Utah-Michigan ITCR Collaboration

FHIR-based document extraction tool for EHRs

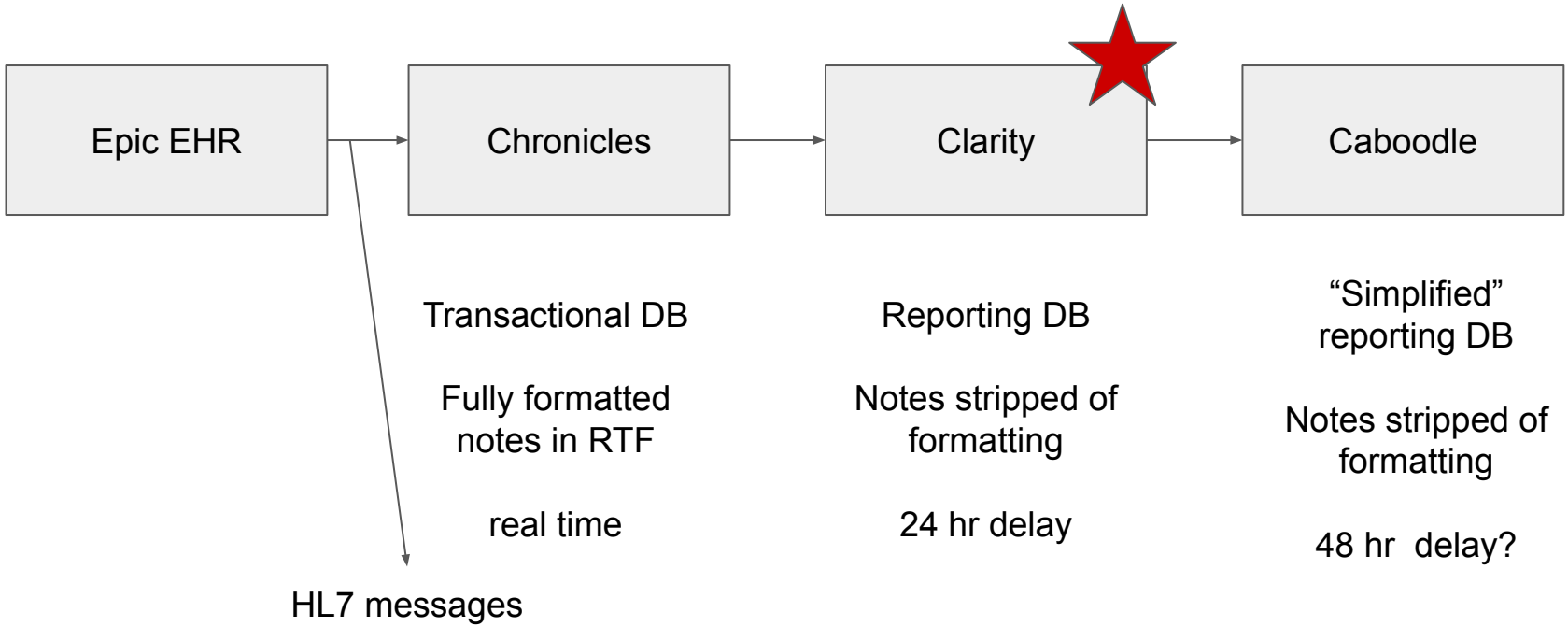
EMERSE

- Search engine designed specifically for EHR data in clinical notes
- ~80% of all EHR data are only in free text
- Groups notes by patient
- Find cohorts
- Supports chart reviews, collaborative work among team members
- Powerful query expansion
- http://project-emerse.org
- Clinical notes are known to be difficult to access, often very locked down
- Many academic institutions have limited capabilities for extracting/utilizing EHR notes

EMERSE used for family history NLP annotation @ Utah

- Goal: develop NLP algorithms to extract family history documentation from clinical notes
- NLP annotation is time consuming and data are sparse
- EMERSE assists annotation process:
  - Identifying most relevant note types
  - Enriching annotation sample
  - Highlighting relevant family history terms in retrieved notes

EHRs (Epic as example)



Note formatting

Eyeglass Final Rx

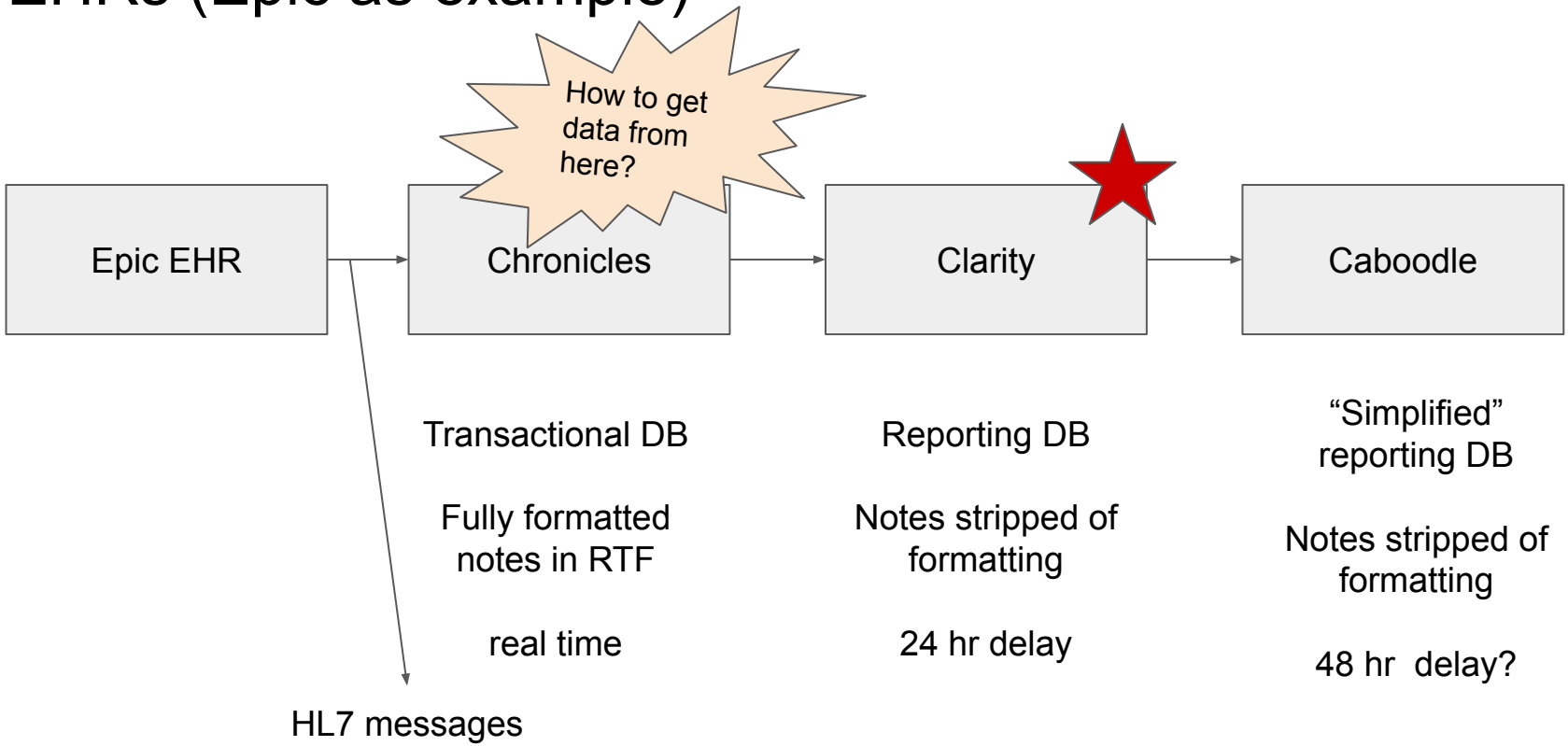
Eyeglass Final Rx

	Sphere	Cylinder	Axis
Right	-1.50	+1.00	165
Left	-2.50	+0.50	28

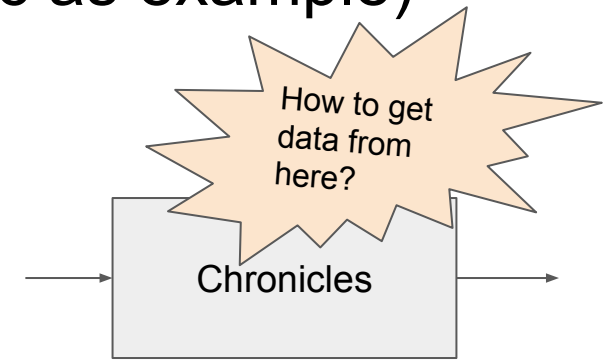
```
Eyeglass Final Rx
Eyeglass Final Rx


Sphere
Cylinder
Axis
Right
-1.50
+1.00
165
Left
-2.50
+0.50
28
```

EHRs (Epic as example)



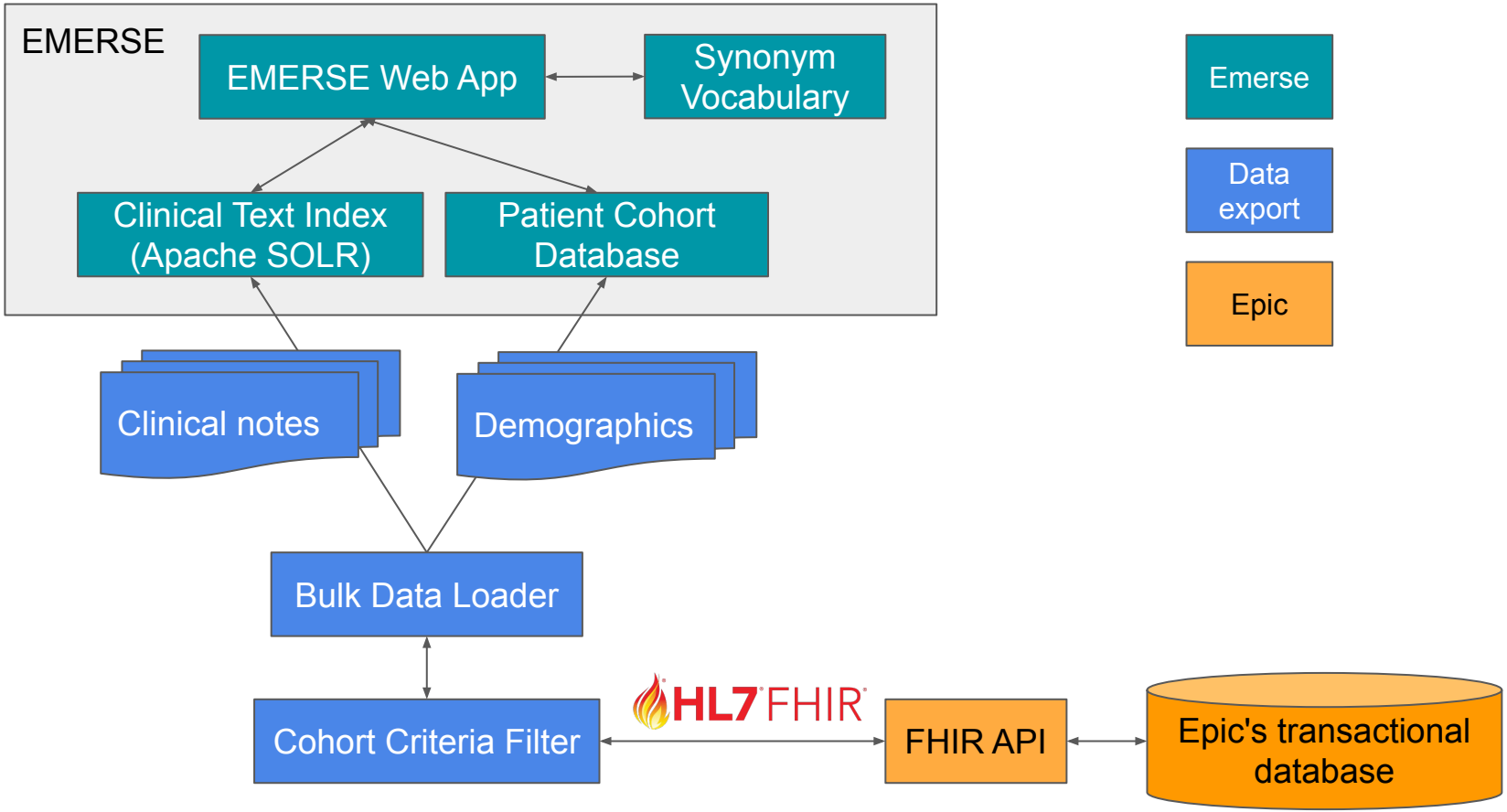
EHRs (Epic as example)



HL7® FHIR®

- Vendor independent standard for interaction with EHRs (mostly data requests)
- Originally envisioned to pull data for a single patient at a time
- For EHR search tool, necessary to “bulk” extract data for many patients
- Utah developed a FHIR-based tool for extracting notes in bulk

FHIR-based clinical note data export tool



What parts of FHIR are used

- Version: STU4
- Resources
  - Patient
  - DocumentReference
  - Binary
- Queries
  - Patient (based on MRN)
  - For each patient, retrieve all DocumentReference resources
- Leveraged EHR's native FHIR API

FHIR experience at Michigan

- FHIR is standard but vendors have flexibility in how and how much of the FHIR spec they implement
- Institutions may also differ in what underlying components are supported/turned on
- Initially had limited environments for testing
  - Had to have FHIR servers enabled for our testing environments
- Could successfully query Chronicles using Utah-developed FHIR API to get fully formatted HTML notes.

FHIR experience at Michigan

- Technical capabilities and portability confirmed
- Benchmarking: 3 documents/second
- Local policies about access not defined
  - How much could we query the primary transaction system? Times of day? Volume/threads?
- Still need additional queries regarding other operational aspects (e.g., deleted/changes notes, split/merged patient IDs, etc)
- Final steps: documenting details, how the tool works, and adding code to GitHub repository