

David Hanauer, MD, MS University of Michigan

hanauer@umich.edu

Twitter: @DavidHanauer2

MARCH Informatics Committee Mtg July 7, 2020

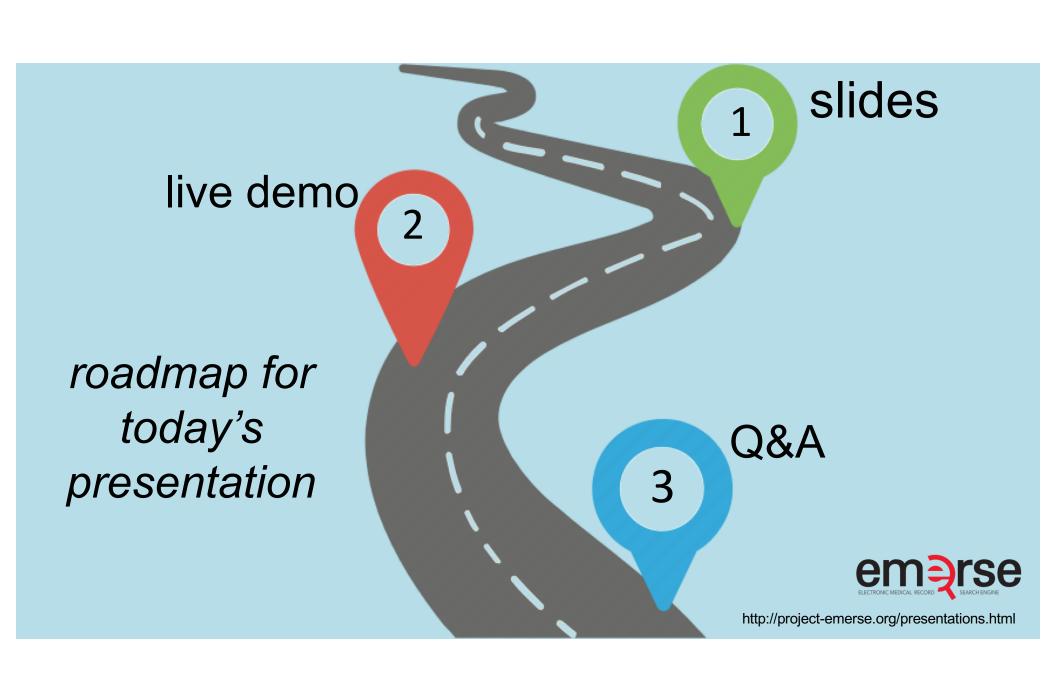
# If you're thinking of taking notes



these slides can be found at:

link will be on most slides





#### 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



Designed to be self-serve\*

\*Once access is approved



#### What is EMERSE?

- A research tool, available at no cost
  - Open source, Apache 2.0 licensing
  - Picking the right tool for your research is important
    - · multiple tools are often needed



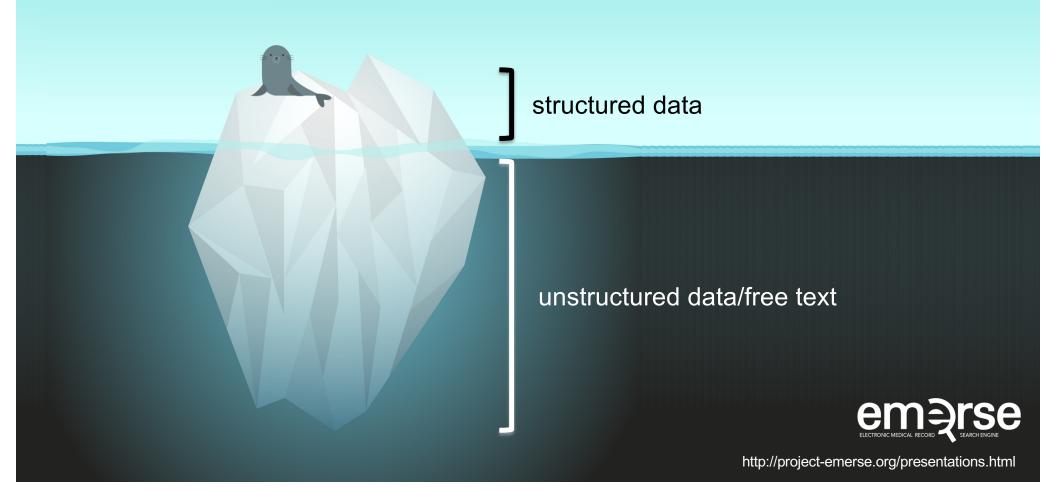


### **Unstructured vs Structured Data**

EMERSE is for this	i2b2 is for this	
Unstructured Data (free text)	Structured Data	
Mrs. Jones is a 56 year old female with a h/o HTN, T2DM, and, and high cholesterol who comes to the clinic today with a 3 day h/o coughing, fever, and severe headache.	WBC: Total cholesterol: Weight: AST: ALT:	5.6 182 67.4 30 52



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



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

"potential cases may be missed or wrongly dated when coded data alone are used.." (PMID: 23964710)

"structured data is best used for first pass filtering of EHR data...prior to a more nuanced second pass using unstructured data" (PCMID: PMC4333685)

"...urinary incontinence, an important patient-centered outcome following prostate cancer treatment, was reported almost exclusively in the free text of EHRs and was rarely coded as an ICD 9 diagnosis code." (PMCID: PMC4899050)

"Claims and structured EHR data give an incomplete picture of burden related to geriatric syndromes. Geriatric syndromes are likely to be missed if unstructured data are not analyzed." (PMID: 29972595)

#### 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: augmented intelligence

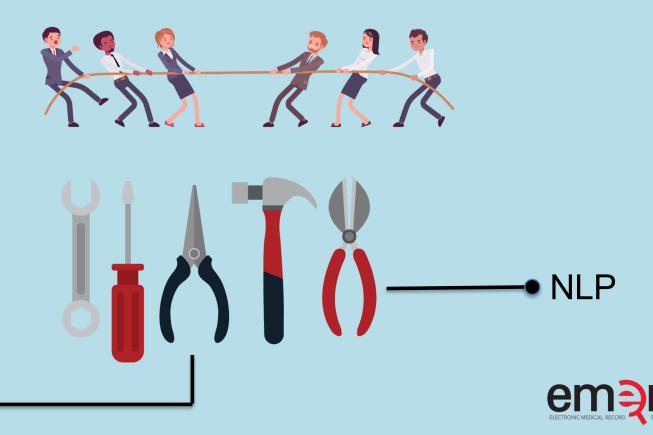


NLP: artificial intelligence



# **EMERSE vs NLP**

**EMERSE** 



#### With EMERSE: find cohorts

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

- diseases
- drugs
- symptoms
- more!





#### **Find cohorts**

It's perfect for finding diagnoses that don't have a specific ICD-10 code... like leiomyosarcoma...

#### ▶ ICD-10-CM Diagnosis Code C49.6 | [convert to ICD-9-CM]

Malignant neoplasm of connective and soft tissue of trunk, unspecified

Malignant neoplasm of conn and soft tissue of trunk, unsp; Angiosarcoma, trunk; Cancer of the soft tissue, trunk; Leiomyosarcoma, trunk; Liposarcoma, trunk; Malignant fibrous histiocytoma, trunk; Primary angiosarcoma of trunk; Primary leiomyosarcoma of trunk; Primary liposarcoma of trunk; Primary malignant fibrous histiocytoma of trunk; Primary malignant neoplasm of soft tissues of trunk; Primary rhabdomyosarcoma of trunk; Primary sarcoma of trunk; Rhabdomyosarcoma, trunk; Sarcoma, trunk; Malignant neoplasm of back NOS



# Highlight documents for chart review

A chest x-ray was obtained and revealed the presence of patchy consolidations in bilateral lower lobes. CT chest/abdomen/pelvis (Fig. 1) was notable for diffuse ground-glass opacities in the periphery of bilateral lungs as well as splenomegaly with a splenic hypodensity measuring 8.6 × 0.7 cm and surrounding peri-splenic edema, consistent with a splenic infarct. Hydroxychloroquine and a one-time dose of Tocilizumab were administered along with supportive measures. A continuous heparin infusion was also initiated and patient was admitted to the general medical ward for further workup.

https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7256550/





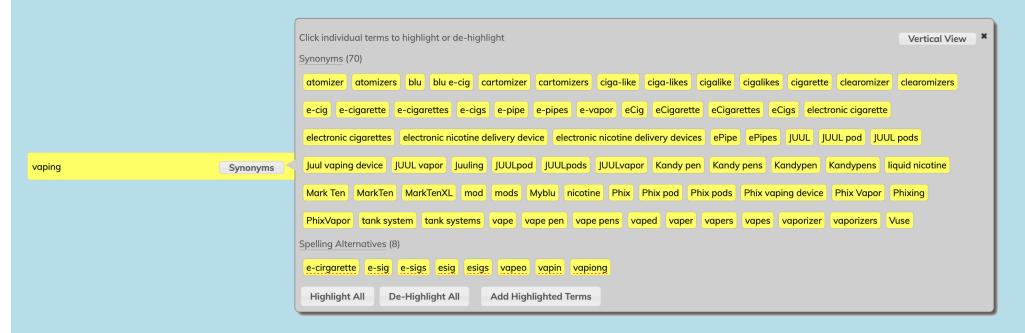
## **Chart review**

Research team at Michigan Medicine used EMERSE to identify presenting symptoms for COVID-19, including:

dyspnea	rhinorrhea or congestion	vomiting	
wheezing	myalgia or arthralgia	abdominal pain	
cough	fatigue or malaise	chest pain or tightness	
sputum production	headache	loss of smell or taste	
blood in sputum	loss of appetite	altered mental status	
sore throat	diarrhea	weakness	
fever or chills	nausea	lightheadedness or syncope	



# **EMERSE** query expansion





# **EMERSE** query expansion

- Can use "EMERSE Synonyms" or upload your own dataset (e.g., ICD-10, human phenotype ontology, MeSH terms)
- Simple file format
  - http://project-emerse.org/documentation/administrator\_guide.html#truepreparing-a-synonyms-file



# **EMERSE Synonyms**

- Manually curated over past 12+ years
- Terms/phrases not found in standard vocabularies, taxonomies, ontologies (but found in clinical notes)
- Academic and non-academic licensing available\*
- > 1.1 million unique rows

\*David Hanauer is entitled to a portion of the license fees

Details are at:

http://project-emerse.org/synonyms.html

To explore a subset, visit:

http://www.davidhanauer.com/EMERSE\_synonyms\_T\_Terms/

# **EMERSE Synonyms**

Category	Example
Acronyms	CXR chest x-ray
Abbreviations	chemotherapy chemo
Professional and consumer terms	emesis puking
Misspellings	diarriea diahhrea
Trade and generic drug names	Tylenol acetaminophen
Chemotherapy regimens	FOLFOX, folinic acid, fluorouracil, oxaliplatin
Word Stems	sleep, sleeping, sleepy

Category	Example
True synonyms	frigid cold
Phrasing variations	lido w/ epi epi-lidocaine
Malapropisms/ usage errors	beast mass prostrate cancer
Root word variations	adenotonsillectomy tonsilloadenoidectomy
Hyphenations	pseudosarcomatous pseudo-sarcomatous
Idioms	under the knife under the weather
Neologisms	trabecularity (vs trabeculation)

## **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/)



#### What's in EMERSE?

#### Free text notes

- admission notes
- history and physical
- progress notes
- pathology reports
- radiology reports
- more!

@ Michigan: 2.5 million patients

170 million notes

1995-present





#### What's not in EMERSE?

#### Structured data

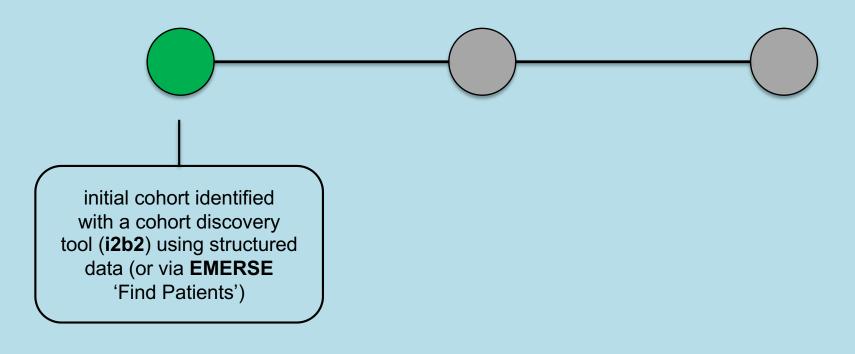
- discrete lab values
- vital signs
- billing codes

Other tools such as **i2b2** exist for structured data



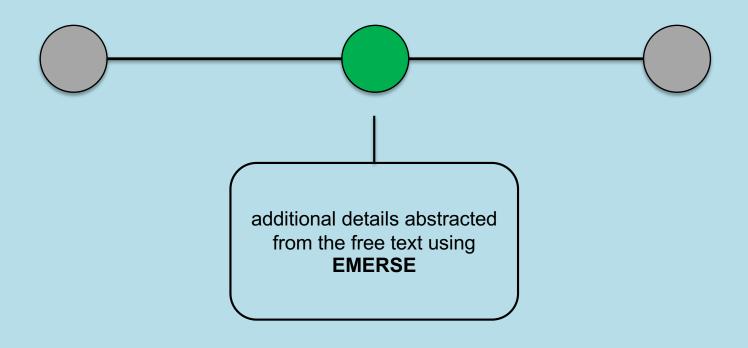


# **Typical workflow**



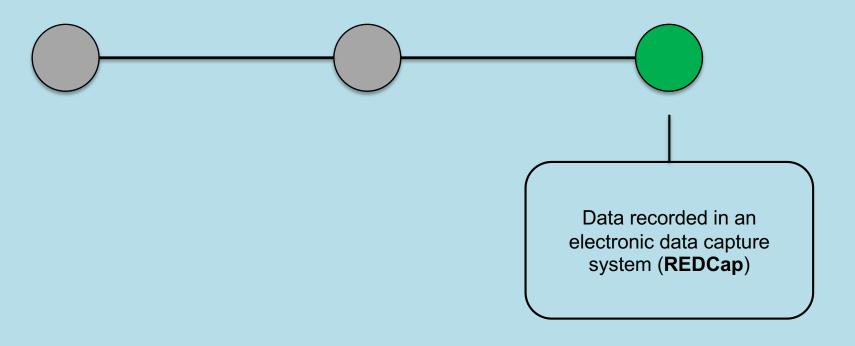


# **Typical workflow**





# **Typical workflow**





# Publications using EMERSE

354 papers54 abstracts



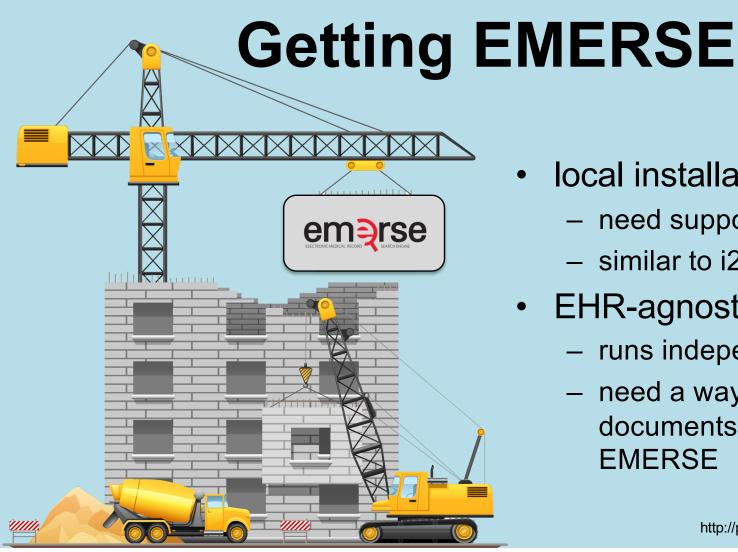
Full list can be found at:

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









- local installation required
  - need support of IT team
  - similar to i2b2
- **EHR-agnostic** 
  - runs independent of EHR
  - need a way to extract documents to send to **EMERSE**

# Test driving EMERSE

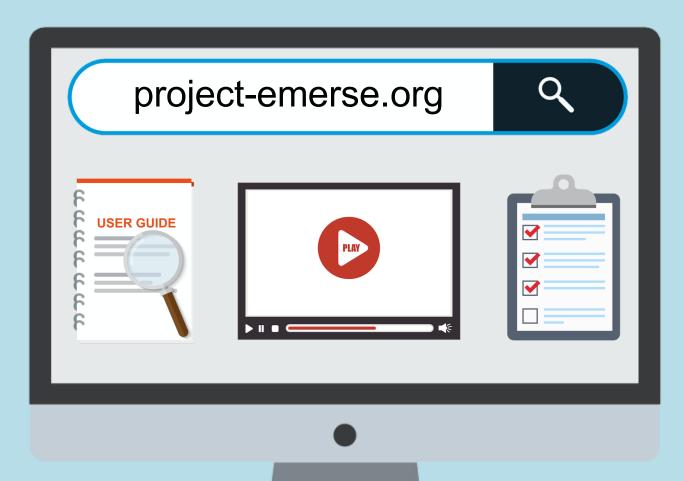
em<sub>3</sub>rse

online demo videos on website virtual machine

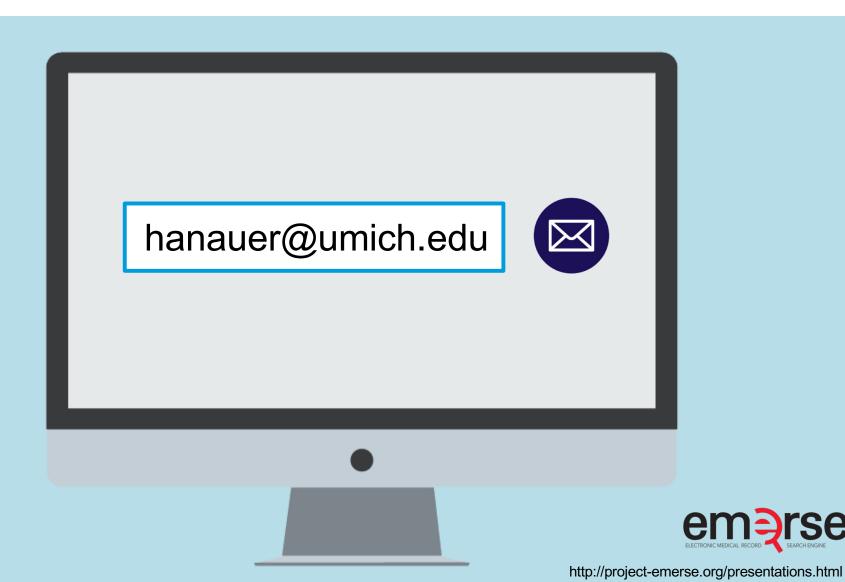
> could be used for individual projects











# Acknowledgements



U of Michigan Rogel Cancer Center [P30CA046592]



Michigan Institute for Clinical and Health Research (MICHR) [UL1TR000433]



U of Michigan Health Information and Technology Services (HITS)



NCI Information Technology for Cancer Research (ITCR) [U24CA204863]

