

Developing Experts for Better Biomedical and Behavioral Research Data: FAIR and AI/ML-Ready Data

Background: Artificial intelligence and machine learning (AI/ML) are a collection of data-driven technologies with the potential to significantly advance biomedical research. Much of this potential is unrealized, however, because biomedical data are not collected and prepared in ways that would allow them to be used efficiently and effectively by AI/ML applications. The task of making data FAIR and AI/ML-ready is not only algorithmic. It requires multi-disciplinary expertise, experimentation and, often, iterative feedback from AI/ML applications and experts. Particularly for biomedical data, AI/ML-readiness should be guided by a concern for human and clinical impact.

Program Goal: To develop new specialists able to make biomedical and behavioral research data FAIR and AI/ML-ready, and to advance the field of data science for AI in biomedicine and behavioral science.

Components of the Proposed Initiative:

- **Inter-disciplinary Education and Training Centers:** The Centers will be comprised of multi-department partnerships that reflect the interdisciplinary skills and competencies needed to make biomedical and behavioral research data FAIR and AI/ML ready. The centers will also be multi-institutional partnerships that are able to recruit new talent into the training programs. Multi-institutional partnerships should also help develop training opportunities that feature relevant and culturally sensitive data and use cases.
- **Coordinating Center:** A coordinating center will organize workshops and other events to share best practices and strengthen the network of Centers. The coordinating center will broker cross-center training opportunities for students, and coordinate annual workshop and meetings for trainees so that a community of experts is developed.

Deliverables:

- **Diverse cohort of new specialists with expertise at the interface of information science, AI, biomedical and behavioral sciences:** These experts will engage with biomedical research teams throughout the data life-cycle, beginning at research planning; help ensure best practices in FAIR data management are followed; and facilitate engagement with AI/ML community. In addition, these experts will gain unique insights into the needs of biomedical and behavioral science researchers and AI/ML communities and will be able to work in interdisciplinary teams to innovate new practices and capabilities for making biomedical and behavioral research data AI/ML ready.

Budget:

\$10M / year to support ~ 10 multi-institutional, 5-year centers; coordination center; and workshops.