ExCITE: A Containerized Open-Source Platform Integrating EHR, FHIR, and OMOP for Biomedical Informatics Education

Robert B Barrett[a], Hayden Spence, MS[a], Benjamin Martin, PhD[a], Teri Sippel Schmidt, MS[a], Saptarshi Purkayastha, PhD[b], Paul Nagy, PhD[a]

[a]Biomedical Informatics and Data Science, Johns Hopkins University, Baltimore, MD, USA

[b]School of Informatics and Computing, Indiana University Purdue University, Indianapolis, IN, USA

Background

Biomedical informatics education increasingly relies on realistic tools and data to train students in health IT and data analytics. Various open-source resources exist to support this training, including electronic health record (EHR) systems like OpenEMR, LibreHealth EHR, and OpenMRS [1-3], synthetic data generators such as Synthea [4], interoperability sandboxes like the SMART on FHIR platform [5], and open clinical datasets exemplified by MIMIC-III/IV [6]. Moreover, the OHDSI initiative provides standardized analytics tools (e.g. ATLAS for OMOP common data model databases) to generate real-world evidence from patient data [7]. However, these educational tools are typically siloed, and there is a notable gap in integrating an EHR interface with standards-based data exchange and analytics for teaching purposes. Existing "educational EHR" solutions have primarily focused on EHR usage training and lack support for data interoperability or population-level analytics [2, 3, 8]. This gap limits hands-on learning of end-to-end workflows - from entering clinical data in an EHR, through exchange via HL7 FHIR, to analysis in an OMOP-format data warehouse.

Methods

ExCITE (Educators Clinical Informatics Training Environment) was developed as a cohesive, containerized platform to bridge these gaps. The platform combines four main components:

1. OpenEMR, a widely used open-source EHR system [1], serving as the clinical front-end
2. HAPI FHIR, to supplement OpenEMR available FHIR APIs, for standards-based data exchange
3. OHDSI Broadsea, which deploys the OMOP Common Data Model database with the WebAPI and ATLAS web interface [7] for observational analytics
4. The Informatics Classroom environment for interactive notebooks and exercises, supported through NoSQL data storage and JupyterLab

All components were configured as Docker containers orchestrated via a common network, enabling seamless communication. The entire architecture can be deployed locally or within a cloud virtual environment, simplifying installation for classroom settings. We documented the system design, selected open standards (FHIR R4, OMOP CDM v5.3), and chose an open-source licensing model (Apache 2.0) to facilitate community adoption and extension.

Results

We achieved a successful deployment of ExCITE across several institutions, demonstrating the portability and utility of this platform for educational informatics. De-identified patient data from the MIMIC-IV dataset was loaded into the OpenEMR database, using a custom ETL developed

by partners. Patient demographics, diagnoses, medications, and observations were then automatically exported via the FHIR server and transformed into the OMOP CDM tables. The resulting standardized database was then accessible through Broadsea utilities, allowing for usage within Atlas and other analytical packages provided under the OHDSI community. In the Informatics Classroom environment, students could programmatically query a FHIR endpoint, reducing the technical burden involved in understanding programmatic methods for doing so. Initial user testing within classroom settings has been positive: learners were able to navigate from viewing a patient's record in the EHR interface, through executing FHIR-based data queries, to understanding and visualizing clinical workflows. These results illustrate the feasibility of a containerized, integrated informatics education platform. The ExCITE system is available as a public GitHub repository [9].

Conclusion

ExCITE addresses a critical unmet need in biomedical informatics training by providing a unified platform that combines an EHR interface, interoperable data exchange, and analytics on a common data model. This all-in-one environment enables students and trainees to gain hands-on experience with a breadth of health data standards and tools - from entering clinical data and using FHIR to harmonize data, to applying OHDSI/OMOP analytical methods. The platform aims to engage the OHDSI community as it cultivates proficiency in standards-based data integration and observational research among new researchers. By open sourcing ExCITE under an Apache license and containerizing its components, we encourage community contributions and easy deployment at academic institutions. Future work will involve expanding the library of synthetic scenarios, evaluating the platform in additional classroom settings, and refining the FHIR-to-OMOP ETL mappings. ExCITE has the potential to enhance training programs in biomedical informatics by immersing learners in a realistic, standards-aligned ecosystem, thereby better preparing the next generation of practitioners and researchers to work with interoperable health data and open-source analytical tools.

Acknowledgments

This research was supported by the NLM, part of the NIH, under award number T15LM013979]

References

1. OpenEMR Project. *OpenEMR – Open Source Electronic Health Record*. OpenEMR Foundation; 2023. [Internet]. Available from: https://open-emr.org
2. Kokkalakonda, Madhuri, et al. "Evaluating Academic Performance in Health Informatics Post-EHR Implementation." *2024 IEEE 12th International Conference on Healthcare Informatics (ICHI)*. IEEE, 2024.
3. Mohammed-Rajput NA, Smith DC, Mamlin B, Biondich P, Doebbeling BN; Open MRS Collaborative Investigators. OpenMRS, a global medical records system collaborative: factors influencing successful implementation. *AMIA Annu Symp Proc*. 2011;2011:960-968.

4. Walonoski J, Kramer M, Nichols J, et al. Synthea: An approach, method, and software mechanism for generating synthetic patients and the synthetic electronic health care record. *J Am Med Inform Assoc*. 2018;25(3):230–238.
5. Mandel JC, Kreda DA, Mandl KD, Kohane IS, Ramoni RB. SMART on FHIR: a standards-based, interoperable apps platform for electronic health records. *J Am Med Inform Assoc*. 2016;23(5):899–908.
6. Johnson AE, Pollard TJ, Shen L, et al. MIMIC-III, a freely accessible critical care database. *Sci Data*. 2016;3:160035.
7. Observational Health Data Sciences and Informatics (OHDSI). ATLAS (Version 2.7.3) [Internet]. OHDSI; 2019 [cited 2025 Jun 17]. Available from: https://github.com/OHDSI/Atlas
8. Welcher, Catherine M., et al. "Barriers to medical students' electronic health record access can impede their preparedness for practice." *Academic Medicine* 93.1 (2018): 48-53.
9. Nagy P. *ExCITE Platform* [Internet]. GitHub repository; 2025 [cited 2025 Jun 17]. Available from: https://github.com/paulnagy/ExCITE