Unleashing Community-Wide Research Potential: OHDSI Lab 2.0 – Empowering Scientists, Eliminating IT Hassles

Kristin Kostka[1,2], David Madigan[1], Christian Reich[1,4], Asieh Golozar[1,3], Brianne Olivieri-Mui[1], Daniel Pitch[1], Justin Manjourides[1], Aleksei Gorodetskii[3], Deepa Reddy[3], Juan Carlos Ñamendi Pineda[3], Osmar Benevidez[3], Konstantin Yaroshovets[3], Peter Berzin[3], Gregory Klebanov[3]

[1]The OHDSI Center at the Roux Institute, Northeastern University, Portland, ME, USA, [2]Nuffield Department of Orthopaedics, Rheumatology, and Musculoskeletal Sciences, University of Oxford, UK, [3]Odysseus Data Services, Inc, [4]Department of Medical Informatics, Erasmus University Medical Center, Rotterdam, The Netherlands

Background

The Observational Health Data Sciences and Informatics (OHDSI) collaborative offers a suite of powerful open-source analytics tools for advancing research in healthcare analytics. [1,2] However, implementing these tools within individual academic labs presents significant challenges, including requiring significant IT expertise as well as issues getting access to actual real-world data sets. Many cannot overcome these barriers and are lost to attrition. As a result, highly skilled researchers are unable to fully leverage the potential of OHDSI. To address this issue, a shared services version of the OHDSI tool stack was developed, known as the OHDSI Lab 2.0. This innovative platform aims to eliminate the burdensome IT requirements and provide researchers with a self-service model for accessing and utilizing a comprehensive ecosystem of resources and tools.

Methods

The OHDSI Lab 2.0 was designed and implemented to provide researchers with an on-demand, scalable model for accessing and utilizing a shared platform to support statistical methods and machine learning applied to clinical research problems within the OHDSI community.

Figure 1 shows the OHDSI Lab 2.0 Instance consists of the following key physical components:

- A cloud EC2 image
- Pre-installed and pre-configured components:
    - ATLAS v2.10.2 – cohort builder and study design, hosted on a shared EC2 instance or cluster and just link be provided in the workspace
    - ACHILLES 1.7 – full database characterization (part of ATLAS)
    - R-Shiny – R-based visualization environment
    - R Studio – R-environment to create R-based modules
    - HADES – OHDSI R-based analytical library
    - Python – support to use Python

Our current deployment is executed within an Amazon Web Services environment (as detailed in Figure 2) with potential for extending this to other popular cloud platforms. The environment offers OMOP-formatted datasets, a pre-specified OHDSI tool stack, and user access and resource management functionality. Researchers can specify their desired data, tools, compute resources, and budget, enabling them to focus on their scientific experiments rather than administrative IT tasks. Figure 2 shows the easy-to-use interface for configuring workspaces on-demand. All research is conducted in compliance with a secure, private cloud infrastructure in compliance with data license agreements and maintains audit trails for usage of underlying data sets. The university maintains relevant Institutional Review Board protocols

for research projects conducted in the environment. All users are required to complete CITI Training and receive onboarding guides to ensure compliant use of the infrastructure.

Results

The first minimum viable product of the OHDSI Lab 2.0 was released in Summer 2022. After a year of iterative testing and continued development with end users, the shared services provided by the OHDSI Lab 2.0 have demonstrate significant benefits for independent researchers. Over 35 users are actively utilizing the environment to advance research project. No longer are researchers required to have their own IT teams or governance processes for onboarding data. The platform has minimized administrative burdens by offering a centralized solution and workflows for getting their hands on OHDSI community tools and allows researchers to dedicate more time to their research efforts. Redundancy of infrastructure investments has been reduced. New faculty report a significant opportunity to optimize start-up packages where funding can be pooled for purchasing data. Furthermore, the provision of secure databases and the elimination of the need for individual research computing groups have created a safer and more efficient environment for researchers. Researchers hold "jam sessions" to share best practices in study design and become more familiar with the array of OHDSI community assets such as validated phenotypes and the HADES methods library. These activities are lowering the barrier to entry for non-OHDSI researchers to understand community jargon and begin to run their own independent research investigations. Several other abstracts submitted to the OHDSI Global Symposium were supported by the OHDSI Lab 2.0 infrastructure.

Conclusion

This is a perfect case study in how a valuable a shared platform can be for independent researchers, streamlining their research efforts and maximizing their productivity. By leveraging this modern, scalable architecture, researchers can tap into a comprehensive ecosystem of resources and tools, while minimizing administrative and IT requirements. The OHDSI Lab 2.0 exemplifies the benefits of shared services in accelerating scientific discovery and advancing healthcare research. There is significant potential to increase the use of OHDSI methods and tools adoption and get these powerful tools into more researchers worldwide. This implementation represents a significant step towards accelerating scientific discovery and promoting global advancements in healthcare research.

(N.B. The original Observational Medical Outcomes Partnership (OMOP) experiment had a communal lab that was dubbed the OMOP Lab to which we pay homage to.)

References

1. Observational Health Data Sciences, Informatics. Chapter 8 OHDSI Analytics Tools. 11 Jan 2021 [cited 16 Jun 2023]. Available: https://ohdsi.github.io/TheBookOfOhdsi/OhdsiAnalyticsTools.html#atlas
2. Methot J, Philofsky M, Bush B, Smith E, Smith D, Nagy P. How health systems can create value by adopting the OMOP CDM. [cited 16 Jun 2023]. Available: https://www.ohdsi.org/wp-content/uploads/2022/10/78-John-Methot_Why_Healthcare_Systems_Adopt_the_OMOP_CDM_2022symposium-John-M.pdf

Figure 1. OHDSI Lab 2.0 Components

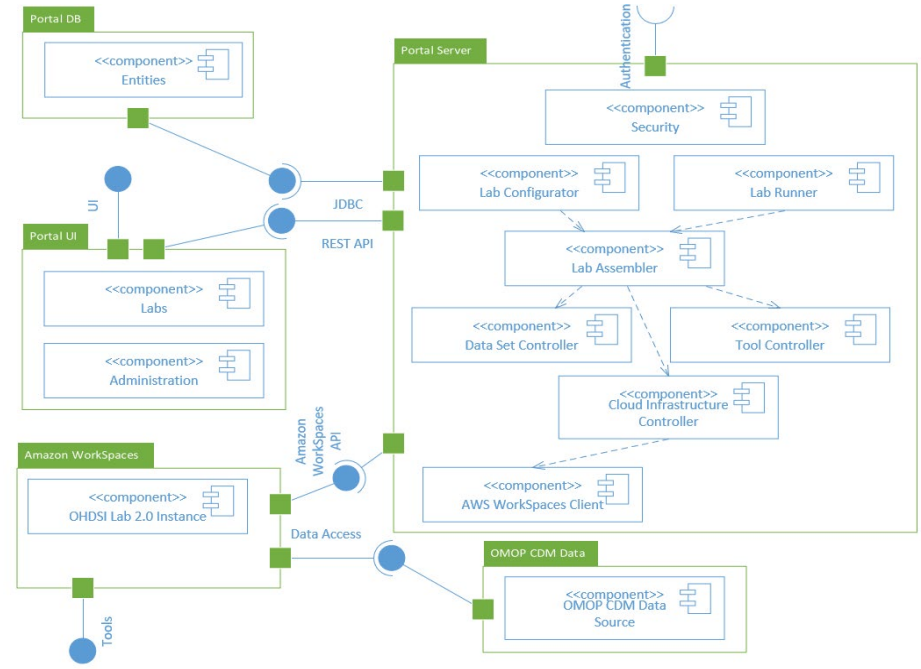

Figure 2. Logical Components within an Amazon Cloud Deployment

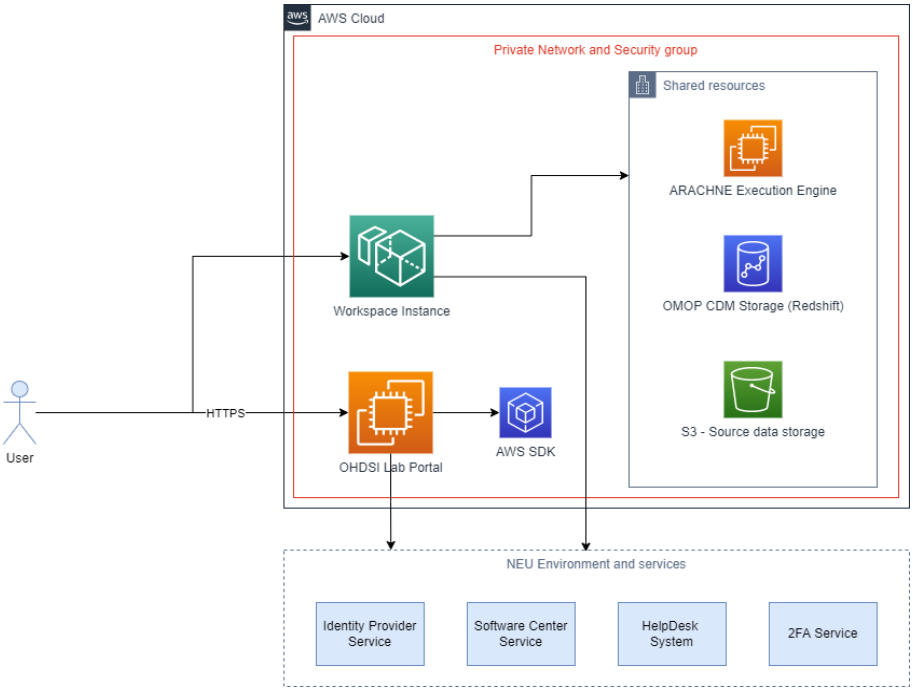

Figure 3. User Interface to Initiate OHDSI Lab Instantiation