Introduction to EMERSE

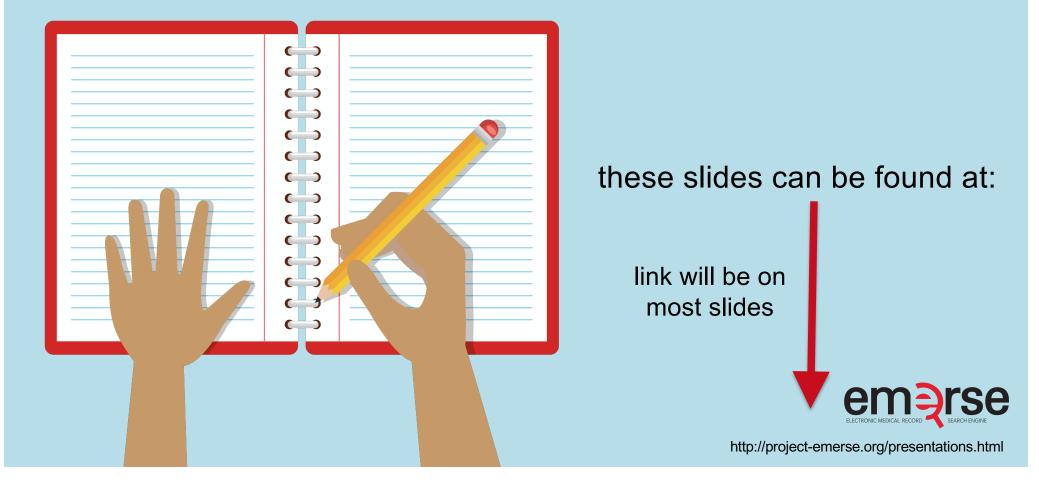


Twitter:

David Hanauer, MD, MS University of Michigan hanauer@umich.edu @informaticsGeek
@projectEMERSE

September 20, 2021

If you're thinking of taking notes



Disclosure

LICENSE

EMERSE is available at no cost under the Apache 2.0 open source license

An optional addition to EMERSE is the "Synonyms Collection", for which David Hanauer is entitled to a portion of the license fees



slides 1 live demo 2 roadmap for today's Q&A presentation 3 se http://project-emerse.org/presentations.html



What is EMERSE?

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





What is EMERSE?

A research tool, available at no cost

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



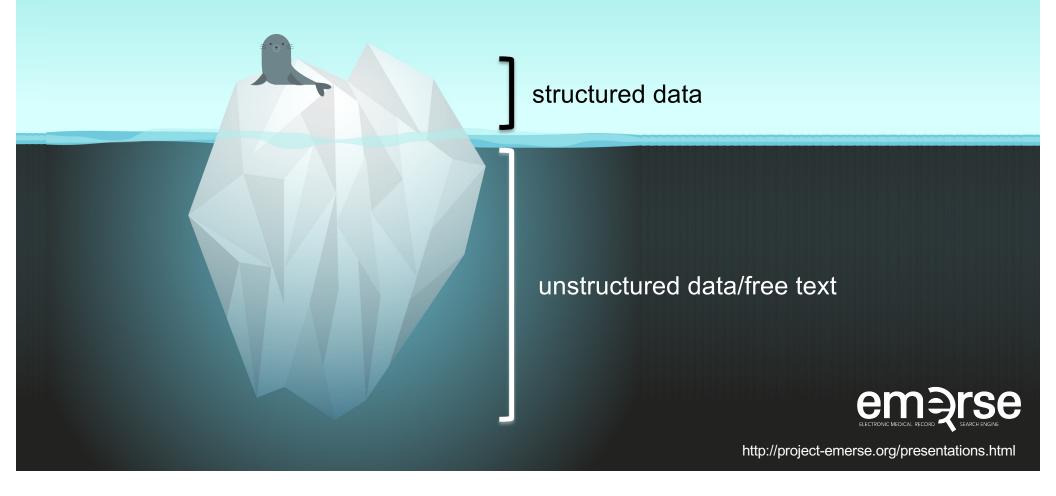


Unstructured vs Structured Data

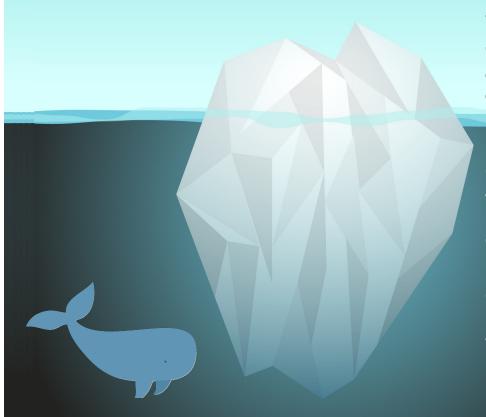
EMERSE is for this	not this	
Unstructured Data (free text)	Structured Data	
Mrs. Jones is a 56 year old female with a history of HTN, hypercholesterolemia, and T2DM who comes to the clinic today with a 3 day h/o dizziness and severe headache on the left side.	WBC: Total cholesterol: Weight: AST: ALT: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



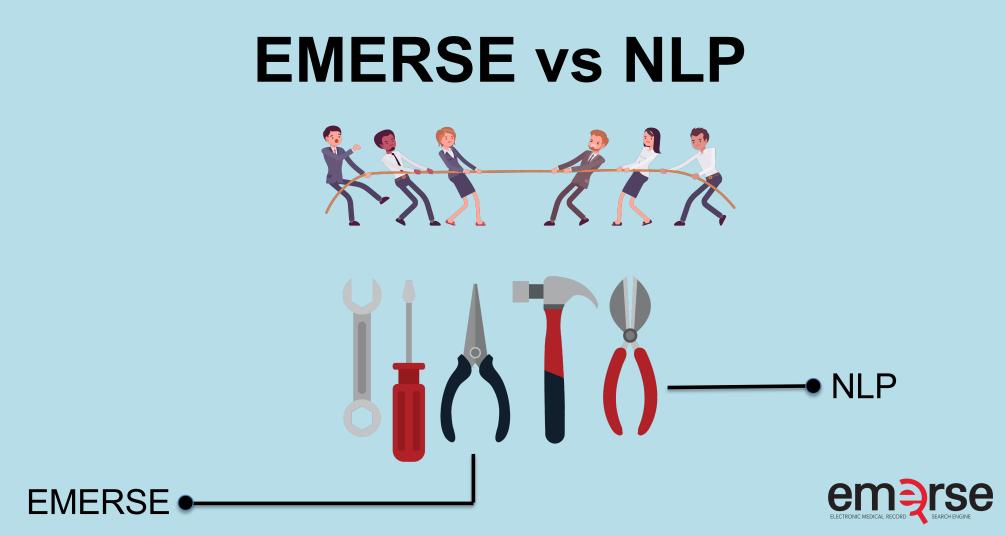


EMERSE: augmented intelligence



NLP: artificial intelligence





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

- diseases
- drugs
- symptoms
- anything*

*if it is mentioned

1 ===



It's perfect for finding rare things...

...like rare cancers



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



https://www.icd10data.com/search?s=leiomyosarcoma

... or Bainbridge-Ropers Syndrome

ICD-10-CM Code **Q87.0**

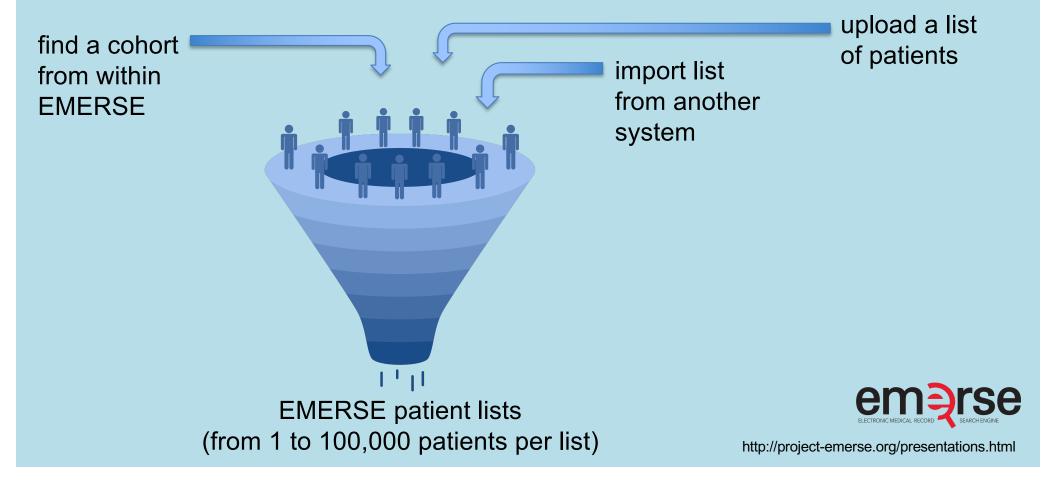
Congenital malformation syndromes predominantly affecting facial appearance

BILLABLE POA Exempt | ICD-10 from 2011 - 2016

...or hundreds of other disorders with no specific code

https://www.icd10data.com/search?s=leiomyosarcoma

Working with cohorts



Highlight documents for chart review

Thoracocentesis confirmed the recurrence of mantle cell lymphoma. Disease restaging work-up revealed multicompartment 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





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 fatique 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 lightheadedness or syncope nausea





https://pubmed.ncbi.nlm.nih.gov/33046294/

EMERSE query expansion

Х

CANCEL





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

Synonyms for "vaping"

Click individual terms to highlight or de-highlight.

e sig e sigs esig esigs

e cirgarette

Synonyms (120) HIGHLIGHT ALL HIGHLIGHT NONE			
atomizer atomizers banana ice blu blu e cig cartomizer cartomizers cherry crush ciga like ciga likes			
cigalike cigalikes cigarette clearomizer clearomizers e cig e cigarette e cigarettes e cigs e hooka			
e hookah e hookahs e hookas e liquid e liquids e pipe e pipes e vapor ecig ecigarette ecigarettes			
ecigs electronic cigarette electronic cigarettes electronic hooka electronic hookah electronic hookahs			
electronic hookas electronic nicotine delivery device electronic nicotine delivery devices eliquid eliquids			
epipe epipes flavor pod flavor pods juice for vape juice for vaping juices for vape juices for vaping			
juul juul pod juul pods juul vaping device juul vapor juuling juulpod juulpods juulvapor kandy pen			
kandy pens kandypen kandypens liquid nicotine mark ten markten marktenxl mod mods myblu			
nicotine njoy phix phix pod phix pods phix vaping device phix vapor phixing phixvapor polar mint			
posh puff puff bar puff bars smok stick smok sticks smok v8 stick smok v8 stick stig tank system			
tank systems vape vape juice vape juices vape liquid vape liquids vape pen vape pens vape stick			
vape sticks vape tank vape tank system vape tank systems vape tanks vaped vaper vapers vapes			
vaping juice vaping juices vaping liquid vaping liquids vaping tank vaping tank system			
vaping tank systems vaping tanks vaporizer vaporizers vivid vanilla vuse			
Spelling Alternatives (8) HIGHLIGHT ALL HIGHLIGHT NONE			

vapeo

vapin

vapiong

ADD HIGHLIGHTED TERMS (128)

EMERSE query expansion

 Can use "EMERSE Synonyms Collection" or administrators can upload additional datasets (e.g., ICD-10, human phenotype ontology, MeSH terms)



EMERSE Synonyms Collection

- Manually curated over past 14+ years
- Terms/phrases not found in standard vocabularies, taxonomies, ontologies (but found in clinical notes)
- Academic and non-academic licensing available
- > 1.2 million unique rows
- 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/PMC3860

em	em		1		
	=				- 1
=		≣		<u>em⊋rs</u> e	\equiv
38601	74/)	=	Ξ		



EMERSE is — fast

Query to identify all patients with the following	Epic Clarity 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



What's in EMERSE?

Free text notes*

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

*depends on local configuration

@ Michigan:

2.6+ million patients210+ million notes1995-present





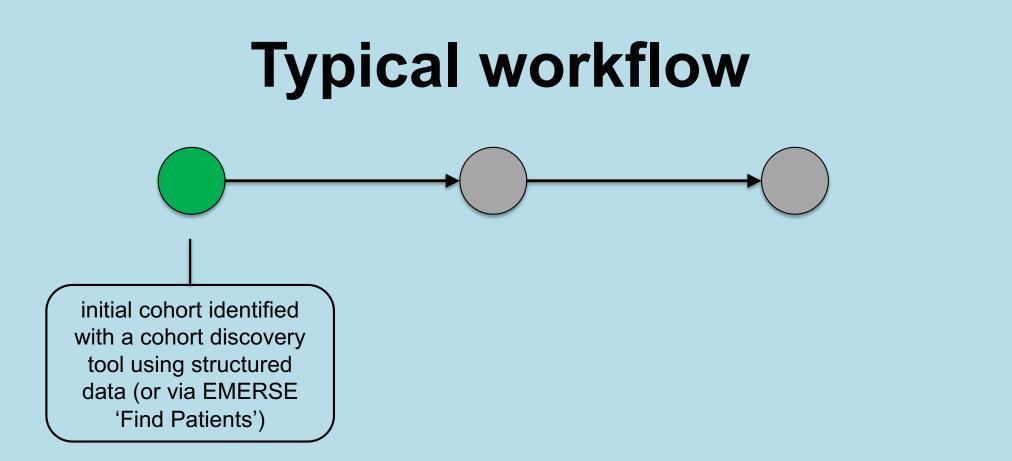
What's not in EMERSE?

Structured data

- discrete lab values
- vital signs
- billing codes

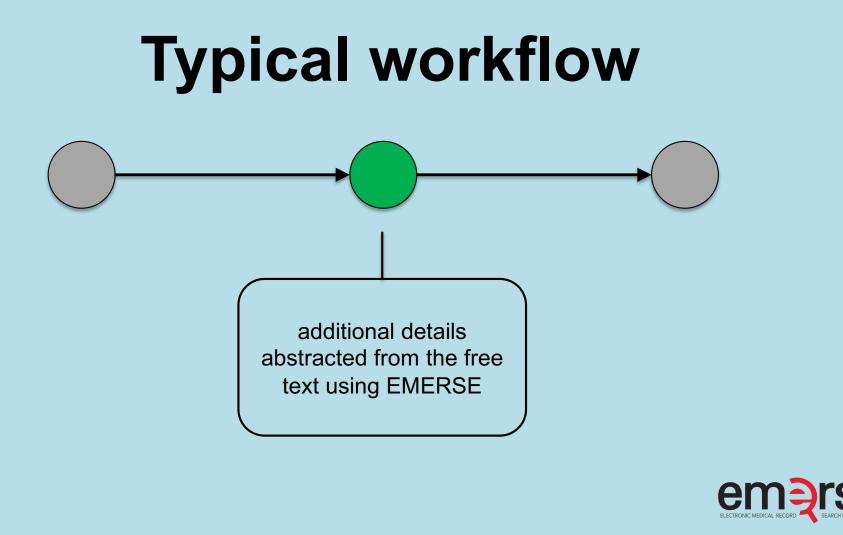


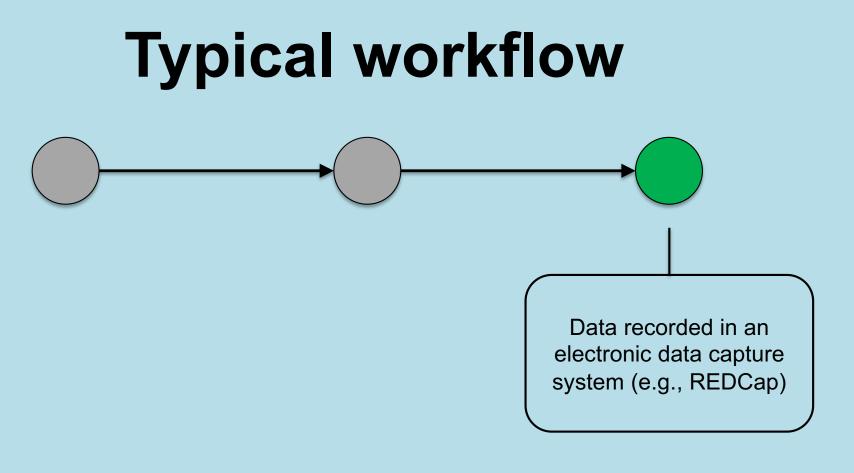
Other tools exist for structured data (DataDirect, i2b2, Leaf, etc.)



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









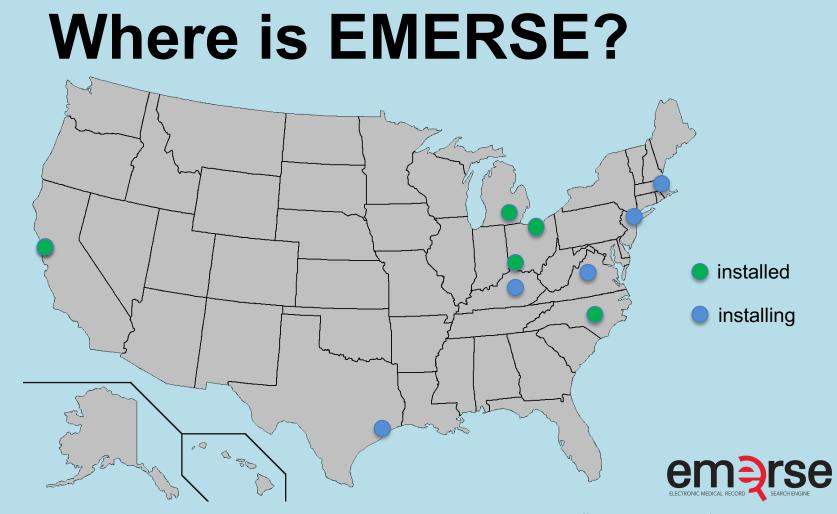
Publications using EMERSE

448 papers82 abstracts

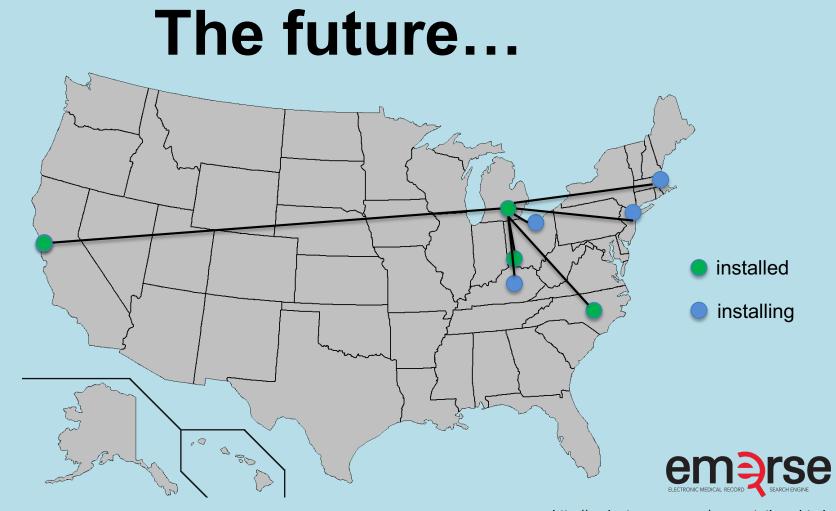


Full list at: http://project-emerse.org/publications.html





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



The future...

If anyone is interested in learning more about the network feature, or to provide your own insights/feedback, we are having a meeting about it *tomorrow* (Tuesday, September 21, 2021) at 11 AM ET.

Register for the meeting at this link:

https://umich.zoom.us/meeting/register/tJEtcumppzwoE tcoSf-bBp38DbolTih8KERe

Who can obtain access?

@ Michigan

- Anyone with EHR access, BUT:
- you need a valid reason such as research, operations, quality
- "curiosity" is not a valid reason*
- all usage is tracked
- @ elsewhere

Will depend on each site's local rules

*but we do have a demo site for curious users







local installation required

need support of IT team



Under the Hood



EMERSE application (Java, Javascript)



Apache Solr



Database



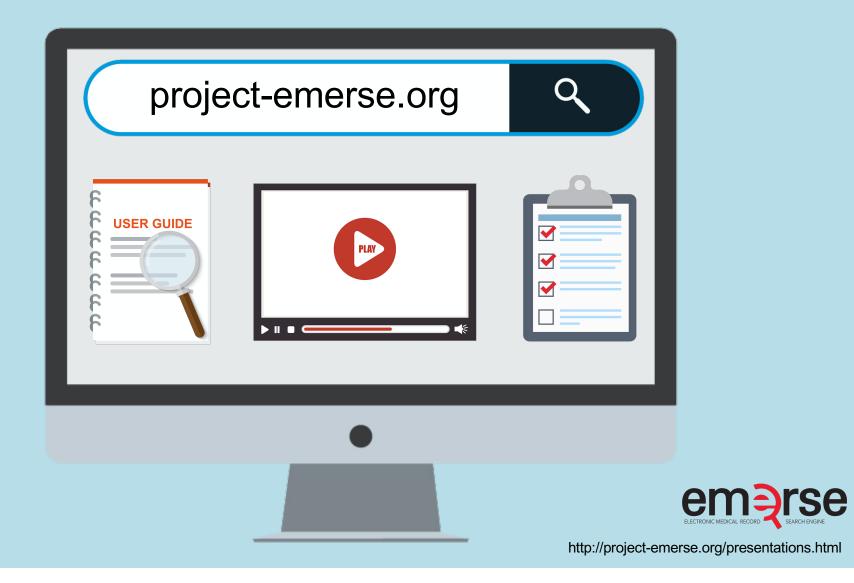


Test driving EMERSE

online demo (contact us) videos on website









Twitter: @projectEMERSE

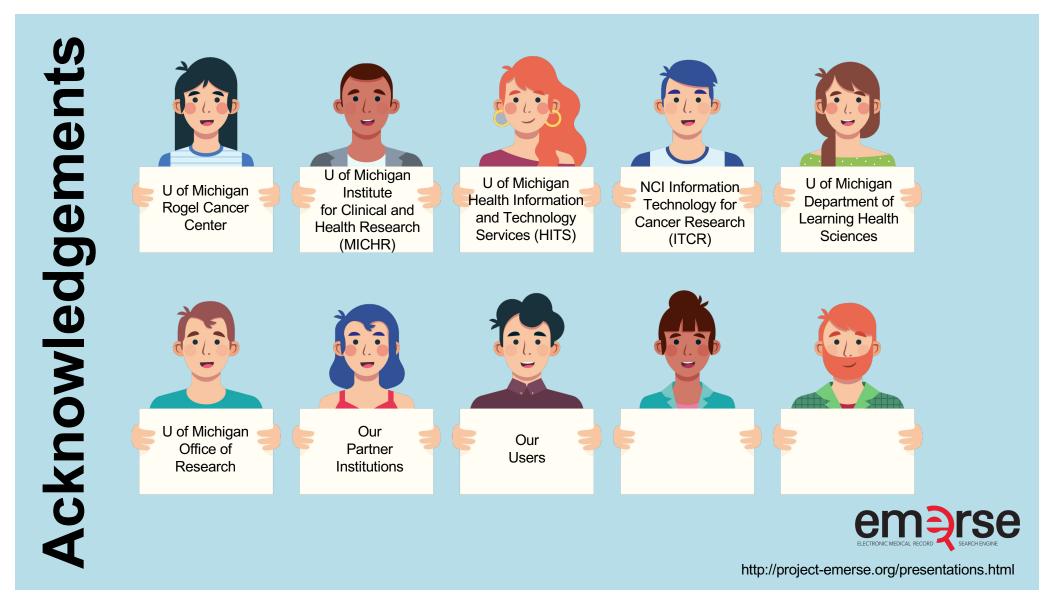
publications software releases announcements webinars



EMERSE-team@umich.edu







slides 1 live demo 2 roadmap for today's Q&A presentation 3 se http://project-emerse.org/presentations.html

Live demonstration



- 10,000 simulated patients
- 500,000 simulated documents
- Look and options will differ depending on version and local setup (some components may not be available to all users/sites)

