EMERSE Community Meeting

2026-Oct-21

https://project-emerse.org

Plan for today

1. Welcome and Housekeeping
2. Announcements
3. **Guest Speakers: Matthew Manganel, Thomas Love, Xiaobo Fuld**
   (Weill Cornell Medicine)
   “Innovations with ETL: Staging to Production”
4. Software Updates
5. Open Forum
6. Adjourn

Housekeeping

Zoom Meeting

More opportunities for interaction

Please stay muted, unless you would like to ask a question or make a comment

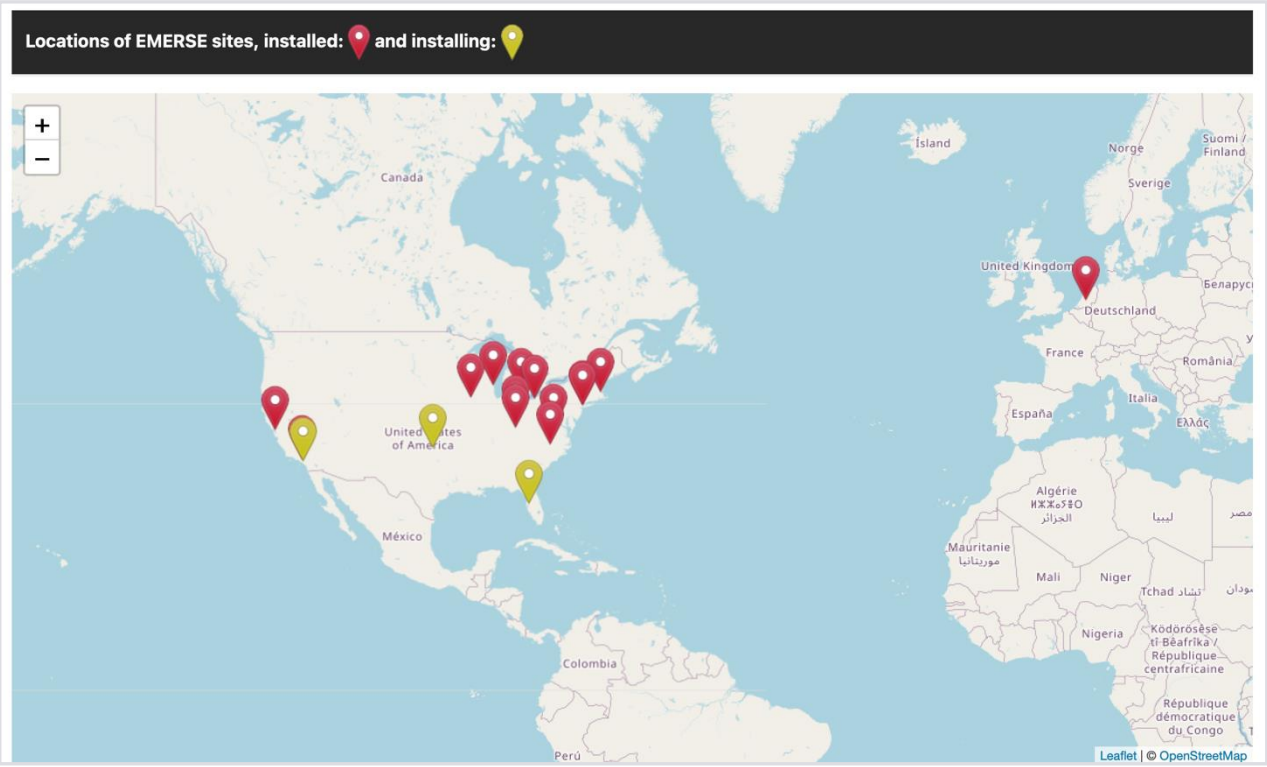
Feel free to use the chat function to type questions or provide comments

We will answer questions throughout

We will record this meeting; it will be available where prior recordings are located:
https://project-emerse.org/presentations.html
https://project-emerse.org/community.html

Community

https://project-emerse.org/community.html



Locations of EMERSE sites

- Installed
- Installing

Acknowledgements

EMERSE has been supported by:

NCI ITCR program

NCATS MICHR CTSA

Michigan Medicine

- Department of Learning Health Sciences
- Office of Research
- Health Information Technology & Services
- Rogel Cancer Center

Publications Using EMERSE



since our last community meeting in May 2026

11

papers

4

abstracts

Full list here: https://project-emerse.org/publications.html



Weill Cornell Medicine

Innovations with ETL: Staging to Production

A robust way to support environments

Matthew Manganel, Thomas Love, Bo Fuld

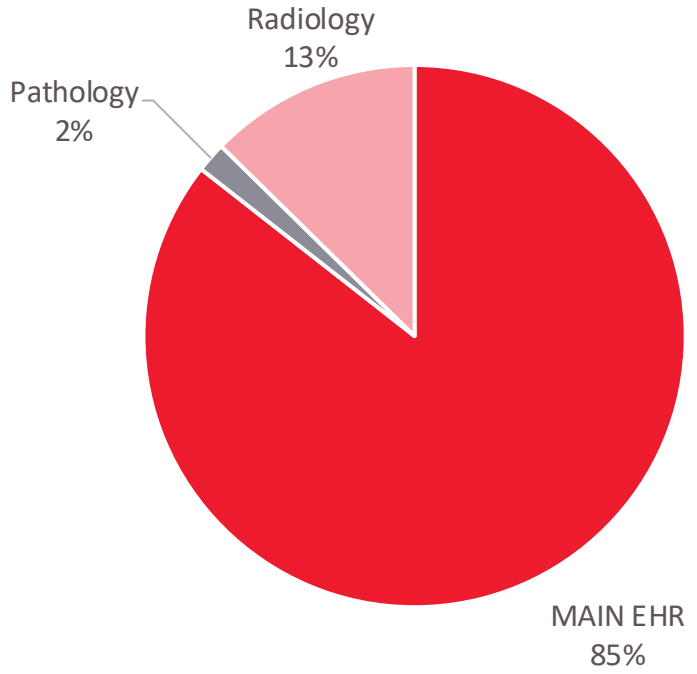
October 21st, 2026

WCM EMERSE

Source	Document Count	Patient Count
Main EHR	226,750,663	3,863,301
Pathology	4,649,856	1,236,999
Radiology	33,611,062	2,795,340
Total	265,011,581	3,997,865 *

*Patients can be present in multiple sources

- Data from 2010 – present
- Currently running EMERSE v6.5.1
- Incrementally refreshed every quarter
- Mirrored across STAGING/PROD instances



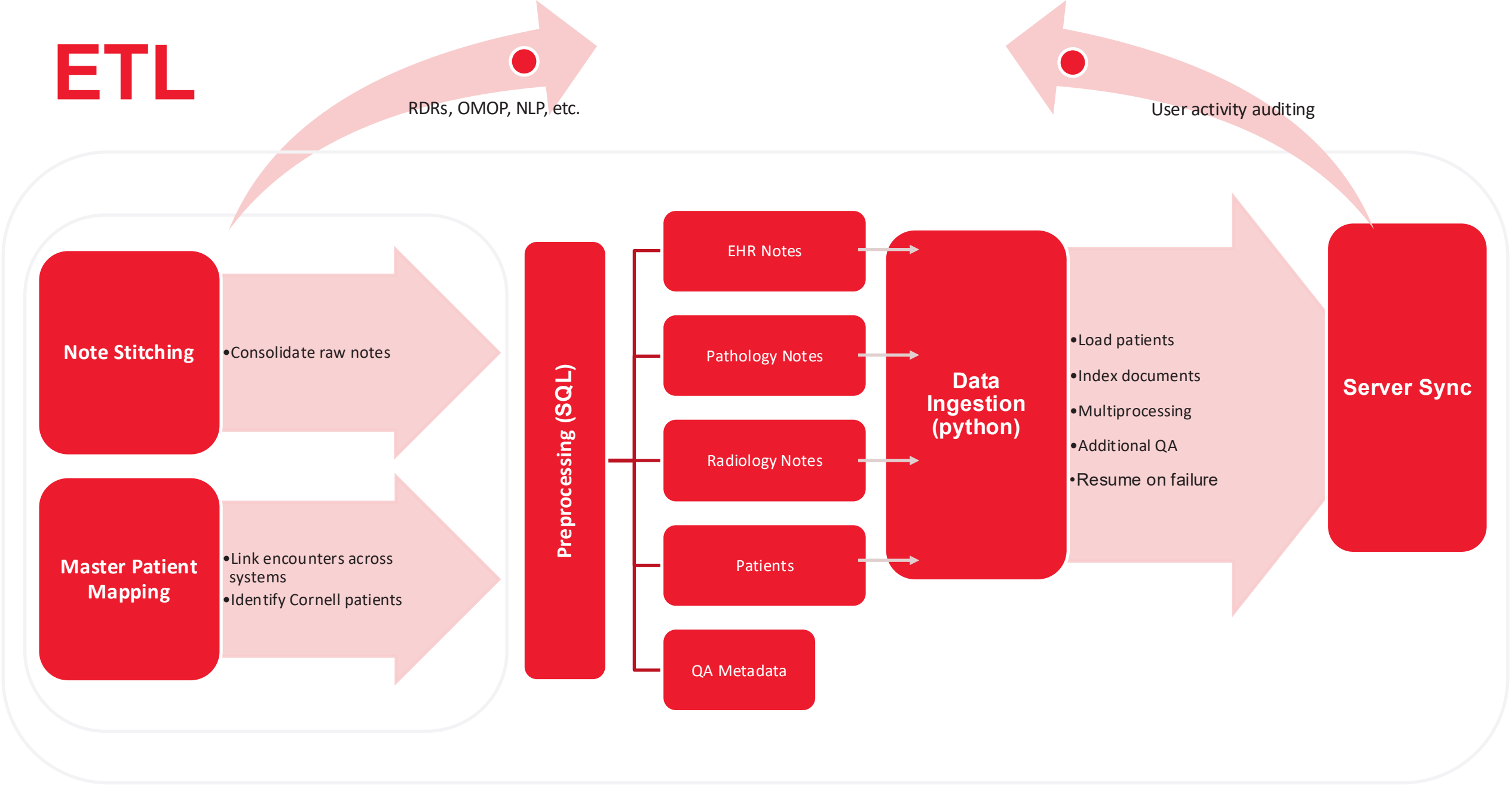


Weill Cornell Medicine

Objectives

- Robust ETL and Infrastructure
- Scalability
- Reliability
- End-to-end Performance

ETL


RDRs, OMOP, NLP, etc.
User activity auditing
Note Stitching
•Consolidate raw notes
Master Patient Mapping
•Link encounters across systems
•Identify Cornell patients
Preprocessing (SQL)
EHR Notes
Pathology Notes
Radiology Notes
Patients
QA Metadata
Data Ingestion (python)
•Load patients
•Index documents
•Multiprocessing
•Additional QA
•Resume on failure
Server Sync



Weill Cornell Medicine

Obstacle

Need to transfer data from Staging to Production

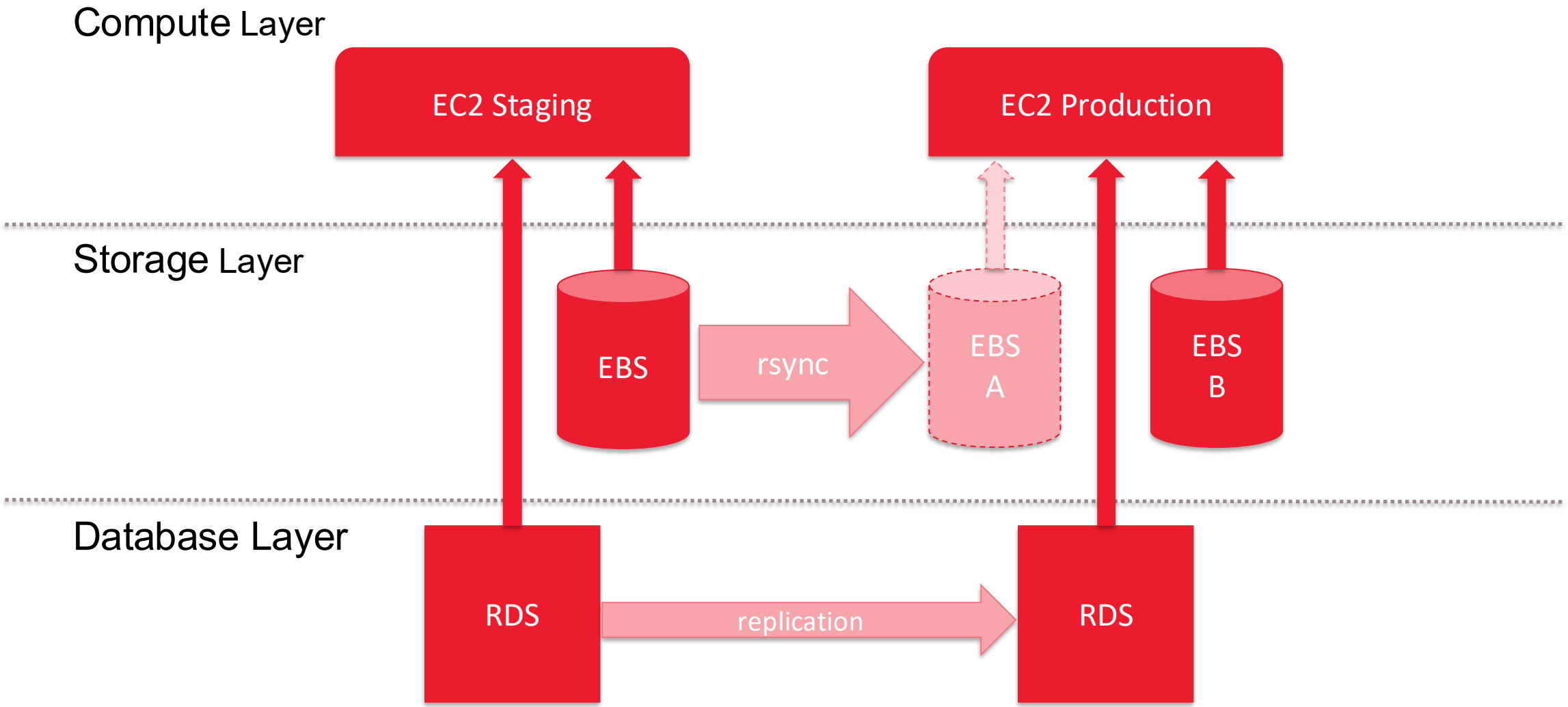


Weill Cornell Medicine

Solution

Infrastructure based on partitioning of Production

Infrastructure Innovations



Benefits

- Data Immutability During Operation
- Reduce Expenses
- Minimize Downtime
- Service Performance
- Easy Rollback
- Disaster Recovery

Weill Cornell Medicine

Future Implications & Improvements

Acknowledgements

- We thank our WCM colleagues for their effort & support:

- Thomas Campion, Tru Tran, Jessie Lee, Tyler Larkworthy, Itzel Nino, Nivedita Chang, Evan Sholle, Carlos Castilla, Omar Rosa

Questions?



Weill Cornell Medicine



Weill Cornell Medicine

EMERSE version 7.2.1

7.2.1

Officially released

- User interface framework updated – improved security
- Timeline
- Customizable table columns in Summaries
- OCR-ready (e.g., thumbnail views, highlighting within documents)
- Documents can be downloaded into zip file
- More compact user interface
- Minor user interface updates (e.g., descriptions no longer required for Saved Patient Lists/Terms)

Brief Demonstration

Note Chat

Work in progress

Note Chat

1. Release at Michigan August 2026 for local testing
2. Seems to be working well
3. Preparing for wider release to other sites

Open Forum

Let’s talk about EMERSE

Ask anything

Share anything

Discuss

Unmute to speak, or type in the Zoom chat

Wrap-up

Thanks for attending!

UPCOMING

Next Meeting

TBD (Spring 2027)

emerse