

# Bringing Data To The People: How A Secure, Self-Service, Free Text Search Tool Can Empower Clinical Research Teams And Improve Productivity



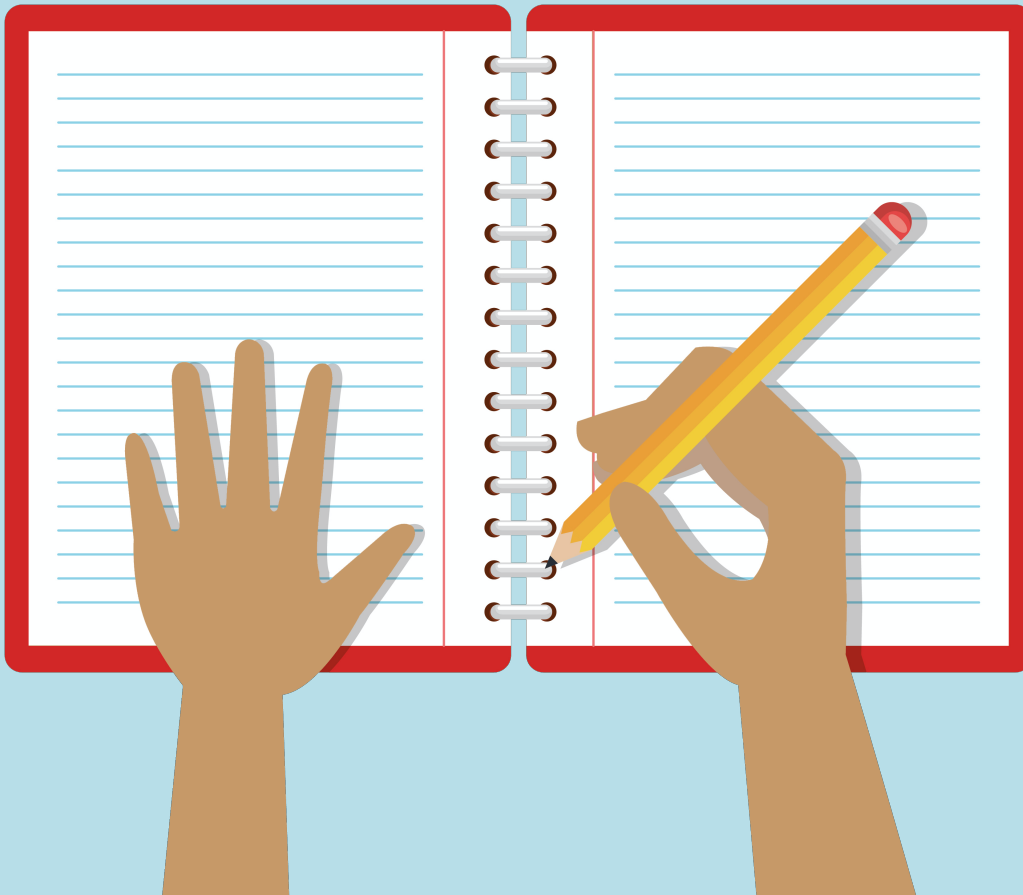
October 19, 2021

David Hanauer, MD, MS  
Dept of Learning Health Sciences  
University of Michigan  
[hanauer@umich.edu](mailto:hanauer@umich.edu)

Twitter: @informaticsGeek  
@projectEMERSE  
Web: project-emerse.org

<http://project-emerse.org/presentations.html>

# If you're thinking of taking notes



these slides can be found at:

link will be on  
most slides



**em@rse**  
ELECTRONIC MEDICAL RECORD SEARCH ENGINE

<http://project-emerse.org/presentations.html>

# Disclosures

Funding: NIH (NCI, NCATS); PCORI

Licenses/Royalties: EMERSE “Synonyms”  
(used for query expansion) which is  
licensed by the U of Michigan



# The Problem(s)

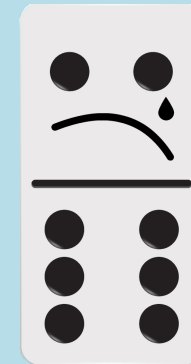
Free text is: complex





# The Problem(s)

Free text is: a mess



“domino pain”

# The Problem(s)

Free text is: difficult to access



# The Problem(s)

Many free text tools are hard to use

Ease of adoption of clinical natural language processing software:  
An evaluation of five systems



Kai Zheng<sup>a,b,\*</sup>, V.G. Vinod Vydiswaran<sup>k</sup>, Yang Liu<sup>b</sup>, Yue Wang<sup>c</sup>, Amber Stubbs<sup>d</sup>, Özlem Uzuner<sup>e</sup>,  
Anupama E. Gururaj<sup>f</sup>, Samuel Bayer<sup>g</sup>, John Aberdeen<sup>g</sup>, Anna Rumshisky<sup>h</sup>, Serguei Pakhomov<sup>i</sup>,  
Hongfang Liu<sup>j</sup>, Hua Xu<sup>f,\*</sup>

"The average ratings provided by the end user evaluators on ease of use and ease of interpreting output... indicat[e] that this group of users generally deemed the systems extremely difficult to work with"

<https://pubmed.ncbi.nlm.nih.gov/26210361/>



<http://project-emerse.org/presentations.html>

# A solution

- A system “for the people”
- Users search the EHR on their own
  - No need to wait in a queue for an analyst or a data scientist
- Data are kept secure within a centralized, audited system
  - No need to download/store the data elsewhere
- Easy-to-use for non-technical researchers



# What is EMERSE?

A search engine, kind of like Google, for free text medical record documents  
(technically, it's an *information retrieval* system)



Simple to use

**em@rse**  
ELECTRONIC MEDICAL RECORD SEARCH ENGINE

<http://project-emerse.org/presentations.html>

# What is EMERSE?

Software we've been developing and refining for 16 years



2005



2021

**em****rse**  
ELECTRONIC MEDICAL RECORD SEARCH ENGINE

<http://project-emerse.org/presentations.html>

# What is EMERSE?

A research tool, available at no cost

- picking the right tool for your research is important; multiple tools are often needed



**em@rse**  
ELECTRONIC MEDICAL RECORD SEARCH ENGINE

<http://project-emerse.org/presentations.html>

# What isn't EMERSE?

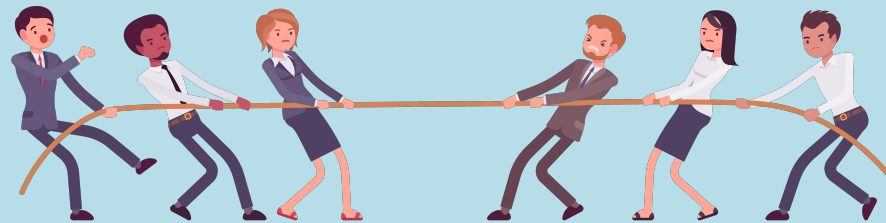
It is not: natural language processing (NLP)

- but full NLP may not always be needed
- and NLP can be hard to use or customize

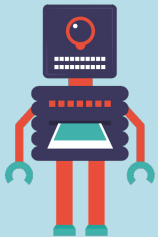




# EMERSE vs NLP



EMERSE: human-in-the-loop *augmented* intelligence

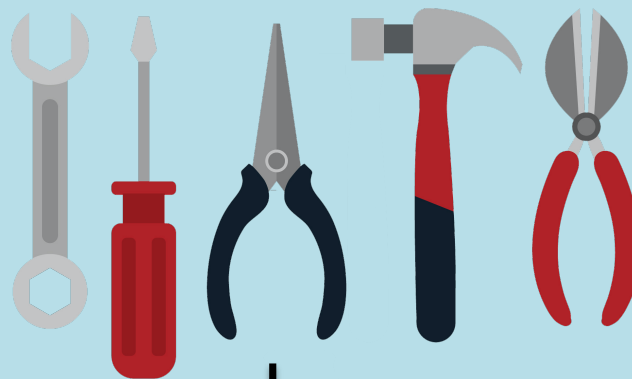


NLP: *artificial* intelligence

**em<sub>ER</sub>se**  
ELECTRONIC MEDICAL RECORD SEARCH ENGINE

<http://project-emerse.org/presentations.html>

# EMERSE vs NLP



EMERSE

NLP

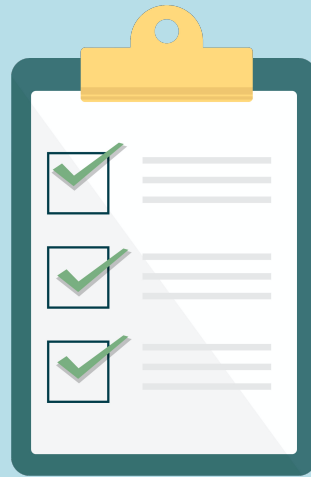
**em@rse**  
ELECTRONIC MEDICAL RECORD SEARCH ENGINE

<http://project-emerse.org/presentations.html>

# Find cohorts

EMERSE allows you to find cohorts based on things mentioned in the notes

- diseases
- drugs
- symptoms
- anything\*

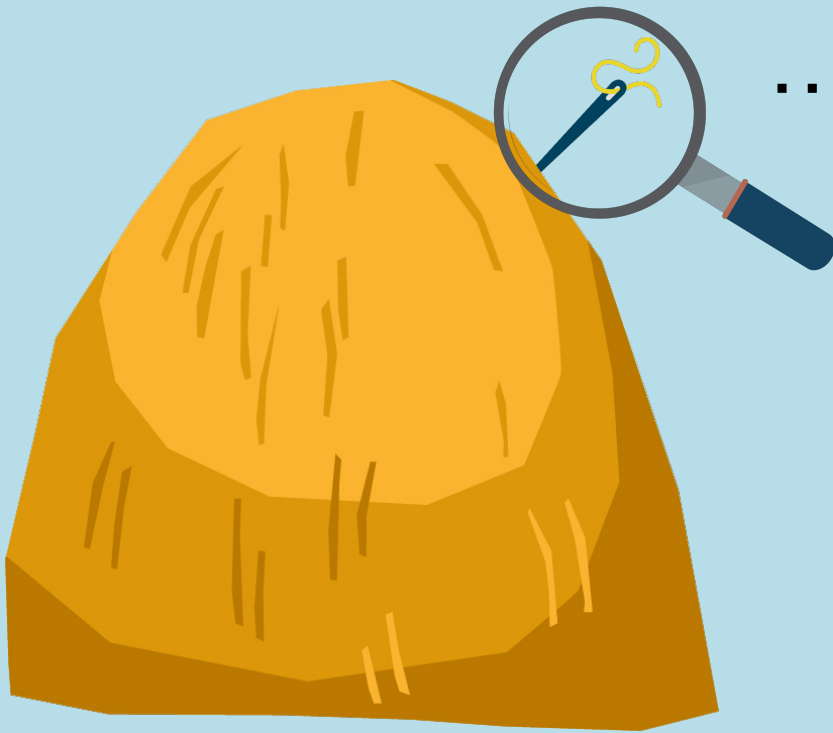


\*if it is mentioned

# Find cohorts

It's perfect for finding rare things...

...like rare cancers



# Highlight documents for chart review

Thoracocentesis confirmed the recurrence of mantle cell lymphoma. Disease restaging work-up revealed multicompartiment lymphadenopathy in the neck, mediastinal, retrocrural, retroperitoneal and pelvic regions. Bone marrow was also involved. The patient was treated with a total of six cycles of rituximab, cyclophosphamide, vincristine, doxorubicin and dexamethasone (R-HyperCVAD) completed in January 2007. That treatment led to complete remission that lasted until October 2008, when the disease was found to have recurred in the left pleural space and retroperitoneum without bone marrow involvement.

<https://jmedicalcasereports.biomedcentral.com/articles/10.1186/1752-1947-4-329>



<http://project-emerse.org/presentations.html>

# EMERSE Bundles

Saved sets of search terms

Shared among users

Helps standardize searches

Supports reproducibility

- Some teams publish Bundles  
in their publication's appendix

(example: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3860174/>)





# EMERSE is *≡fast*

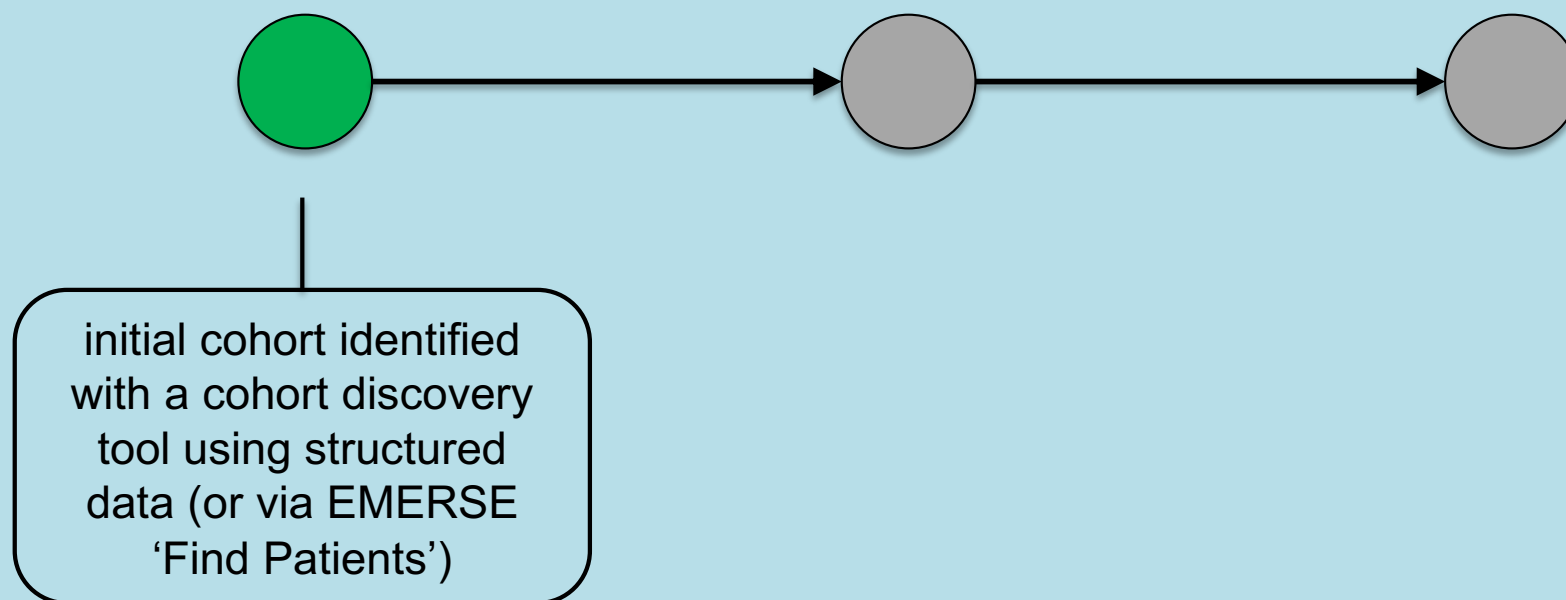
Query to identify all patients with the following	Reporting DB time (s)	EMERSE time (s)	EMERSE advantage
cavernous hemangioma	14,652	2	7,320x
gray platelet syndrome	14,940	2	7,470x
inferior lingular segment of the left upper lobe	17,784	9	1,980x

...enabling real-time querying



<http://project-emerse.org/presentations.html>

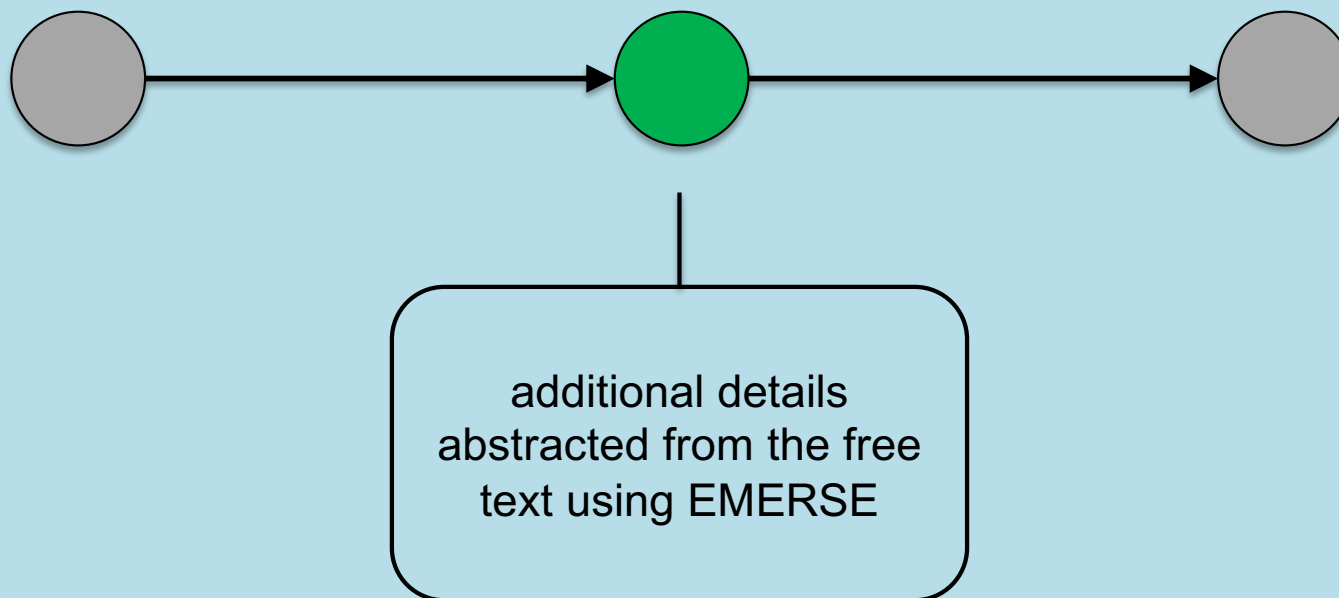
# Typical workflow



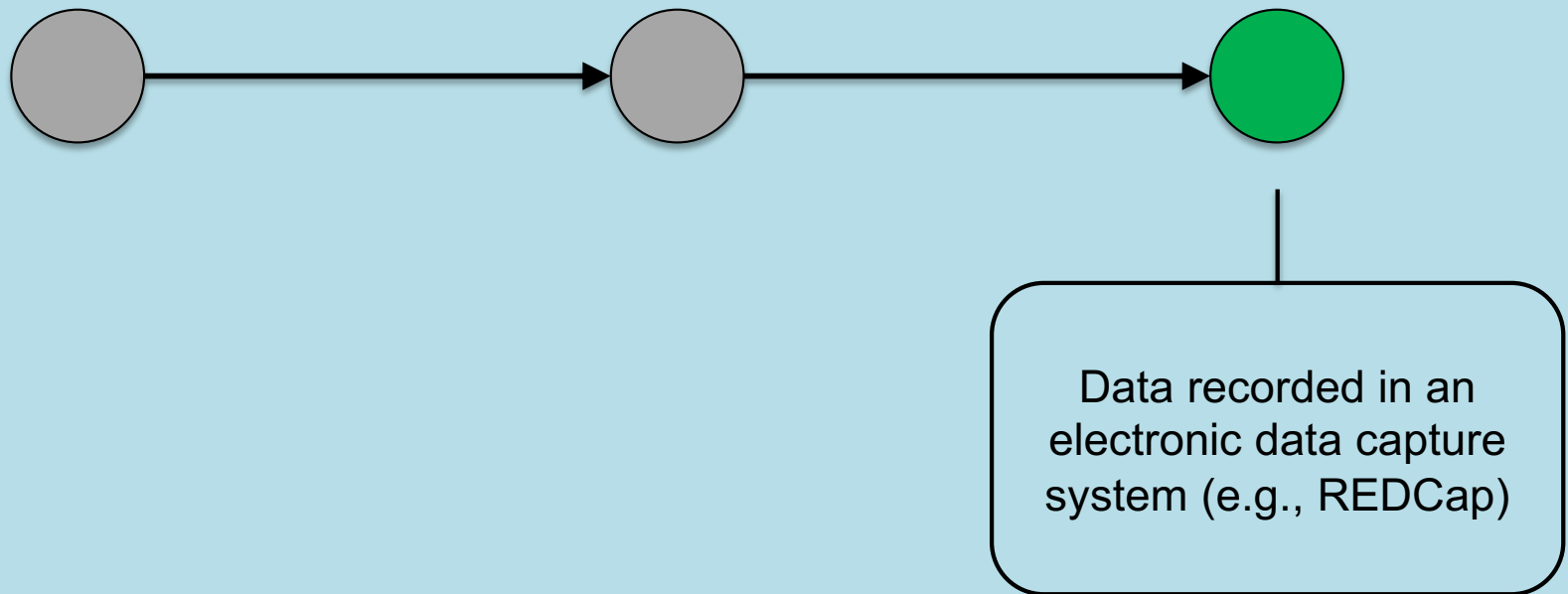
Cohort discovery tools:  
DataDirect @ Michigan  
i2b2, LEAF, others elsewhere



# Typical workflow



# Typical workflow



# Publications using EMERSE

**454** papers

**82** abstracts



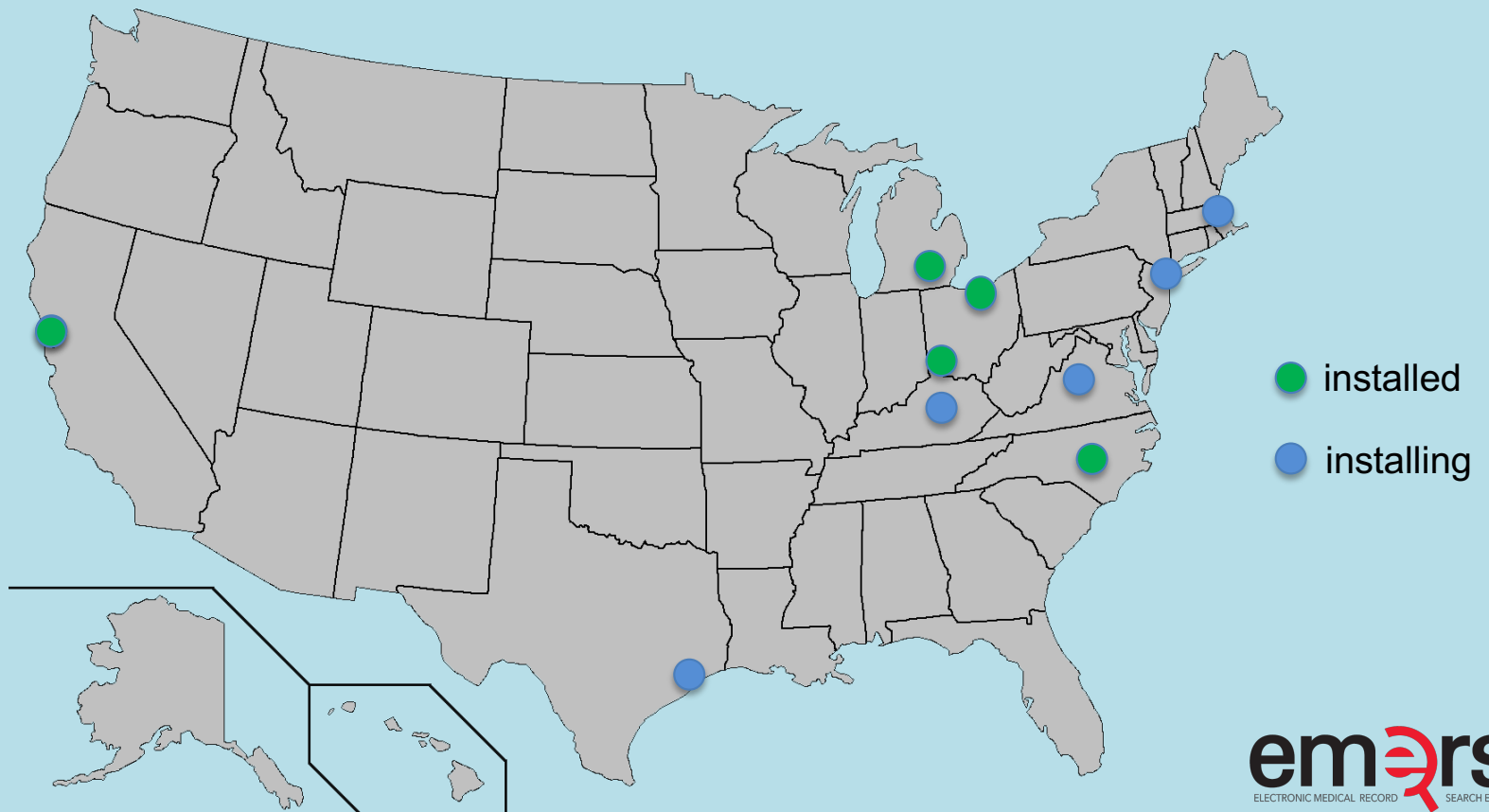
Full list at:

<http://project-emerse.org/publications.html>

**em@rse**  
ELECTRONIC MEDICAL RECORD SEARCH ENGINE

<http://project-emerse.org/presentations.html>

# Where is EMERSE?



# The future...

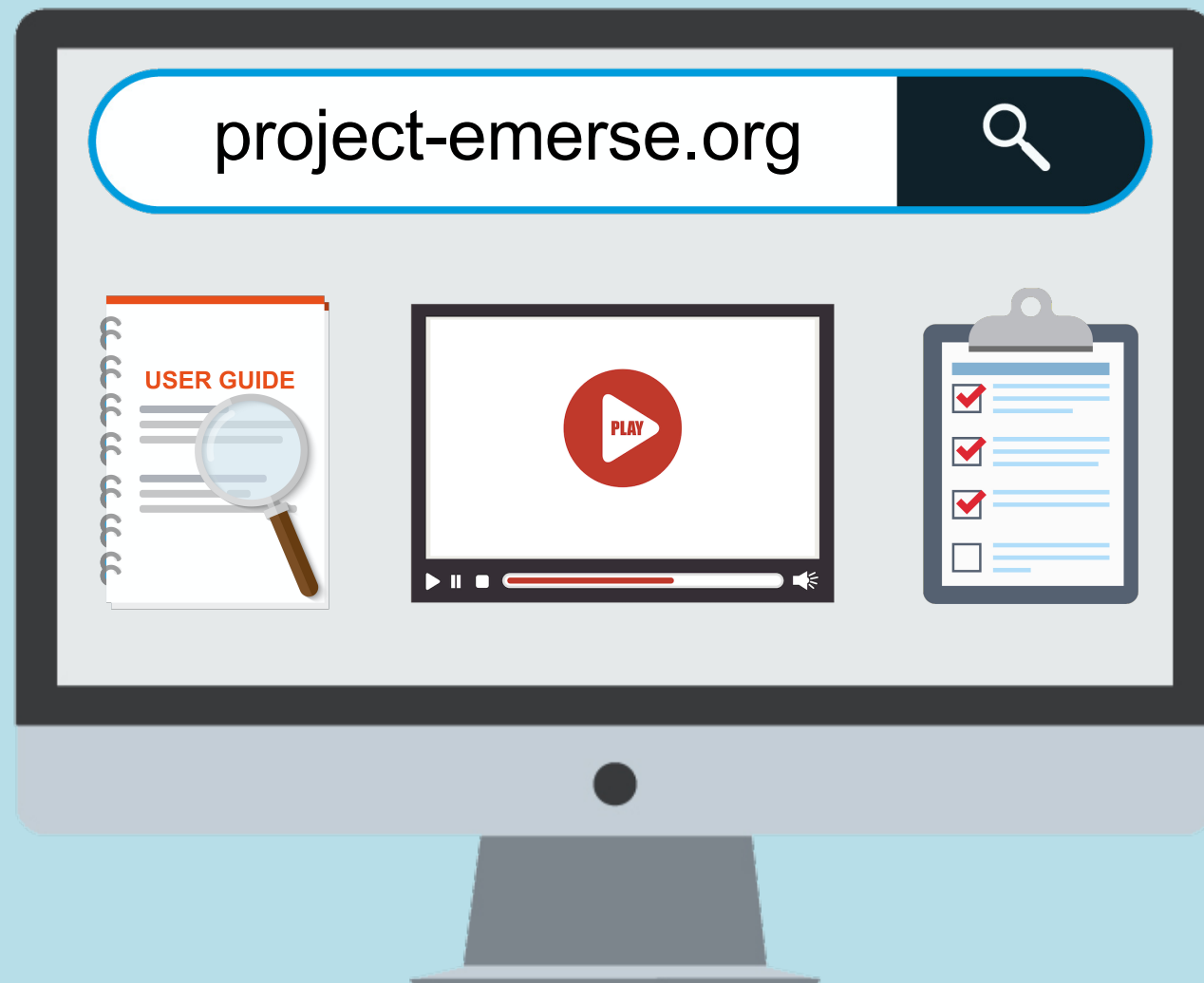
## EMERSE Research Informatics Network

real-time, secure  
cross-site queries



**em****rse**  
ELECTRONIC MEDICAL RECORD SEARCH ENGINE

<http://project-emerse.org/presentations.html>





Twitter: @projectEMERSE

publications  
software releases  
announcements  
webinars

EMERSE-team@umich.edu



Lisa Ferguson  
David Hanauer  
Kellen McClain  
Guan Wang



# Acknowledgements

