NCPI Virtual Touchpoint – Readout (Public Version)

January 19, 2024

The NIH Cloud Platform Interoperability (NCPI) program hosted a Virtual Touchpoint event with over 70 attendees from across different NIH Institutes and Centers (ICs), the NIH Office of the Director (OD) Office of Data Science Strategy (ODSS), and various non-NIH institutions that participate in the development of NCPI Participating Partner Systems. These partner systems are: NHGRI's Genomic Data Science Analysis, Visualization, and Informatics Lab-space (AnVIL), NHLBI's BioData Catalyst (BDC), NCI's Cancer Research Data Commons (CRDC), the Common Fund-supported Gabriella Miller Kids First Data Resource Center (KFDRC), and NLM-NCBI's database of Genotypes and Phenotypes (dbGaP). The goals of the event included:

- bringing the NCPI community together to discuss interoperability topics,
- providing a forum for NCPI contributors to share thoughts, ideas, and concerns, and
- maintaining momentum in the community until the next NCPI Workshop.

This summary was prepared by the NCPI Administrative Coordinating Center (NCPI ACC).

Participant Summary

Approximately half of the Touchpoint session participants were from ICs or institutions not directly involved with NCPI program management or coordination, as shown in Figure 1.
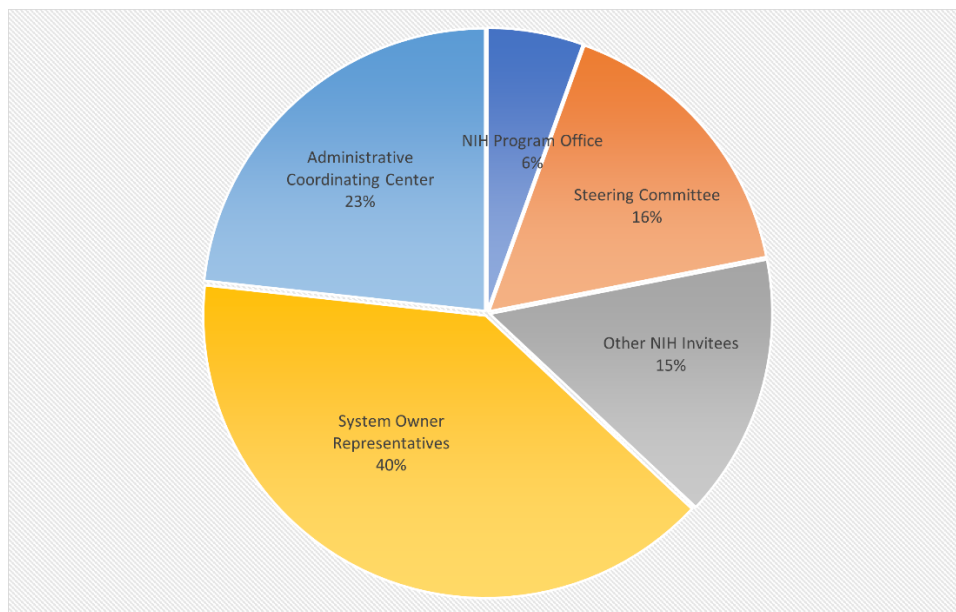

Figure 1: Participant Representation

Welcome & Overview of NCPI Program

The NIH ODSS/NIH NCPI Program Official (PO) opened the Touchpoint event with an overview of the NCPI program. The NCPI program is overseen by NIH ODSS and is a collaboration between NIH data repositories and analysis systems. This collaboration includes NIH representatives, analysis platform team members, and researchers using those platforms. The NCPI program exists to help realize the NIH's vision of a trans-NIH ederated data ecosystem with the collaboration of NCPI Partner Systems: AnVIL, BDC, CRDC, dbGaP, and KFDRC. The program's goal is to expand and connect additional data systems across NIH—aligned with the principles of data findability, accessibility, interoperability, and reusability (FAIR).

The NIH ODSS/NIH NCPI PO also highlighted strategic direction for the NCPI Program as the focus of the Touchpoint, as well as its purpose to serve a precursor for a deeper dive into strategy during the upcoming NCPI Workshop.

NCPI Guiding Principles: Brief Reflection

An overview of the NCPI Guiding Principles was provided to "plant the seed" for broader discussion of revising them over the coming year. Themes surfaced from this session included:

- data flows are essential, not siloes – and how do we enhance flows, rather than exchanging blobs of data?
- support and re-using current API and data standards,
- the importance of focusing on use cases, and user needs,
- prioritize, don't get distracted by "shiny things", and
- we are in this together – collaborate!

Revising the principles will be informed by the NCPI Program's vision and mission. Additionally, the principles can inform metrics that will determine whether NCPI program goals have been met.

Strategy-Based Discussion

The Strategy-Based Discussion session involved a series of breakout sessions, with touchpoint attendees randomly assigned to breakout rooms. Discussion topics included interoperability projects, working groups, intra-NCPI collaboration, and the approach to NIH Researcher Auth Service (RAS) integration across the NCPI ecosystem. Participants contributed thoughts and ideas associated with strengths, areas of improvement, ideas, and questions associated with a set of high-level NCPI-relevant topics.

During the session, attendees noted that the existence of the NCPI program has improved interoperability across NIH biomedical data platforms, and that the ecosystem could benefit from alignment on development timelines, tools, and data elements. Similarly, participants expressed a desire to improve knowledge sharing across different teams that were handling RAS integrations on different systems and noted that engagement with GA4GH has been useful in informing system implementations to improve interoperability (e.g., via use of the GA4GH Passport and the Data Repository Service standards).

NCPI Working Groups

This session included a review of the background of NCPI Working Groups (WGs) and take-aways from early conversations between the ACC and WG chairs/co-chairs followed by a group discussion of successes, challenges, and opportunities to ensure successful working groups. The session concluded with high-level next steps for NCPI WGs. The five original WGs were created by the NCPI program to address program needs for:

- Governance guiding principles
- Technology solutions for sharing complex clinical and phenotypic data
- Training and outreach resources to acclimate users to the NCPI ecosystem
- Dataset and cohort discovery
- Technical standards for cross-platform data exchange

During the Touchpoint discussion, attendees noted that the WG discussions allowed for key technical contributions from platform representatives, as well as a project-agnostic shared vision of targeting interoperability infrastructure via researcher-driven use cases and coordinated implementation rollout. They also indicated a desire for increased clarity of WG scope and that WGs should account for different stakeholders as well as the interplay between scientific, infrastructure, and programmatic goals.

NCPI and GA4GH

This session highlighted the collaborative relationship between GA4GH and NCPI, implementation of GA4GH standards, and a discussion about potential priorities and future direction for engagement. GA4GH contributes to NCPI with global collaborations, data sources, and critical standards that facilitate NCPI's goals. Conversely, NCPI provides GA4GH with real-world use cases, data, and practical implementations that help in testing and refining GA4GH's standards. The group also discussed several GA4GH products that NCPI is currently advancing, including [Passports](#), the [Data Use Ontology](#), the [Data Repository Service (DRS)](#), and genomic file standards (e.g., [CRAM](#)), as well as additional products that NCPI can focus on while optimizing NCPI's engagement with GA4GH.

Upcoming NCPI Workshops

This session included a presentation of a tentative agenda for upcoming NCPI Workshops and solicitation of feedback from attendees. Planning for the next workshop will be collaborative. The NCPI ACC plans to form an Advisory/Planning Committee to help establish the agenda and sessions, and the results of the Virtual Touchpoint will be used to plan and prepare for the workshop. The next workshop will be limited to NCPI core participants. We are aiming to include sessions for NCPI Interoperability Projects to engage with the larger community and across projects. We are also planning to include work on reviewing and discussing the NCPI vision and mission.

Closeout Summary

The Touchpoint Closeout included a summary of touchpoint themes and next steps that emerged from the day's sessions. Key points included the observation that considering user needs, avoiding silos in data flows, and leveraging standards and APIs will be important to improve interoperability. Finally, on the upcoming Workshop, input from the NCPI community will be critical as well as having a "North Star" diagram for interoperability. The NCPI community will need to define concrete products expected from the Workshop, focus on specific interoperability use cases, and identify cross-cutting challenges. A final theme from the Workshop session was associated with consideration of growth—bringing other NIH efforts into the NCPI ecosystem.

Closing Remarks

In closing remarks, the NIH ODSS/NIH NCPI PO thanked the attendees and the NCPI ACC for their support in making the event a success. They noted a common desire for policy, timeline, and goal alignment across organizations within the NCPI ecosystem. The PO reiterated the NCPI program's goal of enabling and enhancing a federated data ecosystem across NIH and cited Tim Berners-Lee: "data is a precious thing and will last longer than the systems themselves". This concept applies across the NCPI community and FAIR stewardship of data should be maintained. Finally, the PO closed the Virtual Touchpoint mentioning the upcoming NCPI Workshop and thanked the attendees for their participation.