

# Unlocking Insights from Clinical Notes with the EMERSE Text Processing Tool

Presentation for NCI/CBIIT

September 30, 2024



Twitter/X: @projectEMERSE

Web: project-emerse.org

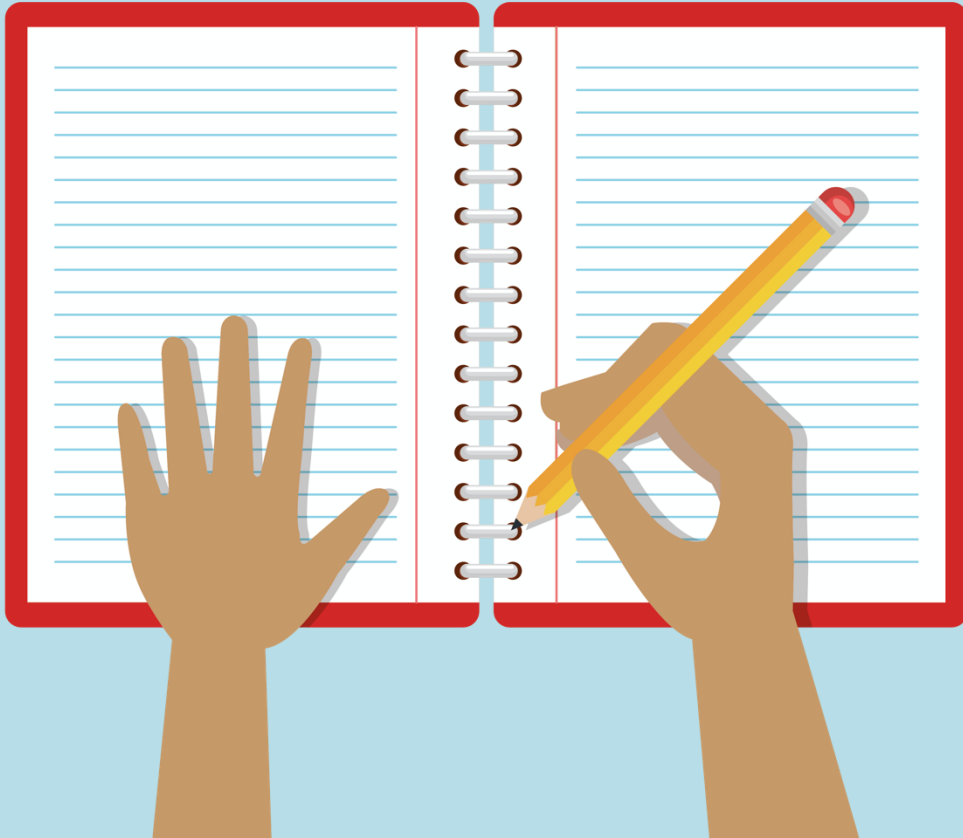
Email: hanauer@umich.edu  
emerse-team@umich.edu

David Hanauer, MD, MS

Dept of Learning Health Sciences

University of Michigan

# If you want to take notes or visit links



these slides can be found at:

this link will be  
on most slides



**em****erse**  
ELECTRONIC MEDICAL RECORD SEARCH ENGINE

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

# Thank You

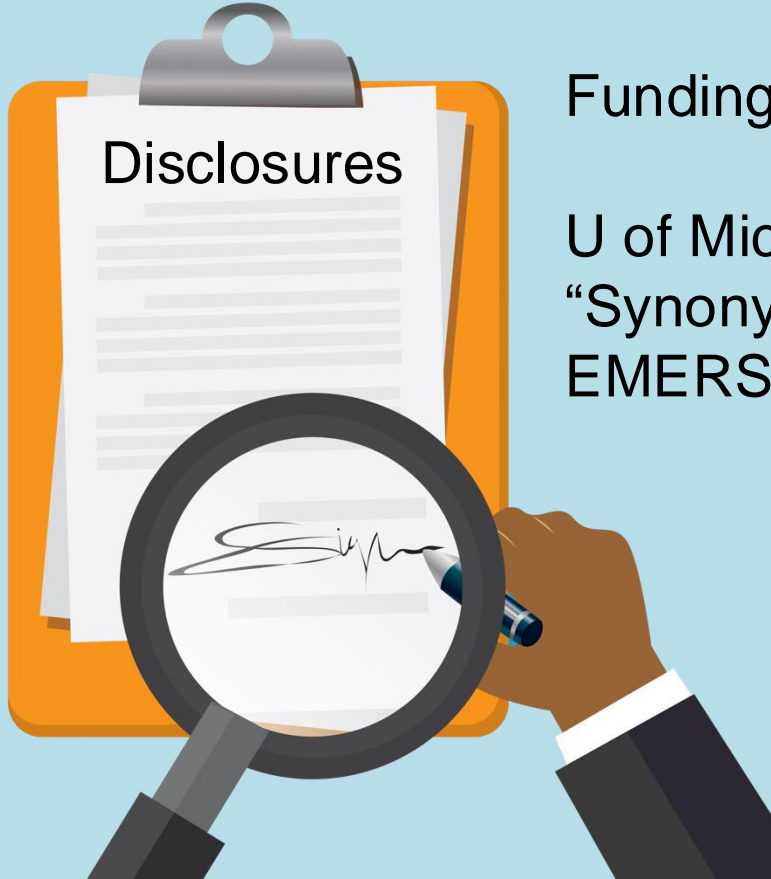
To the NCI ITCR program  
for helping our project grow



# Disclosures

Funding: NIH (NCI-ITCR, NCATS-CTSA)

U of Michigan Royalties/Licensing:  
“Synonyms” dataset—optional “plugin” for  
EMERSE.



A hand holding a megaphone on the left side of the image. The megaphone is white with a grey rim and a red interior. Several white and orange triangular shapes radiate from the megaphone, suggesting sound or energy. The background is black with a red diagonal stripe on the left. A large white banner is positioned diagonally across the center, containing the text "I am here to tell you about EMERSE".

I am here to tell you  
about EMERSE

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

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

**2021 study out of UC Irvine:** *Design, Implementation, and Usability of the Electronic Medical Record Search Engine (EMERSE) Tool*

<https://escholarship.org/uc/item/44p23878>

“Users unanimously responded that they would recommend the system to others, and...for a tool they found so useful, they believed that **far too few people both within and outside of their network knew about the tool’s existence.**”



# Unstructured vs Structured Data

EMERSE is for this...	...not this
<i>Unstructured Data (free-text)</i>	<i>Structured Data</i>
Mrs. Jones is a 56 year old female with a history of HTN, hypercholesterolemia, and T2DM who comes to the clinic today with a 3 day h/o dizziness and severe headache on the left side.	WBC: 5.6 Total cholesterol: 182 Weight: 67.4 AST: 30 ALT: 52

# 80% of EHR data are in unstructured free-text



} structured data

} unstructured data/free text



# Most medical centers lack tools for free-text



# The EMERSE solution



- A system “for the people”
- Users search the notes on their own
  - No need to wait in a queue for an analyst or a data scientist
- Easy-to-use for non-technical users
- Unlike with some EHRs, EMERSE can search across all notes and all patients at once
- Continuous refinements for almost 19 years



# The EMERSE solution



- A system “for the admins”
- Enterprise grade, easy to support
- Configurable with granular roles/privileges to control access (e.g, all patients vs. some patients)
- Data are kept secure within a centralized, audited system
  - No need to download/store the data elsewhere

# EMERSE is a useful tool

- It will be one of many you need to build/support your project





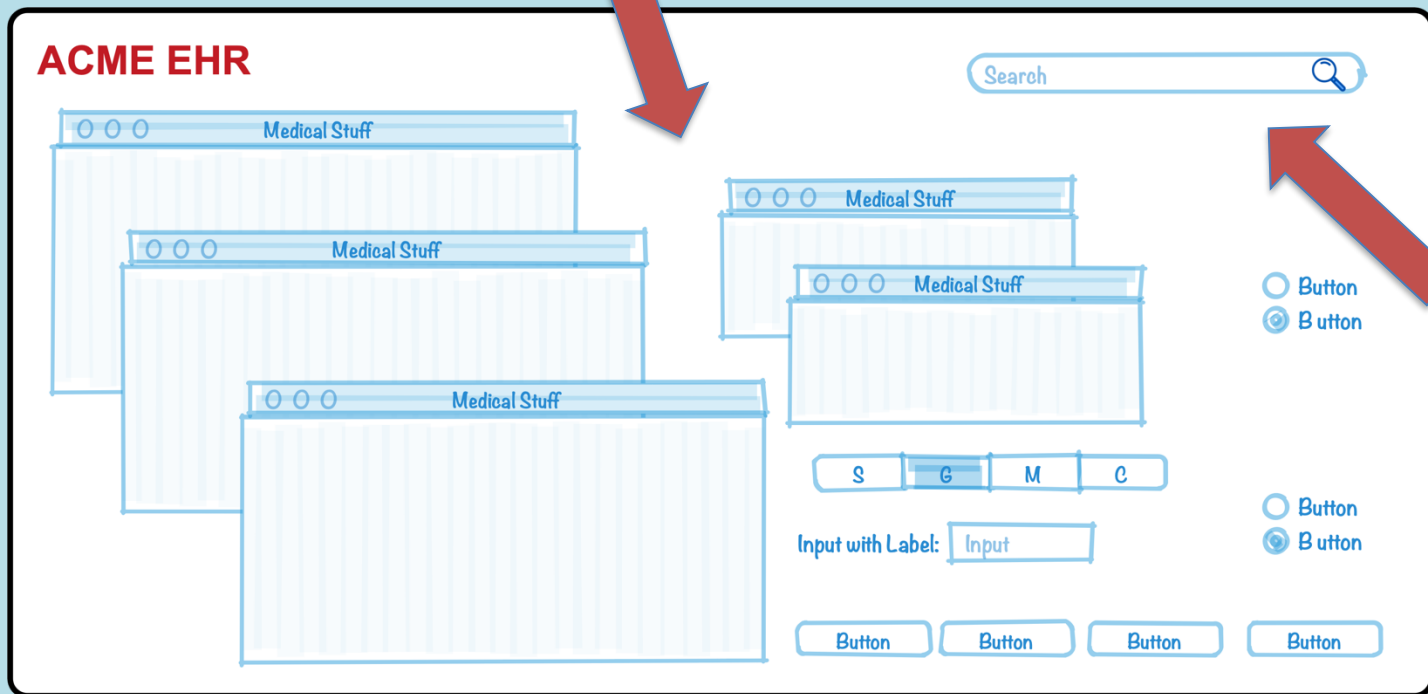
# We think EMERSE is really good

- We're proud of what we created
- We think it's better than other software tools that cost \$\$\$



# How do we know it's better? *We've been laser-focused on it*

*EHRs have a bazillion features*



*The EMERSE team has invested years on just this one feature (search) because it's so important*

# How do we know it's better?

## *We've studied it*

- “new users of the EMERSE system are able to complete basic but critical workflow tasks in the system with a high rate of success...are **highly satisfied with the interface**, and have highly positive perceptions of its expected utility in their work.”
- “survey results indicate **very high ratings of usability and satisfaction** with the EMERSE system by new users after only one session of use.”



# How do we know it's better?

## *Our users tell us*

Thank you for...this important tool which is **proving extremely valuable** in enhancing patient safety and quality of care delivered at Michigan Medicine.

..continued gratitude for the availability of this powerful research tool. As always, **thank you for this innovation!**

EMERSE is **working out great for casefinding**...have found quite a bit of cases that we would have missed otherwise. **Thank you and everyone on your team!**

To say **“it is the most useful tool I use in my job”** would be an understatement, so **thank you** for creating and sharing it with others!!

EMERSE is **an absolute gem** @umichmedicine. The functionality is very friendly and it saved hours of time during the data collection process. **Thank you to the team** that created this powerful research tool!



# How do we know it's better?

## *Researchers mention it in their publications*

- “Reviewers used the EMERSE search tool to **ensure thorough review** of the available documentation...” [PMID 36119396]
- “the **tool avoids the pitfalls of diagnostic inaccuracy** seen with tools querying on ICD and billing codes...” [PMID 36114099]
- “[EMERSE] provides software features to **comprehensively scan all clinical documents**...for keywords and phrases to ensure that even rarely mentioned events are detected.” [PMID 36550198]
- “...information [was] captured via EMERSE...in order to **obtain the most accurate and complete information** per patient.” [PMID 36752520]


# How do we know it's better?

## *Top-tier medical centers trust it*

Installed	Installing
U of Michigan	Weill Cornell Medical Center
Harvard U – Dana Farber Cancer Center	U of Virginia
Columbia U Cancer Center	Moffitt Cancer Center, Tampa, FL
U of North Carolina – Chapel Hill	U of California – Irvine
U of California – San Francisco	Children's Health Orange County, Calif.
U of Kentucky	
U of Cincinnati	
Case Western Reserve U	
U of Minnesota – Minneapolis	
Utrecht University, Netherlands	
U of Iowa	

# What can EMERSE do?

Lots of things!



EMERSE for Cardiac Surgery Research  
David Hanauer

Linda Farhat, Clinical Research Coordinator  
Department of Cardiac Surgery, Michigan Medicine

01:28

**Cardiac Surgery Research**

"It allows us to basically target exactly what we need to know."

--Linda Farhat, Clinical Research Coordinator



EMERSE for Quality Analytics  
David Hanauer

Andrew Heiler, Clinical Quality Coordinator

01:18

**Quality Analytics**

"I can't imagine having to do what I do without having EMERSE."

--Andrew Heiler, Quality Analytics Coordinator



EMERSE for Infection Control  
David Hanauer

Jayna Berger Heiler, Infection Prevention Coach

01:33

**Infection Prevention**

"The value of EMERSE is immeasurable."

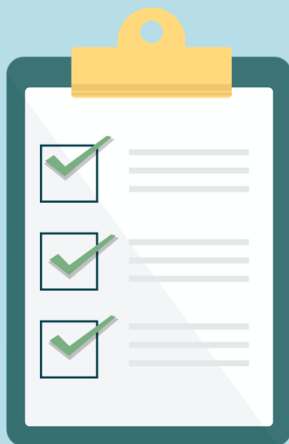
--Jayna Berger Heiler, Infection Prevention Coach

Watch our videos: [https://project-emerse.org/use\\_cases.html](https://project-emerse.org/use_cases.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 such as  
cutaneous leiomyosarcoma

See this talk for more details:

<https://vimeo.com/677482835>

“Using EMERSE to Improve Research  
Involving 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.

# Our philosophy

It's important to view the terms/concepts in the context of the original text to truly understand the clinical meaning.







Location_Start	Location_End	Semantic	Entity
28	63	problem	a palpable mass of the right breast
51	63	bodyloc	right breast
130	136	bodyloc	breast
130	146	problem	breast carcinoma
170	180	treatment	lumpectomy
185	211	test	sentinel lymph node biopsy
213	240	treatment	Frozen section of the tumor
249	262	bodyloc	sentinel node
273	299	problem	granulomatous inflammation
308	325	test	gross examination
353	373	problem	tuberculous mastitis
396	421	treatment	anti-tuberculosis therapy
426	436	temporal	six months
442	444	negation	no
445	457	problem	side effects
461	486	problem	any further complications

## CASE PRESENTATION

The patient presented with a palpable mass of the right breast with clinical, laboratory and mammographic findings indicative of breast carcinoma. The patient underwent lumpectomy and sentinel lymph node biopsy. Frozen section of the tumor and the sentinel node revealed "granulomatous inflammation", while gross examination confirmed the diagnosis of tuberculous mastitis. The patient received anti-tuberculosis therapy for six months with no side effects or any further complications.

<https://jmedicalcasereports.biomedcentral.com/articles/10.1186/1752-1947-2-34>



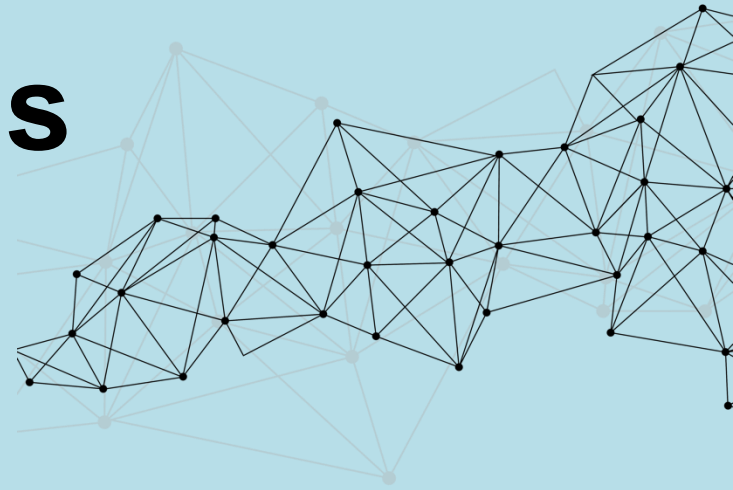
# EMERSE is $\equiv$ fast

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



# Synonyms



- Used for query expansions
- User-controlled
- Multiple datasets can be included
- EMERSE Synonyms
  - acronyms, abbreviations, professional/consumer terms, misspellings, trade/generic drug names, species, chemo regimens, phrase variations, malapropisms, idioms, neologisms, organizations, companies, & more
  - 2.6 million unique entries

# Synonyms for "smoker"



Click individual terms to highlight or de-highlight.

Synonyms (126)

[SORT A TO Z](#)

[HIGHLIGHT ALL](#)

[HIGHLIGHT NONE](#)

- abuse nicotine
- abuses nicotine
- addicted to nicotine
- addicted to tobacco
- addiction to nicotine
- addiction to tobacco
- ash tray
- ash trays
- ashtray
- ashtrays
- chain smoker
- chain smokers
- chain smoking
- chainsmoker
- chainsmokers
- chainsmoking
- cigar
- cigar day
- cigar per day
- cigarette day
- cigarette pack day
- cigarette packs each day
- cigarette packs every day
- cigarette packs per day
- cigarette per day
- cigarette smoke
- cigarette smoker
- cigarette smokers
- cigarette smoking
- cigarettes
- cigarettes day
- cigarettes per day
- cigars
- cigars days
- cigars per day
- continue to smoke
- continued to smoke
- continues to smoke
- continuing to smoke
- current every day smoker
- current everyday smoker
- current smoker
- currently smokes
- dependence on cigarettes
- dependence on nicotine
- dependence on tobacco
- dependent on cigarettes
- dependent on nicotine
- dependent on tobacco
- nicotine
- nicotine abuse
- nicotine abuser
- nicotine abusers
- nicotine addiction
- nicotine addictions
- nicotine dependence
- nicotine dependent
- pack a day
- pack day
- pack each day
- pack history
- pack per day
- pack year
- pack year smoker
- pack years
- pack yr
- pack yrs
- packs
- packs a day
- packs day
- packs each day
- packs of cigarette day
- packs of cigarette each day
- packs of cigarette per day
- packs of cigarettes day
- packs of cigarettes each day
- packs of cigarettes every day
- packs of cigarettes per day
- packs per day
- packs per year
- packyear
- packyears
- packyr
- packyrs
- pipe
- pipes
- pk year
- pk years
- pk yr
- pk yrs

# Synonyms for "social determinant of health"



Click individual terms to highlight or de-highlight.

Synonyms (364)

[SORT A TO Z](#)

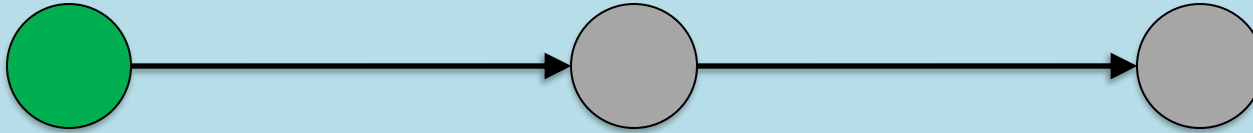
[HIGHLIGHT ALL](#)

[HIGHLIGHT NONE](#)

- abused
- academic problem
- adoptive parent
- afford
- affordable medication
- affordable prescription
- alcoholism
- amphetamine abuse
- antisocial
- appears depressed
- appears sad
- appears to be depressed
- appears to be sad
- assisted living facilit
- assisted living facility
- assistive device
- bankruptcy
- banned
- bath bench
- bath not on first floor
- bed and bath are not on the first floor
- boarding house
- burglary
- care giver support
- cared for by a friend
- cared for by a relative
- charity care
- cheaper
- childhood abuse
- chipped paint
- cigarettes
- cluttered
- co pay
- cocaine abuse
- common law partnership
- community support
- computer illiterate
- condominium
- copay
- correctional facility
- court date
- court mandated
- court ordered
- crime victim
- death of family member
- death of friend
- death of partner
- dementia
- detox
- difficult working conditions
- difficulty maintaining employment
- disability income
- disabled partner
- disabled partners
- disabled relative
- disorderly
- divorced
- drug seeking
- drunk
- dysthymia
- dysthymic
- economic problems
- education level
- education problem
- education problems
- elderly parents
- employment difficulties
- employment income insecurity
- employment insecurity
- estranged
- evict
- evicted
- experiencing discrimination
- exposed to violence
- extended care facility
- extended tub bench
- feeling down
- feels down
- felony
- filthy living
- finances
- financial concerns
- financial constraints
- financial difficulties
- financial issues
- financial strain
- financial stress
- financial stressors
- financially
- fixed income
- flooded
- food bank
- food insecure
- food insecurity
- food pantries
- food stamps
- foreclosed
- forgery
- forging
- foster care
- foster family
- fruit and vegetable intake
- gender identity
- general financial insecurity
- grab bars present
- group home
- gunpoint
- hand held shower
- health insurance
- healthcare problem



# Typical workflow

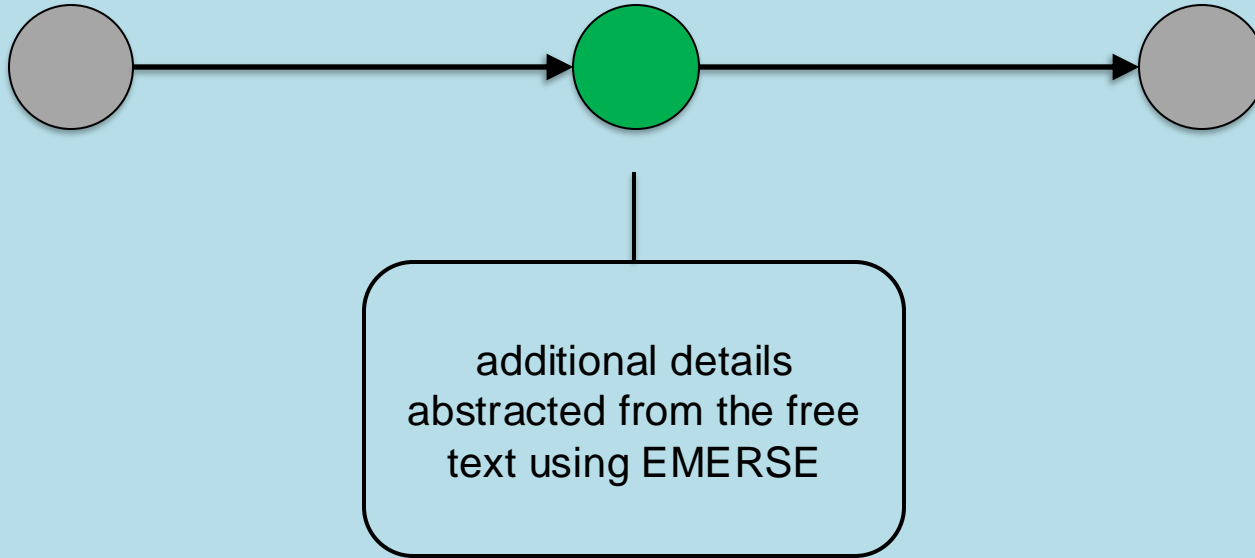


initial cohort identified with a cohort discovery tool using structured data (or via EMERSE 'Find Patients')

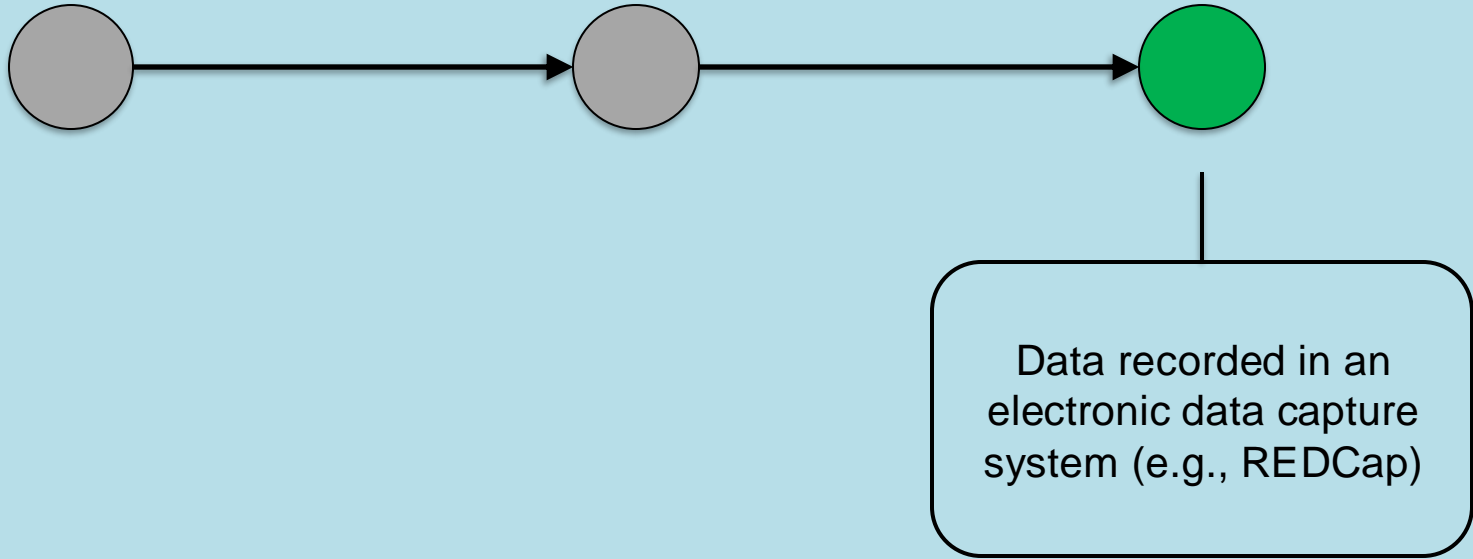
Cohort discovery tools:  
i2b2/ENACT, Leaf, etc.



# Typical workflow



# Typical workflow



# Publications using EMERSE

665

papers and abstracts



Full list at:

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

# Just released...

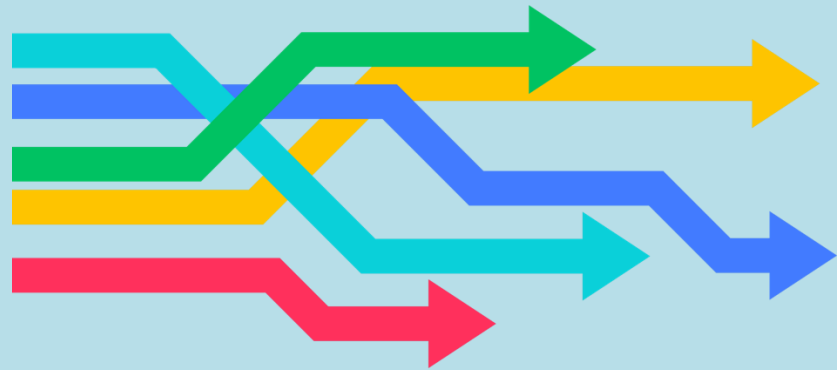
Incorporation of NLP features

- negation
- uncertainty
- subject (patient vs other)
- named entity recognition/mapping to ontologies



If you don't like our built-in NLP, add your own

# Future Work



Data extraction from templated notes

Integration with ChatGPT or similar tools

& more...

# How much does it cost?

Software	\$0.00
Installation Guidance	\$0.00
Access to Documentation	\$0.00
Training (within reason)	\$0.00
Synonyms dataset (optional)	\$0.00*



\* For research use within EMERSE

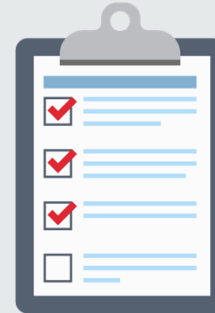
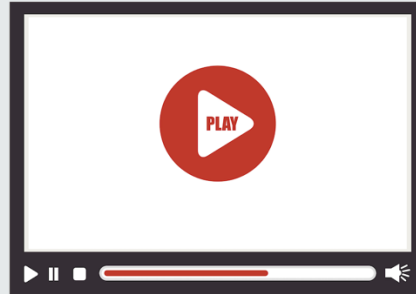


# Fine Print

- There will be local operational costs
  - free kitten < free pony
- Sites need to determine how to extract their data to get it into EMERSE
  - We can provide recommendations
- Sites install EMERSE themselves
  - on premise or in the cloud
- Our team does not have access to anyone's data other than U of Michigan



project-emerse.org







Twitter: @projectEMERSE

publications

software releases

announcements

webinars

# Interested in EMERSE?



Contact us to schedule a time with your team for:

- Discussions about usage strategies
- Training
- Live demonstrations (abstractors, IT teams, etc)

# Community Meeting



TOMORROW: October 1, 2024, 1-2 PM ET

Open to everyone, find registration link at

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



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

EMERSE-team@umich.edu



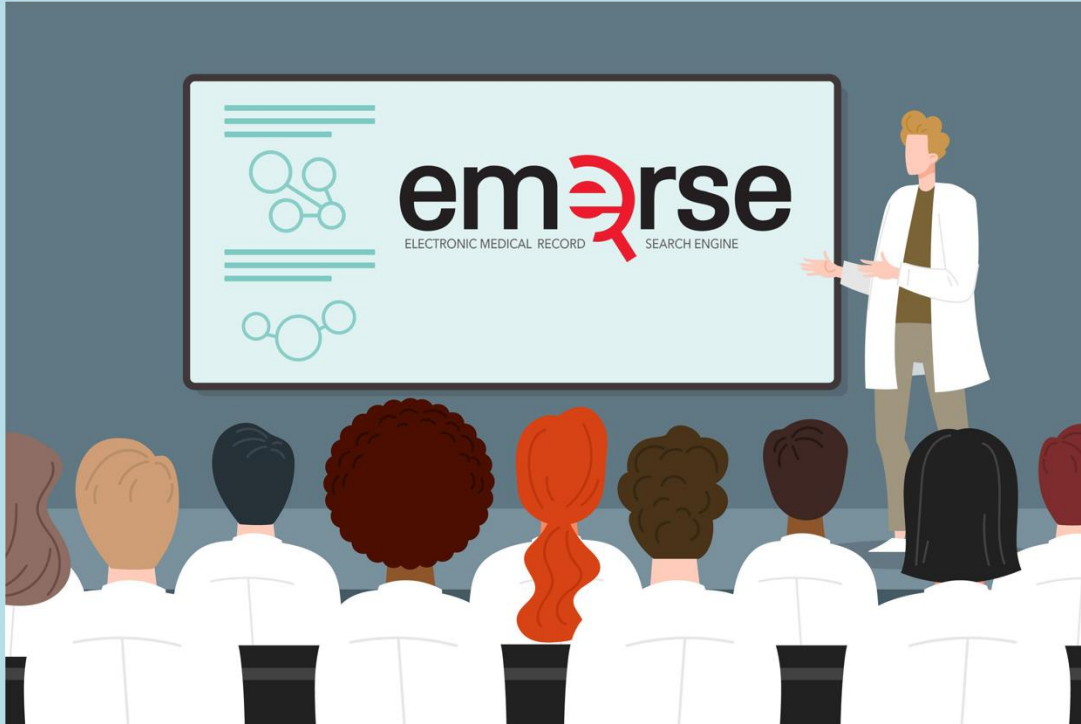
Lisa Ferguson

David Hanauer

Kellen McClain

Guan Wang

# Live Demo



- No real names
- No PHI
- Publicly available
- Abstracts & case reports