig@niaid.nih.gov)



National Institute of  
Allergy and  
Infectious Diseases







PERMIT  
PARKING

5615  
ROSELAND  
12712  
BUCK CREEK  
HILL ROAD

Fishers

12712





# Office of Data Science and Emerging Technologies = NIAID Office of Data Science



**DAIDS**

Division of AIDS



**DMID**

Division of Microbiology and Infectious Diseases

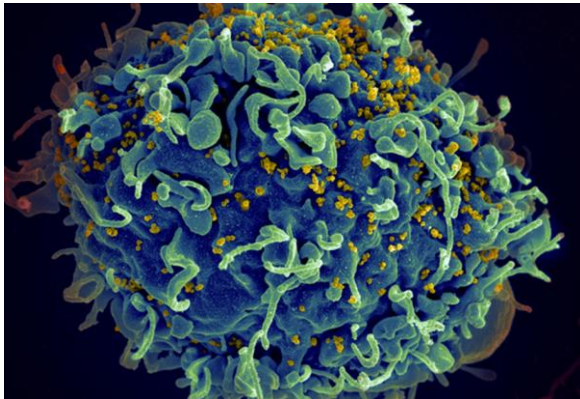


**DAIT**

Division of Allergy, Immunology, and Transplantation

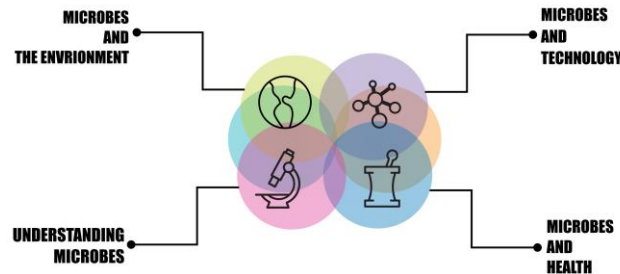
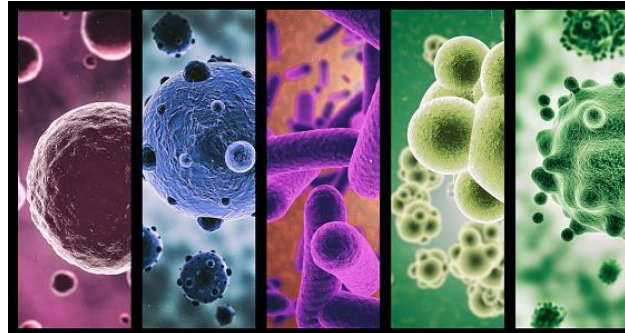
# DAIDS

Division of AIDS



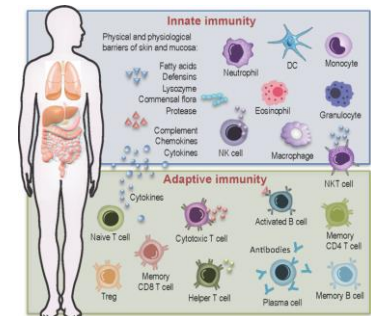
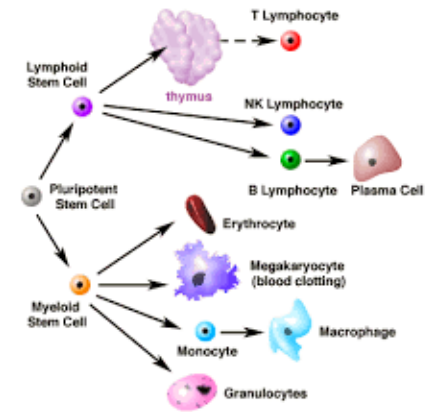
# DMID

Division of Microbiology and Infectious Diseases

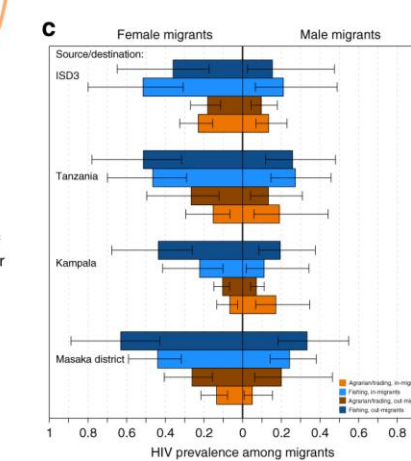
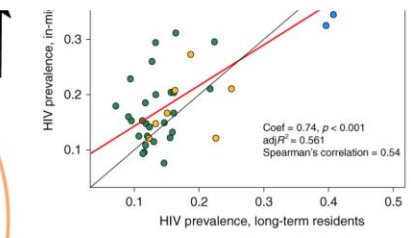
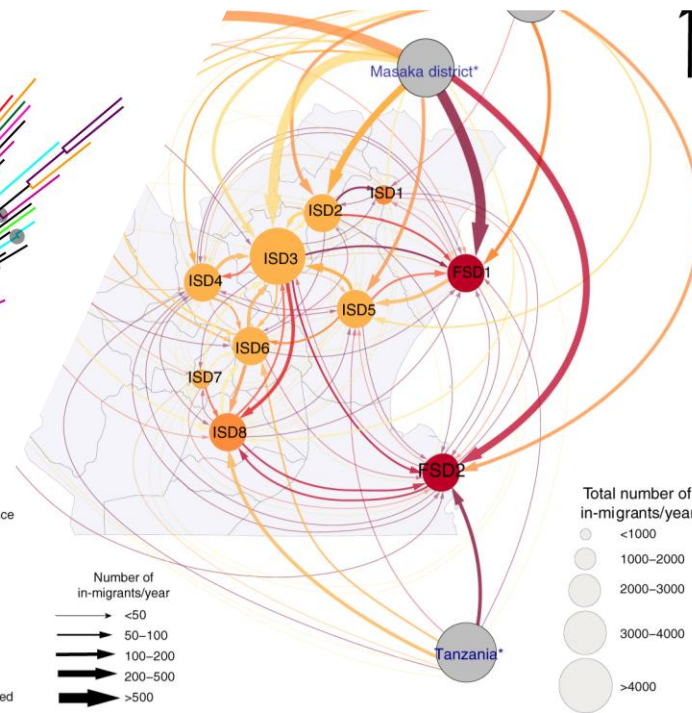
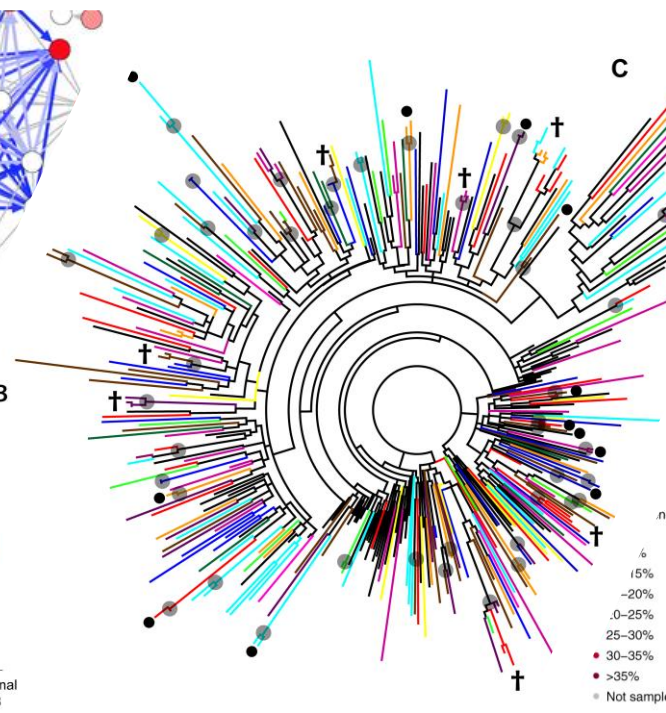
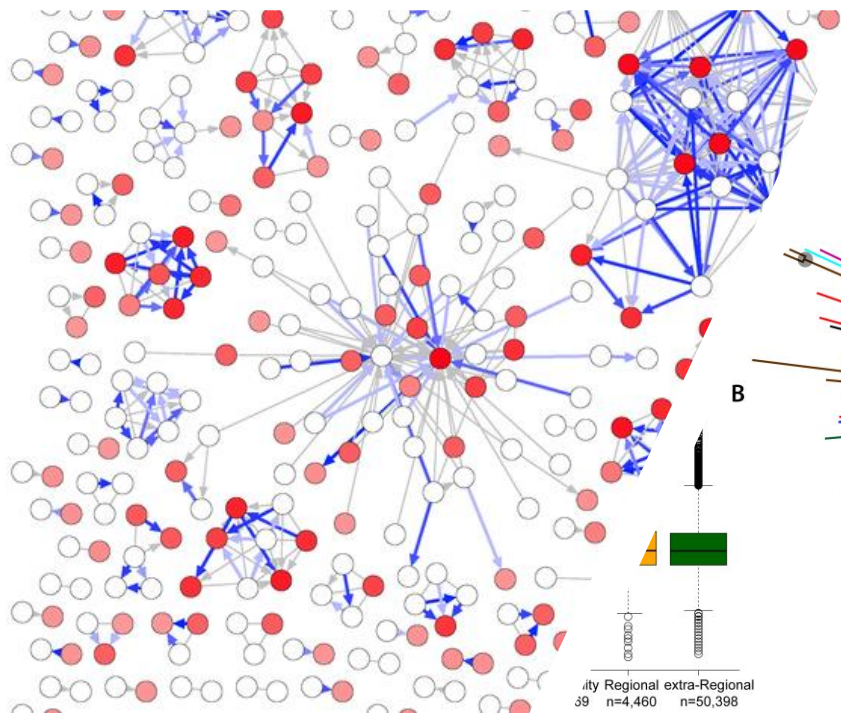


# DAIT

Division of Allergy, Immunology, and Transplantation










## Challenges and best practices for digital unstructured data enrichment in health research: A systematic narrative review

Jana Sedlakova, Paola Daniore, Andrea Horn Wintsch, Markus Wolf, Mina Stanikic, Christina Haag, Chloé Sieber, Gerold Schneider, Kaspar Staub, Dominik Alois Ettlin, Oliver Grübner, Fabio Rinaldi, Viktor von Wyl ,  
for the University of Zurich Digital Society Initiative (UZH-DSI) Health Community

Published: October 11, 2023 • <https://doi.org/10.1371/journal.pdig.0000347>



**1. Lack of meta-information**

**2. Standardization issues**

**3. Data quality: inaccuracies, bias, confounding**

**4. Infrastructure and Human Resources**

**5. Finding suitable analysis tools, methods, and techniques**

**6. Alignment with a research design and/or research question**

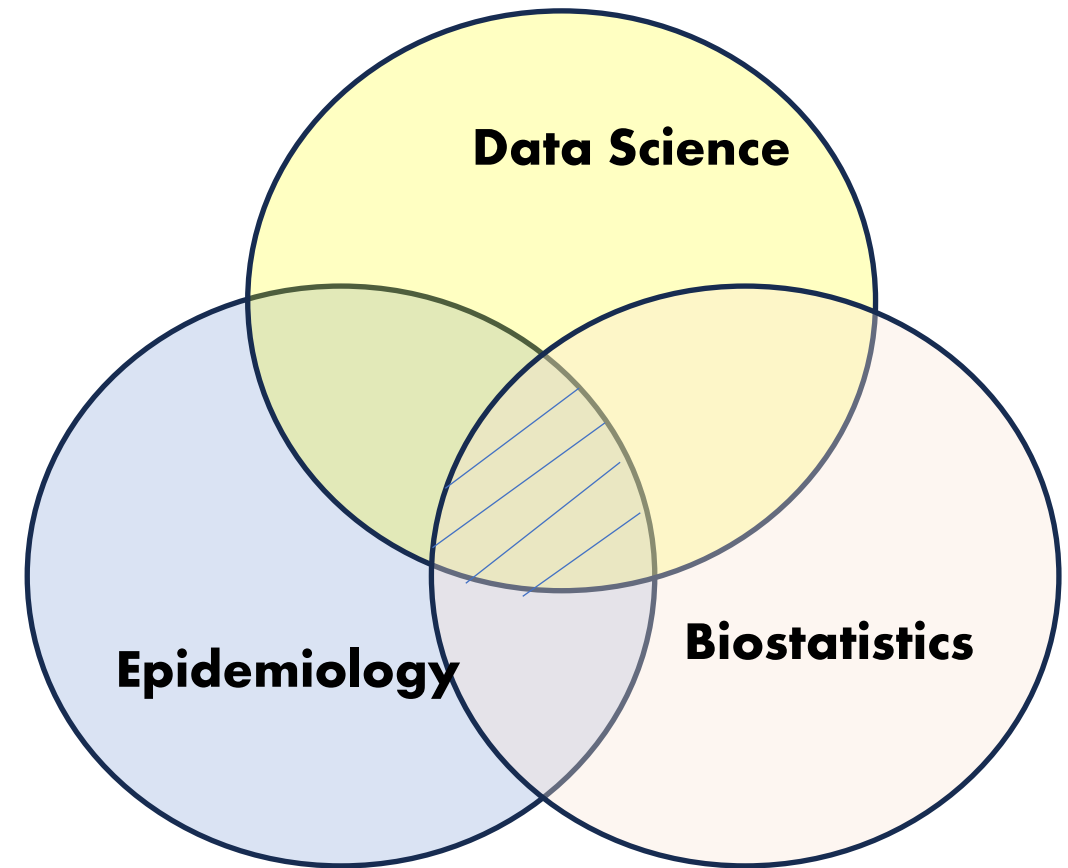
**7. Ethics & Legal Issues**

# On the Convergence of Epidemiology, Biostatistics, and Data Science

Neal D. Goldstein<sup>1</sup> Michael LeVasseur<sup>1</sup> Leslie A. McClure<sup>1</sup>

<sup>1</sup>Department of Epidemiology and Biostatistics, Dornsife School of Public Health, Philadelphia, Pennsylvania, United States of America

Published on: Sep 29, 2020



“It is important to note that while increasing sample sizes (as the ‘big data’ movement is witness to) can increase the precision of an estimate, it does nothing to mitigate the effects of bias.”





[Home](#) > [EP Home](#)

# NLM's University-based Biomedical Informatics and Data Science Research Training Programs

<https://grants.nih.gov/grants/guide/rfa-files/RFA-LM-21-001.html>

The National Library of Medicine supports research training in biomedical informatics and data science at eighteen educational institutions in the United States. NLM's T15 program offers graduate education and postdoctoral training and research experiences in a wide range of areas including health care informatics, translational bioinformatics, clinical informatics, public health informatics, and consumer health informatics. Trainees will receive exposure to a core curriculum focused on biomedical data science concepts and methods as well as develop skills needed to lead independent future research. Some T15 programs offer an additional area of training in HIV/AIDS. Please contact T15 awarded institutions for questions related to trainee selection, eligibility, program specifics, and levels of support. The location of NLM T15 training institutions and contact information for awardees is provided in the interactive map below.