NCPI Workshop Fall 2024: Readout (Public Version)

The National Institute of Health (NIH) Cloud Platform Interoperability (NCPI) 2024 Fall Workshop held on September 25-26, 2024, in Rockville, MD, convened researchers, developers, and stakeholders from NCPI systems and NIH institutes. The workshop aimed to advance NCPI's mission of enhancing data interoperability, collaboration, dismantling silos, and fostering connectivity within the biomedical research ecosystem.

Goals of the Workshop

The primary objectives of this workshop were to:

1. Define NCPI and its offerings.
2. Hear updates from the NCPI interoperability projects.
3. Coordinate product roadmaps among NCPI partners.
4. Foster a culture of collaboration and communication.
5. Identify and begin to address critical challenges in data integration and interoperability.

Key Themes and Discussions

Discussions and Sessions at the September 2024 NCPI workshop addressed the following key themes:

- **Defining NCPI and Enhancing Collaboration:** Participants emphasized the need for clear definitions of NCPI's offerings and coordinated product roadmaps. Roundtable discussions highlighted practical solutions for data integration and the importance of community engagement. A systematic approach to tackling data integration challenges incrementally was agreed upon.
- **Interoperability Project Progress and Challenges:** Project members noted challenges with data model harmonization. The emerging theme from these discussions was that data harmonization is vital to increasing interoperability across partner systems. As the NCPI partner systems utilize the Fast Healthcare Interoperability Resources (FHIR) standards to address this issue, it was noted that the intentionally broad and varied nature of the FHIR standard requires specific harmonization to ensure consistency. Projects also highlighted the difficulties of protected data management across platforms. While open-access datasets are useful in pipeline development, appropriate datasets for development and testing are not always open-access. More information about these projects is available at the NCPI GitHub.
- **Establishing Standards for Data and Tools:** The workshop stressed the importance of standardizing data and tools across platforms to ensure interoperability. Collaboration and technical innovation were highlighted as crucial for comprehensive data analysis to make genomic resources accessible to a broader audience.
- **Sustainable Funding and Cost Management:** NCPI Partner Systems and Interoperability Projects discussed the challenges in developing a sustainable funding model and

managing costs, including egress fees and cloud credits. The need for comprehensive training, documentation, and tutorials to expand community knowledge on cost management was highlighted.

- **NCPI GA4GH Driver Project:** Workshop attendees also discussed the role of NCPI as a driver project for the Global Alliance for Genomics and Health (GA4GH). As NCPI is the largest repository of biomedical data currently available, it presents a unique opportunity to test, implement, and refine the standards of GA4GH to the benefit of both organizations. This workshop specifically focused on updates to GA4GH's Data Repository Standard (DRS) and discussions around the future of the GA4GH Passport specification. These updates are important for NCPI as DRS is a fundamental service enabling data set retrieval and the GA4GH passport specification is the basis for the Researcher Auth Service being implemented across NCPI to enable controlled data access through researcher authentication and authorization.

Strategic Plans and Future Directions

Looking ahead, the NCPI program aims to:

1. Define and progress along a flexible and adaptable tactical roadmap that addresses both immediate and long-term needs.
2. Enhance data discoverability and access interoperability.
3. Harmonize data and tools across platforms.
4. Develop comprehensive documentation of current partner systems and platforms, including exposed system endpoints and implemented technology.
5. Collaborate with the Global Alliance for Genomics and Health (GA4GH) to take advantage of currently available interoperability standards and to identify gaps in the current standards.

The NCPI 2024 Fall Workshop highlighted the importance of collaboration, technical innovation, and community engagement in advancing the program's mission. By moving towards an NCPI "map" through a bottom-up approach, addressing cross-system and cross-platform needs, and incorporating pre-existing interoperability solutions, the NCPI program can enhance data interoperability and drive scientific innovation in the biomedical research community. These efforts will ensure that the NCPI program remains at the forefront of biomedical research, fostering an environment where data is findable, accessible, interoperable, and reusable, ensuring it is utilized to its fullest potential to address complex health challenges and improve patient outcomes.