

## Accepted Manuscript

Automated Population of an i2b2 Clinical Data Warehouse from an openEHR-based Data Repository

Birger Haarbrandt, Erik Tute, Michael Marschollek

PII: S1532-0464(16)30084-3  
DOI: <http://dx.doi.org/10.1016/j.jbi.2016.08.007>  
Reference: YJBIN 2615

To appear in: *Journal of Biomedical Informatics*

Received Date: 21 January 2016  
Revised Date: 24 July 2016  
Accepted Date: 5 August 2016

Please cite this article as: Haarbrandt, B., Tute, E., Marschollek, M., Automated Population of an i2b2 Clinical Data Warehouse from an openEHR-based Data Repository, *Journal of Biomedical Informatics* (2016), doi: <http://dx.doi.org/10.1016/j.jbi.2016.08.007>

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.



# Automated Population of an i2b2 Clinical Data Warehouse from an openEHR-based Data Repository

Birger Haarbrandt\*, Erik Tute, Michael Marschollek

Peter L. Reichertz Institute for Medical Informatics, University of Braunschweig – Institute of  
Technology and Hanover Medical School, Hanover, Germany

## Summary

*Background:* Detailed Clinical Model (DCM) approaches have recently seen wider adoption. More specifically, openEHR-based application systems are now used in production in several countries, serving diverse fields of applications such as health information exchange, clinical registries and electronic medical record systems. However, approaches to efficiently provide openEHR data to researchers for secondary use have not yet been investigated or established.

*Methods:* We developed an approach to automatically load openEHR data instances into the open source clinical data warehouse i2b2. We evaluated query capabilities and the performance of this approach in the context of the Hanover Medical School Translational Research Framework (HaMSTR), an openEHR-based data repository.

*Results:* Automated creation of i2b2 ontologies from archetypes and templates and the integration of openEHR data instances from 903 patients of a paediatric intensive care unit has been achieved. In total, it took an average of ~2527 seconds to create 2.311.624 facts from 141.917 XML documents. Using the imported data, we conducted sample queries to compare the performance with two openEHR systems and to investigate if this representation of data is feasible to support cohort identification and record level data extraction.

*Discussion:* We found the automated population of an i2b2 clinical data warehouse to be a feasible approach to make openEHR data instances available for secondary use. Such an approach can facilitate timely provision of clinical data to researchers. It complements analytics based on the Archetype Query Language by allowing querying on both, legacy clinical data sources and openEHR data instances at the same time and by providing an easy-to-use query interface. However, due to different levels of expressiveness in the data models, not all semantics could be preserved during the ETL process.

---

\*Corresponding author

Download English Version:

<https://daneshyari.com/en/article/6927671>

Download Persian Version:

<https://daneshyari.com/article/6927671>

[Daneshyari.com](https://daneshyari.com)