

Reconstruction of the MIMIC-III Database for Data Analytics

By: Joseph Miles

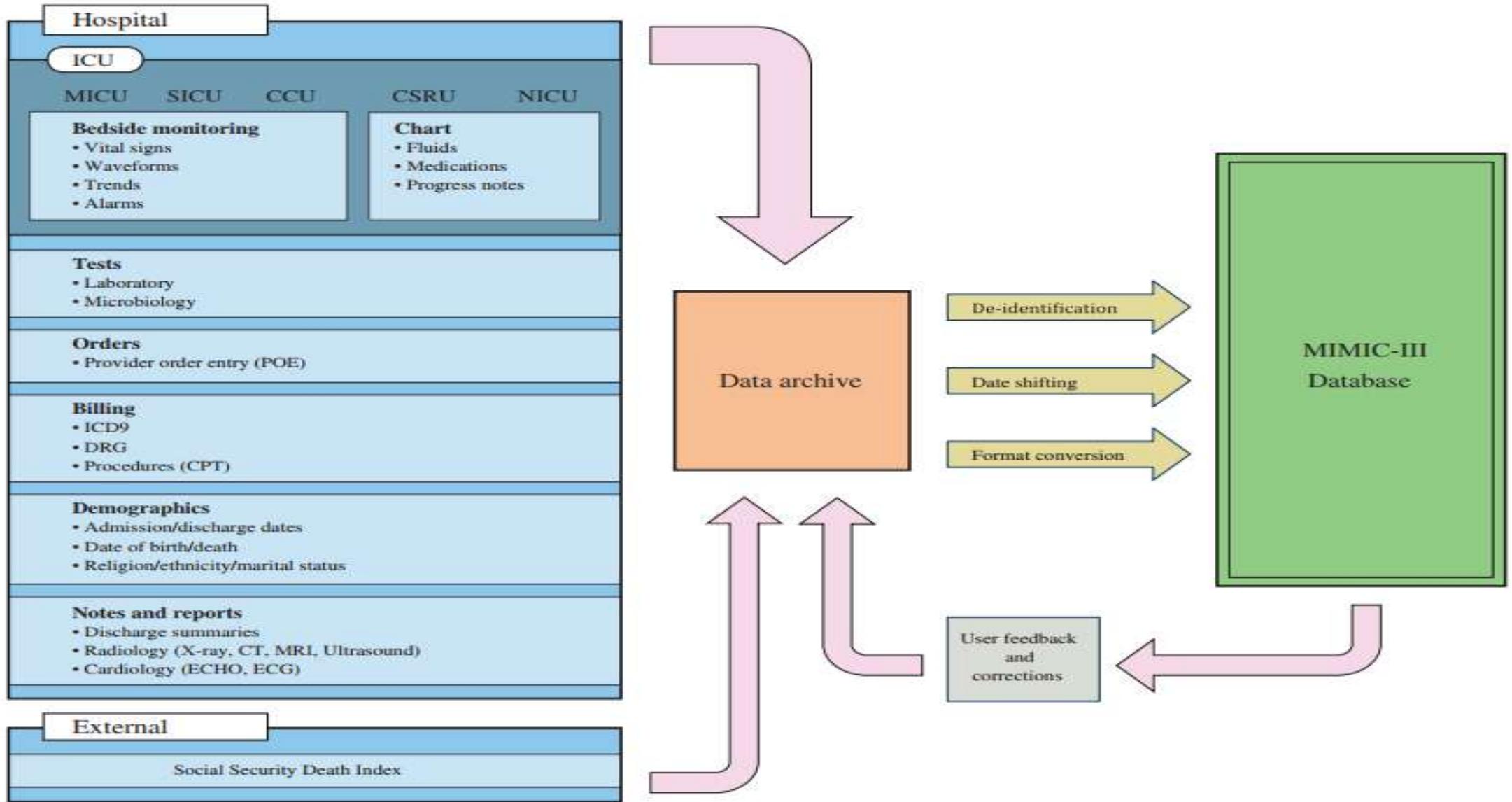
jmiles3@oswego.edu

SUNY Oswego, BHI600IS: Database Applications

Introduction

- ▶ MIMIC-III is a comprehensive collection of deidentified data from 53,423 distinct critical care hospital admissions from 38,597 distinct adult
 - ▶ from Beth Israel Deaconess Medical Center in Boston, Massachusetts
- ▶ 26 tables, 324 attributes
 - ▶ Average of 4579 charted observations and 380 laboratory measurements for each hospital admission
 - ▶ Over 2-million rows of unstructured textual data from various healthcare provider notes and analyses

MIMIC-III overview from Johnson and Associates, 2016



Client/User Background

- ▶ Researcher/Data Analyst
 - ▶ project creation (example: create hypotheses)
 - ▶ project validation (example: compare study data with data in MIMIC-III)
 - ▶ data mining exercises (example: semantic analysis of provider notes).

Project Overview

- ▶ After downloading and analyzing the MIMIC source tables, implementation occurs in 5 additional steps:
 - ▶ 1. Create tables with attribute rules (data types) and identify the primary key for each table.
 - ▶ 2. Load records from csv files into each table.
 - ▶ 3. Declare the indexes for each table.
 - ▶ 4. Define foreign keys in each table and establish table relationships.
 - ▶ 5. Implement user interface (with appropriately granted permissions) for the database.

Conceptual Data Model

- ▶ Initial Data Dictionary
 - ▶ <https://mimic.physionet.org/mimictables/admissions/>
- ▶ Entity Relationship Class Diagram:
 - ▶ <http://pi.cs.oswego.edu/~jmiles3/mimic/assets/MIMIC-ER-DIAGRAM.jpg>
- ▶ Data Dictionary
 - ▶ <http://pi.cs.oswego.edu/~jmiles3/mimic/DataDictionary/>
- ▶ Use:
 - ▶ Researcher/Data Analyst
 - ▶ Ad hoc SQL interface
 - ▶ Implemented as “read only”

Database Interface

MIMIC-III

Graphical Interface for the MIMIC database
by Joseph Miles

2 queries performed; 2 rows returned; elapsed time: 110.16 milliseconds.

Enter MySQL code below.

Multiple lines and queries allowed.
Click "Process" to execute code.

Must use:
SELECT
SHOW or
DESCRIBE

For example:
SHOW tables;

```
SELECT COUNT(DISTINCT SUBJECT_ID) FROM ADMISSIONS;  
SELECT COUNT(*) FROM PATIENTS;
```

Process

Query 1:

```
COUNT(DISTINCT SUBJECT_ID)
```

46520

Query 2:

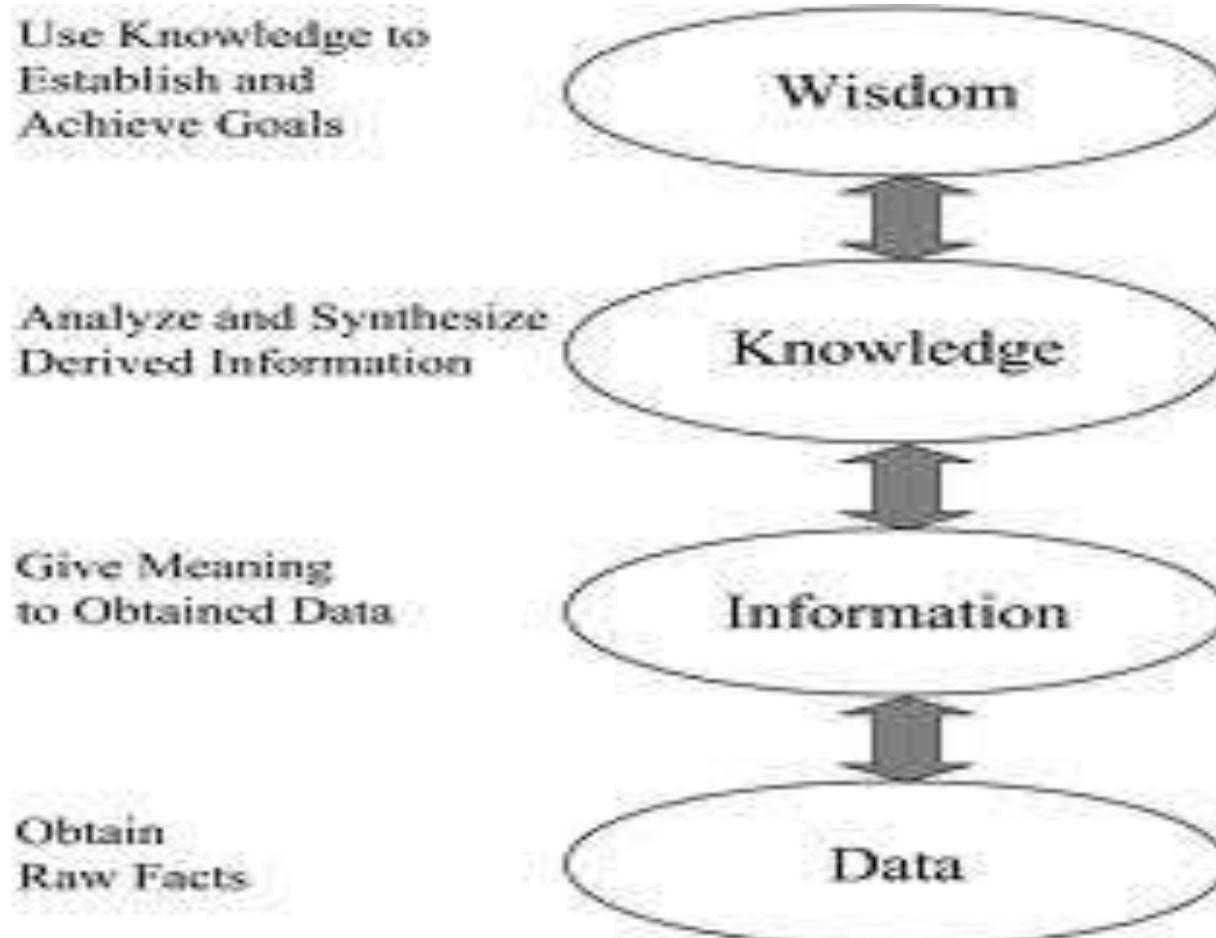
```
COUNT(*)
```

46520

[MIMIC Data Dictionary](#)

[MIMIC Entity Relationship Diagram](#)

Data-Information-Knowledge-Wisdom Continuum (Rhona Carretas, 2012)



Discussion

- ▶ The goal of health informatics could be defined as refining **data** to produce new **information** which can be synthesized into better **knowledge** which develops into **wisdom** utilized for future success
 - ▶ Future success in health informatics will be in producing wisdom for improved patient care
- ▶ Beyond the hope that medical wisdom can be derived from the multiple instantiations of MIMIC databases utilized around the world, hopefully the existence of the MIMIC database and the success it continues to create for those who utilize the data's information will lead other institutions to similarly share and utilize patient data towards the goal of increasing patient-care-wisdom.

References

- ▶ Carretas, R. (2012). Nursing Informatics. Retrieved 8/9/2017 from <http://rhonacarretas.blogspot.com/2012/05/theories-in-nursing-informatics.html>
- ▶ Johnson AEW, Pollard TJ, Shen L, Lehman L, Feng M, Ghassemi M, Moody B, Szolovits P, Celi LA, and Mark RG. (2016). MIMIC-III, a freely accessible critical care database. *Scientific Data* DOI: 10.1038/sdata.2016.35. Available from: <http://www.nature.com/articles/sdata201635>
- ▶ Weinman, B. SQL Essential Training. Lynda.com. “SID” source code retrieved 8/6/2017 from <https://www.lynda.com/SQL-tutorials/SQL-Essential-Training/139988-2.html>