Improving Team Science Through "Thons"

Reflections on the April Olympians Community Event

Clair Blacketer[1,2]*, Melanie Philofsky[3]*, Evanette Burrows[1]
Maxim Moinat[2], Katy Sadowski[4]
[1]Janssen Research & Development, Raritan, NJ, [2]Department of Medical Informatics, Erasmus MC, Rotterdam, NL, [3]Odysseus Data Services, Inc., Cambridge, MA, USA, [4]Boehringer Ingelheim Pharmaceuticals, Inc., Ridgefield, CT
**co-first authors*

Background

The Observational Health Data Sciences and Informatics (OHDSI) community has long championed collaborative efforts to accelerate progress in the health data science domain. Regular community events, such as hack-a-thons, Phenotype Phebruary[1], and study-a-thons, serve as platforms for intense, short-term, team-driven science. Each event is tailored to maximize productivity through diverse methodologies suited to distinct group objectives.

In April 2024, the Common Data Model (CDM), THEMIS, and Data Quality (DQ) working groups convened for a month-long initiative called April Olympians. The aim was to develop a comprehensive community resource library outlining conventions for data transformation to the CDM. This abstract encapsulates our methodology, results, and insights on optimizing global team science.

Methods

Preparation Phase

Leading up to the event, the CDM, DQ, and THEMIS working group leaders engaged in weekly one-hour planning sessions over several weeks. These meetings were crucial for aligning goals and preparing necessary materials. The entire event's workflow was meticulously tracked using a GitHub project, with tasks segmented into three specialized teams:

1. Hunter Team: Focused on identifying and collecting ratified CDM conventions.
2. Writer Team: Tasked with documenting the conventions for inclusion in the library.
3. Builder Team: Responsible for building out the resource library.

Detailed field guides were crafted for each team, supplemented by templates to ensure consistent and thorough documentation. Team leads were appointed to manage

[1] https://www.ohdsi.org/phenotype-phebruary-2024/

participants and address queries, with the overarching strategy vetted by the OHDSI Steering Committee. Multiple communication channels, including twice a week check-in calls, email, posts to the collaborative Teams workspace, direct calls and messages, were established to maintain seamless interaction.

Two weeks prior to kick-off, the event was announced on an OHDSI community call and the intent of April Olympians was detailed as well as the team names and leads. A form was shared allowing interested parties to sign up for the team of their choice. This helped to gauge interest and determine which teams could expect to have the most involvement.

Execution Phase

The event commenced with a one-hour kickoff meeting to establish clear expectations. To accommodate global schedules, bi-weekly 30-minute check-ins were held in different time zones. Tasks were deliberately broken down into manageable 15-minute segments to promote steady progress and engagement. Participants were empowered to autonomously open GitHub tickets, as detailed in the preparatory materials. However, early on we found permissions challenging as we had only tested with members who already had access to the OHDSI GitHub environment.

The team leads prioritized rapid response to questions and issues, leveraging both Teams and GitHub for communication. Comprehensive debates and discussions were held to ensure ratified CDM conventions were clear and universally understood.

Throughout the month the event's progress was communicated to the broader OHDSI community through weekly calls, fostering transparency and inclusion. In addition, each call was focused on a different topic that was relevant to data transformation. These included highlighting available tools to aid in the ETL process, community-based data characterization, data quality methods[1], and an overview of the standard vocabularies[2]. Keeping the topics in line with the mission of April Olympians helped to show the importance of the work and maintain the mindset of the event.

Results

By the end of April 2024, the collaborative effort yielded a robust knowledge base of CDM conventions, as depicted in Figure 1. The group successfully closed over 80 GitHub issues, driven by 20 consistent contributors from the OHDSI community.

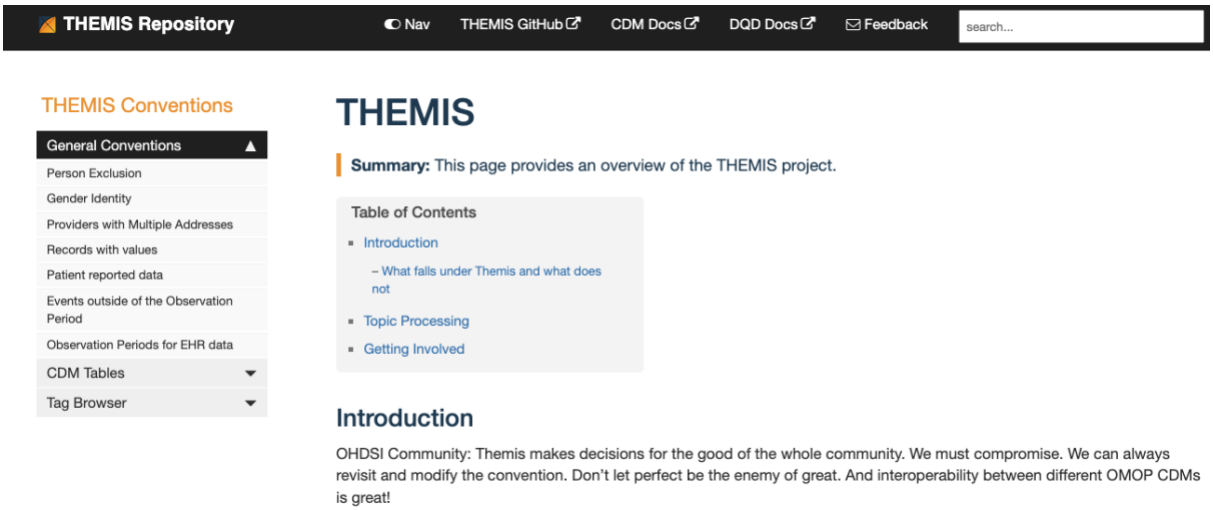

Figure 1. THEMIS Repository (https://ohdsi.github.io/Themis/)

Insights from Team Science

Throughout the process, we gleaned several key insights into effective team science practices.

Do:

- **Prepare comprehensive materials**: Detailed guides and templates are crucial for participant engagement and productivity.
- **Be responsive**: Quick responses to queries and issues help maintain momentum and morale.
- **Break tasks into small chunks**: Dividing work into 15-minute segments makes it manageable and less intimidating.
- **Test permissions**: Ensure the process for gaining access to the tools is well understood and tested.
- **Empower participants**: Allowing autonomy fosters a sense of ownership and accountability.
- **Recognize contributions**: Acknowledging efforts encourages continued participation.
- **Regular read-outs**: Weekly updates keep the team accountable and informed.

Don't:

- **Use intimidating language**: Tasks perceived as requiring extensive knowledge or long-term community involvement can deter participation.
- **Over-schedule meetings**: Well-prepared materials and responsive leads can reduce the need for frequent meetings, allowing for more asynchronous work.
- **Create complex task descriptions**: Simplified task descriptions can lower the barrier to entry and encourage broader participation.

Conclusion

The month-long collaborative event culminated in the creation of a much-needed THEMIS convention library for the OHDSI community. The initiative not only achieved its primary objective but also provided valuable lessons on the nuances of team science. Comprehensive preparation and clear, accessible communication were identified as critical factors for success, while avoiding overly complex or responsibility-laden language is essential to maintain high levels of engagement.

References

1 Blacketer C, Defalco FJ, Ryan PB, *et al.* Increasing trust in real-world evidence through evaluation of observational data quality. *Journal of the American Medical Informatics Association : JAMIA*. Published Online First: 27 July 2021. doi: 10.1093/jamia/ocab132

2 Reich C, Ostropolets A, Ryan P, *et al.* OHDSI Standardized Vocabularies—a large-scale centralized reference ontology for international data harmonization. *Journal of the American Medical Informatics Association*. 2024;31:583–90.