Session 2: ETL Discussion at AMIA

i2b2 Community Discussion at AMIA: i2b2 and ETL Processes

Date: November 14, 2016

Presenters:

Diane Keogh, Managing Director, i2b2 Foundation

Presentation: i2b2 Foundation Update

James R. Campbell, MD, University of Nebraska

Presentation: Enhancing interoperation: an i2b2 ETL schema for Epic EHRs

Andrew Post, MD, PhD, Atlanta Clinical & Translational Science Institute (ACTSI)

Presentation: i2b2 ETL - Eureka Clinical Analytics Project at Emory University

Russ Waitman, PhD, University of Kansas Medical Center

Presentation: HERON ETL for i2b2 and Discussion on how we might support reproducible science as a community

Mike Mendis, Partners HealthCare – i2b2

Presentation: ETL with Business Intelligence Development Studio (BIDS)

Session Summary

The second i2b2 Educational Seminar was held on November 14, 2016 at the AMIA Conference in Chicago, IL. The main topic of this educational seminar was a community discussion around i2b2 and ETL processes. Several community members presented information about their experience using ETL tools with their EMR and i2b2. These discussions evolved into how the community can share the ETL scripts with the rest of the community.

Presentations (5)

i2b2 Foundation Update

Presenter: Diane Keogh, Managing Director, i2b2 Foundation

Discussion Highlights:

- Accomplishments to date
- Source code repositories
- Sharing i2b2 plugins
- Community Support Efforts – Educational webinar series

- Exploring partnership between i2b2 and tranSMART foundations

Presentation Slides



Enhancing interoperation: an i2b2 ETL schema for Epic EHRs

Presenter: James R. Campbell, MD, University of Nebraska

Discussion Highlights:

- ETL Library for Epic EHRs
- Organizing i2b2 for interoperation
- i2b2 ETL architecture for Epic EHRs
- Data warehouse extracts vs CCDA vs FHIR interface for transportability of code

Presentation Video

Presentation Slides



---

i2b2 ETL - Eureka Clinical Analytics Project at Emory University

Presenter: Andrew Post, MD, PhD, Atlanta Clinical & Translational Science Institute

Discussion Highlights:

- ETL tools used to load from central warehouse into i2b2
- EMR is Cerner
- Implemented incremental loading

Presentation Video

Presentation Slides



HERON ETL for i2b2 and Discussion on how we might support reproducible science as a community

Presenter: Russ Waitman, PhD, University of Kansas Medical Center

Discussion Highlights:

- HERON ETL Data Flow
- Jenkins used to automate the ETL process
- Data Sources:
    - REDCap
    - Clarity
    - IDX
    - University Health system consortium data

Presentation Video

Presentation Slides



Additional Resources

KUMC maintains overview of our HERON ETL process here:

https://informatics.kumc.edu/work/wiki/HERON

Links to an older code repository:

https://informatics.kumc.edu/work/browser/heron_load

Through our work with the GPC we are now maintain our current code repository on Github

https://github.com/kumc-bmi

Access to Repository

The actual HERON ETL is in a private project https://github.com/kumc-bmi/heron. Once you have a GitHub account, contact **Dan Connolly** and Heron-Admin@kumc.edu to get access to the latest code.

ETL with Business Intelligence Development Studio (BIDS)

Presenter: Mike Mendis, Partners HealthCare – i2b2

Discussion Highlights:

- BIDS is Visual Studio development environment
- Microsoft offers SQL Server and Business Intelligence stack provides
    - SQL Server / T-SQL query engine
    - Integration Services provide ETL tools
    - Analysis Services provides tools to build multidimensional databases…
    - Reporting Services provides tools to build reports and a portal to deploy reports

- i2b2 ETLsMicrosoft Connectors available for Oracle database support
    - Can be used to populate i2b2 data marts
    - Run stored procedure to convert to standard format

Presentation Video

Presentation Slides

