EMERSE: an easy-to-use, self-service search engine and chart review tool for EHR notes
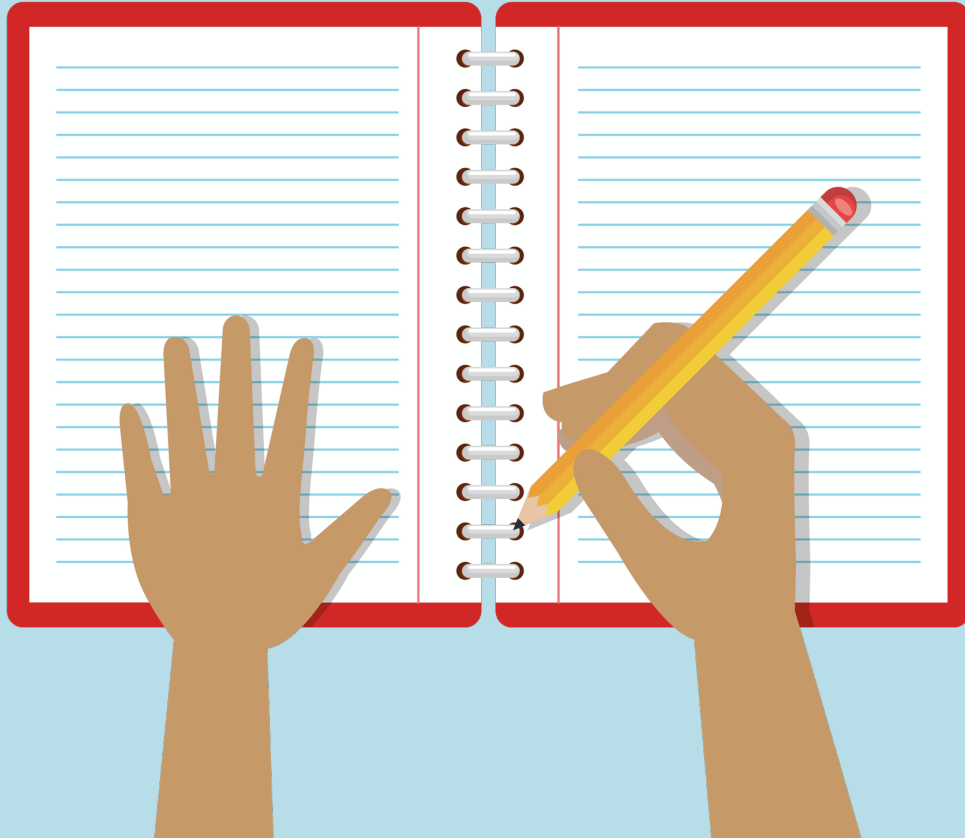
Rogel Cancer Center
Cancer Control and Population Sciences Retreat
September 28, 2023

Twitter: @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're thinking of taking notes or want to visit links

these slides can be found at:

this link will be on most slides

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

Disclosures

Disclosures

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

Royalties: EMERSE "Synonyms" (used for query expansion) which the U of Michigan licenses

Free for academic use within EMERSE

I am here to tell you about EMERSE

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 <u>far too few people both within and outside of their network knew about the tool's existence</u>."

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: 5.6 Total cholesterol: 182 Weight: 67.4 AST: 30 ALT: 52

emerse
ELECTRONIC MEDICAL RECORD   SEARCH ENGINE

80% of EHR data are in unstructured free-text

structured data

unstructured data/free text
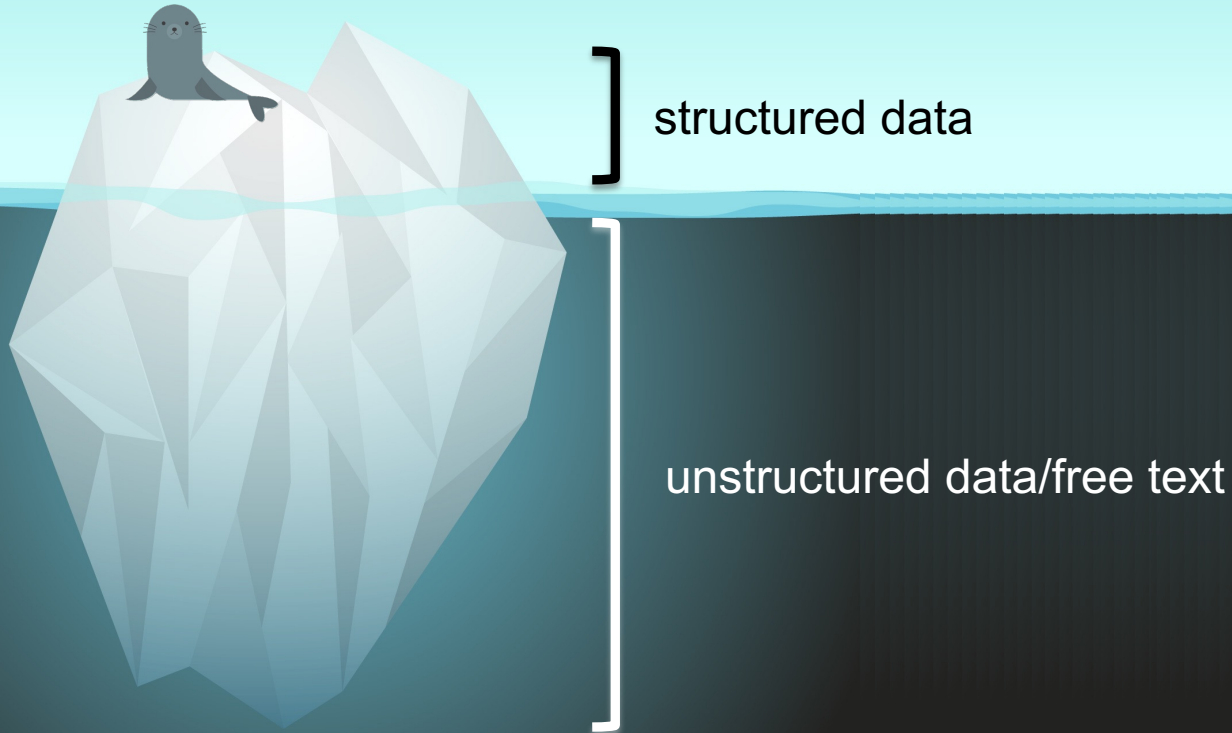
Most medical centers have no self-service tools to explore free-text data

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 researchers

- Unlike with some EHRs, EMERSE can search across all notes and all patients at once

Find cohorts

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

– diseases

– drugs

– symptoms

– anything*
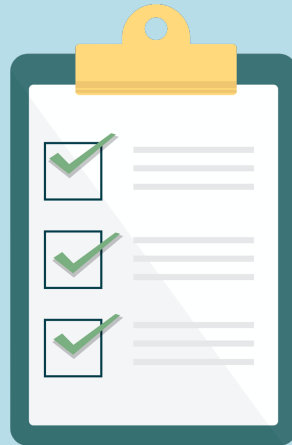
*if it is mentioned

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"

emerse
ELECTRONIC MEDICAL RECORD SEARCH ENGINE

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.

Our philosophy

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

emerse
ELECTRONIC MEDICAL RECORD   SEARCH ENGINE

MetaMap output

Column 1:

Meta Mapping (1000):
  1000 **^patient (Patients) [Patient or Disabled Group]
Meta Mapping (702):
  742 Presented (Presentation) [Idea or Concept]
  742 PALPABLE (Palpable) [Qualitative Concept]
  784 right Breast mass (Lump in right breast) [Finding]
Meta Mapping (1000):
  1000 Clinical Laboratory (Clinical Laboratory Services) [Health Care Activity]
Meta Mapping (1000):
  1000 Clinical Laboratory (Laboratories, Clinical) [Health Care Related Organization,Manufactured Object]
Meta Mapping (947):
  947 Mammography finding [Finding]
Meta Mapping (743):
  715 CARCINOMA OF BREAST (Breast Carcinoma) [Neoplastic Process]
Meta Mapping (1000):
  1000 **^patient (Patients) [Patient or Disabled Group]
Meta Mapping (1000):
  1000 LUMPECTOMY (Lumpectomy of breast) [Therapeutic or Preventive Procedure]
Meta Mapping (1000):
  1000 SENTINEL LYMPH NODE BIOPSY (Sentinel Lymph Node Biopsy) [Diagnostic Procedure]
Meta Mapping (745):
  806 Frozen Section (Frozen Sections) [Tissue]
  593 Tumour (Neoplasms) [Neoplastic Process]
Meta Mapping (745):
  806 Frozen Section (Frozen Sections) [Tissue]
  593 Tumor (Tumor tissue sample) [Tissue]
Meta Mapping (745):
  806 Frozen Section (Frozen Sections) [Tissue]
  593 Tumor (Specimen Source Codes - tumor) [Intellectual Product]
Meta Mapping (745):
  806 Frozen Section (Frozen Sections) [Tissue]
  593 Tumor (Tumor Mass) [Finding]
Meta Mapping (745):
  806 Frozen Section (Cryoultramicrotomy) [Laboratory Procedure]
  593 Tumour (Neoplasms) [Neoplastic Process]
Meta Mapping (745):
  806 Frozen Section (Cryoultramicrotomy) [Laboratory Procedure]
  593 Tumor (Tumor tissue sample) [Tissue]
Meta Mapping (745):
  806 Frozen Section (Cryoultramicrotomy) [Laboratory Procedure]
  593 Tumor (Specimen Source Codes - tumor) [Intellectual Product]

Column 2:

Meta Mapping (745):
  806 Frozen Section (Cryoultramicrotomy) [Laboratory Procedure]
  593 Tumor (Tumor Mass) [Finding]
Meta Mapping (1000):
  1000 Sentinel node (Sentinel node (disorder)) [Disease or Syndrome]
Meta Mapping (1000):
  1000 Sentinel Node (Sentinel Lymph Node) [Body Part, Organ, or Organ Component]
Meta Mapping (1000):
  1000 Revealed [Qualitative Concept]
Meta Mapping (1000):
  1000 Granulomatous Inflammation (Granulomatous inflammation) [Pathologic Function]
Meta Mapping (1000):
  1000 Gross examination (Sample macroscopy) [Laboratory Procedure]
Meta Mapping (1000):
  1000 confirmed (Confirmed by) [Qualitative Concept]
Meta Mapping (1000):
  1000 Confirmed (Confirmation) [Finding]
Meta Mapping (696):
  760 DIAGNOSIS (Diagnosis) [Diagnostic Procedure]
  593 Tuberculous (Tuberculosis) [Disease or Syndrome]
  593 MASTITIS (Mastitis) [Disease or Syndrome]
Meta Mapping (696):
  760 DIAGNOSIS (Diagnosis) [Diagnostic Procedure]
  593 Tuberculous (Tuberculosis) [Disease or Syndrome]
  593 Mastitis (Inflammatory disorder of breast) [Disease or Syndrome]
Meta Mapping (696):
  760 DIAGNOSIS (Diagnosis) [Diagnostic Procedure]
  593 Tuberculous (Tuberculosis) [Disease or Syndrome]
  593 Mastitis (Inflammation of non-human mammary gland) [Disease or Syndrome]
Meta Mapping (696):
  760 Diagnosis (Disease) [Disease or Syndrome]
  593 Tuberculous (Tuberculosis) [Disease or Syndrome]
  593 MASTITIS (Mastitis) [Disease or Syndrome]
Meta Mapping (696):
  760 Diagnosis (Disease) [Disease or Syndrome]
  593 Tuberculous (Tuberculosis) [Disease or Syndrome]
  593 Mastitis (Inflammatory disorder of breast) [Disease or Syndrome]
Meta Mapping (696):
  760 Diagnosis (Disease) [Disease or Syndrome]
  593 Tuberculous (Tuberculosis) [Disease or Syndrome]
  593 Mastitis (Inflammation of non-human mammary gland) [Disease or Syndrome]

Column 3:

Meta Mapping (696):
  760 DIAGNOSIS (Diagnosis Code) [Intellectual Product]
  593 Tuberculous (Tuberculosis) [Disease or Syndrome]
  593 MASTITIS (Mastitis) [Disease or Syndrome]
Meta Mapping (696):
  760 DIAGNOSIS (Diagnosis Code) [Intellectual Product]
  593 Tuberculous (Tuberculosis) [Disease or Syndrome]
  593 Mastitis (Inflammatory disorder of breast) [Disease or Syndrome]
Meta Mapping (696):
  760 DIAGNOSIS (Diagnosis Code) [Intellectual Product]
  593 Tuberculous (Tuberculosis) [Disease or Syndrome]
  593 Mastitis (Inflammation of non-human mammary gland) [Disease or Syndrome]
Meta Mapping (696):
  760 diagnosis (diagnosis aspect) [Functional Concept]
  593 Tuberculous (Tuberculosis) [Disease or Syndrome]
  593 MASTITIS (Mastitis) [Disease or Syndrome]
Meta Mapping (696):
  760 diagnosis (diagnosis aspect) [Functional Concept]
  593 Tuberculous (Tuberculosis) [Disease or Syndrome]
  593 Mastitis (Inflammatory disorder of breast) [Disease or Syndrome]
Meta Mapping (696):
  760 diagnosis (diagnosis aspect) [Functional Concept]
  593 Tuberculous (Tuberculosis) [Disease or Syndrome]
  593 Mastitis (Inflammation of non-human mammary gland) [Disease or Syndrome]
Meta Mapping (696):
  760 DIAGNOSIS (Diagnosis Study) [Research Activity]
  593 Tuberculous (Tuberculosis) [Disease or Syndrome]
  593 MASTITIS (Mastitis) [Disease or Syndrome]
Meta Mapping (696):
  760 DIAGNOSIS (Diagnosis Study) [Research Activity]
  593 Tuberculous (Tuberculosis) [Disease or Syndrome]
  593 Mastitis (Inflammatory disorder of breast) [Disease or Syndrome]
Meta Mapping (696):
  760 DIAGNOSIS (Diagnosis Study) [Research Activity]
  593 Tuberculous (Tuberculosis) [Disease or Syndrome]
  593 Mastitis (Inflammation of non-human mammary gland) [Disease or Syndrome]
Meta Mapping (696):
  760 Diagnosis (Date of diagnosis) [Temporal Concept]
  593 Tuberculous (Tuberculosis) [Disease or Syndrome]
  593 MASTITIS (Mastitis) [Disease or Syndrome]
Meta Mapping (696):
  760 Diagnosis (Date of diagnosis) [Temporal Concept]
  593 Tuberculous (Tuberculosis) [Disease or Syndrome]
  593 Mastitis (Inflammatory disorder of breast) [Disease or

Column 4:

Syndrome]
Meta Mapping (696):
  760 Diagnosis (Date of diagnosis) [Temporal Concept]
  593 Tuberculous (Tuberculosis) [Disease or Syndrome]
  593 Mastitis (Inflammation of non-human mammary gland) [Disease or Syndrome]
Meta Mapping (1000):
  1000 **^patient (Patients) [Patient or Disabled Group]
Meta Mapping (1000):
  1000 RECEIVED (Receive) [Qualitative Concept]
Meta Mapping (719):
  753 TUBERCULOSIS (Tuberculosis) [Disease or Syndrome]
  753 therapy (therapeutic aspects) [Functional Concept]
  623 Six months [Temporal Concept]
Meta Mapping (719):
  753 TUBERCULOSIS (Tuberculosis) [Disease or Syndrome]
  753 Therapy (Therapeutic procedure) [Therapeutic or Preventive Procedure]
  623 Six months [Temporal Concept]
Meta Mapping (719):
  753 TUBERCULOSIS (Tuberculosis) [Disease or Syndrome]
  753 Therapy (Therapy Object (animal model)) [Finding]
  623 Six months [Temporal Concept]
Meta Mapping (1000):
  1000 side effects (aspects of adverse effects) [Functional Concept]
Meta Mapping (1000):
  1000 N side effects (Adverse event) [Pathologic Function]
Meta Mapping (1000):
  1000 N Side effects (Adverse effects) [Pathologic Function]
Meta Mapping (888):
  694 Further [Spatial Concept]
  861 N Complications (Complication) [Pathologic Function]
Meta Mapping (888):
  694 Further [Spatial Concept]
  861 complications (Complication Aspects) [Functional Concept]

emerse
ELECTRONIC MEDICAL RECORD   SEARCH ENGINE

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

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.
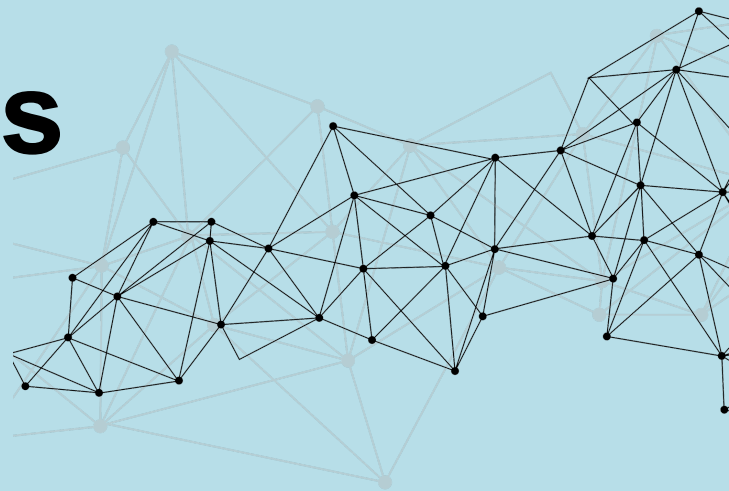
EMERSE is *fast*

| 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

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
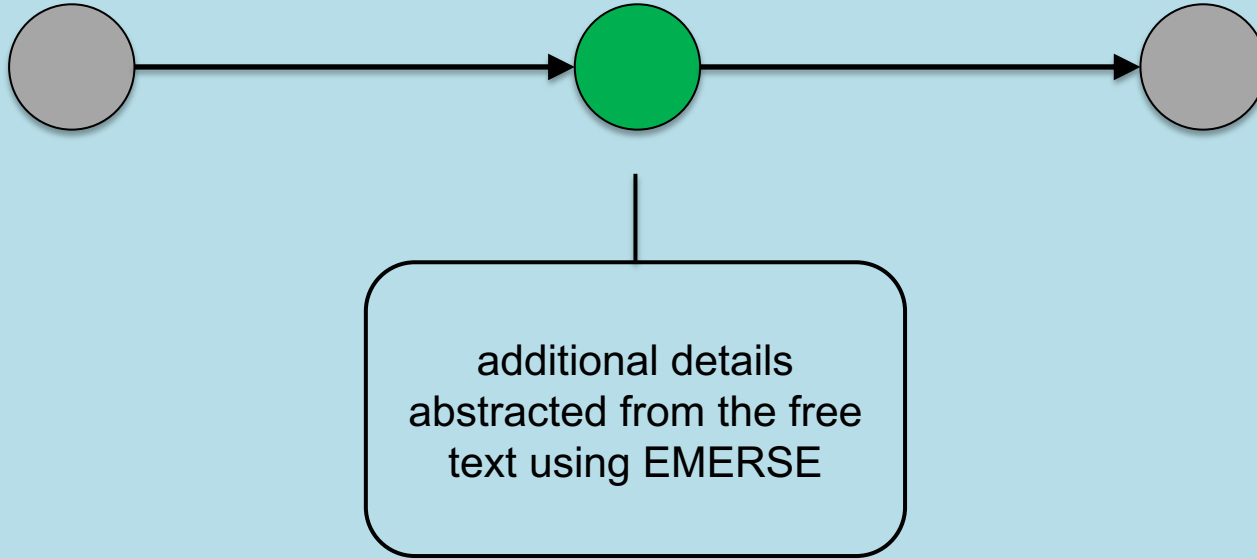 - 1.9 million unique entries (~85% not within UMLS)
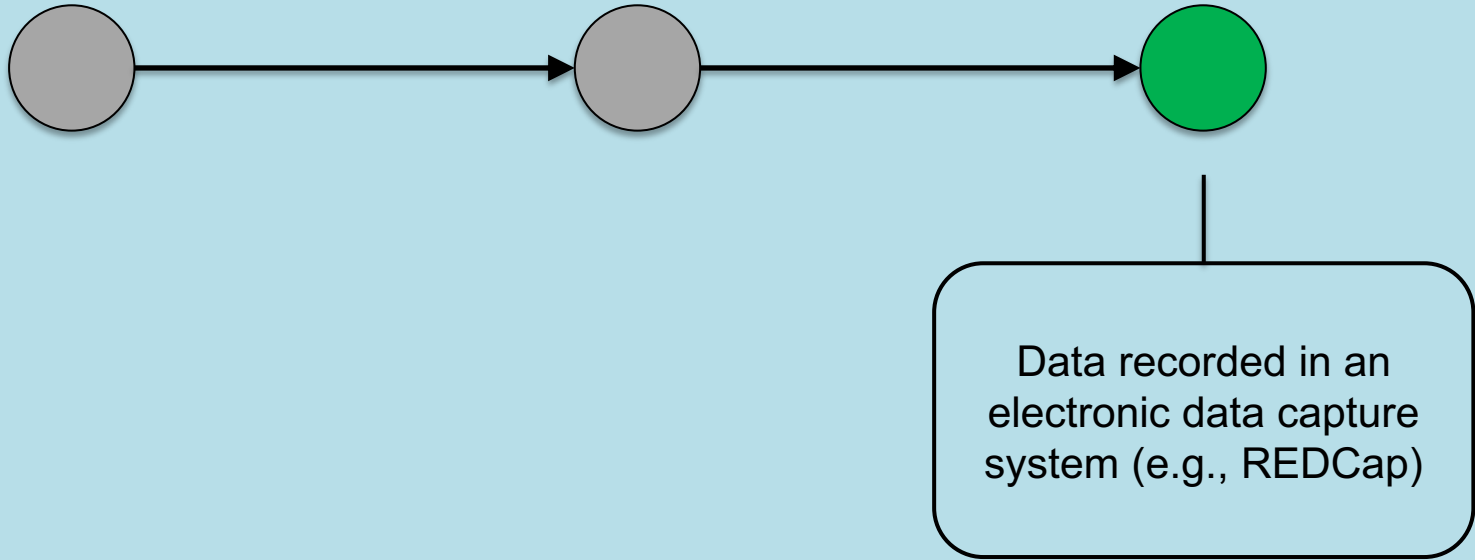
Typical workflow
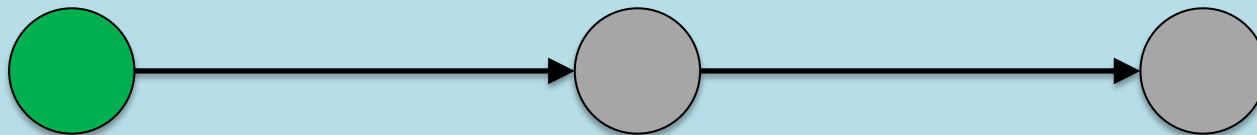
initial cohort identified with a cohort discovery tool using structured data (or via EMERSE 'Find Patients')

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

emerse
ELECTRONIC MEDICAL RECORD SEARCH ENGINE

Typical workflow

additional details abstracted from the free text using EMERSE

Typical workflow

Data recorded in an electronic data capture system (e.g., REDCap)

Publications using EMERSE

605

papers and abstracts

Full list at:

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

Recent papers

- A retrospective cohort study of genetic referral and diagnosis of lynch syndrome in patients with cutaneous sebaceous lesions (36437392)

- Sex differences in adults with acute myeloid leukemia and the impact of sex on overall survival (36419242)

- Clinical characteristics and outcomes in patients with metastatic breast cancer and pseudocirrhosis: a single center retrospective cohort study (36319907)

- Pneumocystis jirovecii Infection in autologous hematopoietic stem cell transplant recipients (36550198)

- Diagnosis of Clival Cancer and Sixth Nerve Palsy (35830685)

Recent papers

- Allostatic load and cardiovascular outcomes in males with prostate cancer (36752520)
  - "All the information obtained was complemented with electronic health record (EHR) information captured via EMERSE (Electronic Medical Record Search Engine) in order to obtain the most accurate and complete information per patient."

Where is EMERSE?

Locations of the sites that have EMERSE installed (📍) or are currently working on installing (📍)

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

Where is EMERSE?

Locations of the sites that have EMERSE installed (📍) or are currently working on installing (📍)

Coming soon?

EMERSE Research Informatics Network

real-time, secure
cross-site queries

emerse
ELECTRONIC MEDICAL RECORD SEARCH ENGINE

http://project-emerse.org/presentations.html
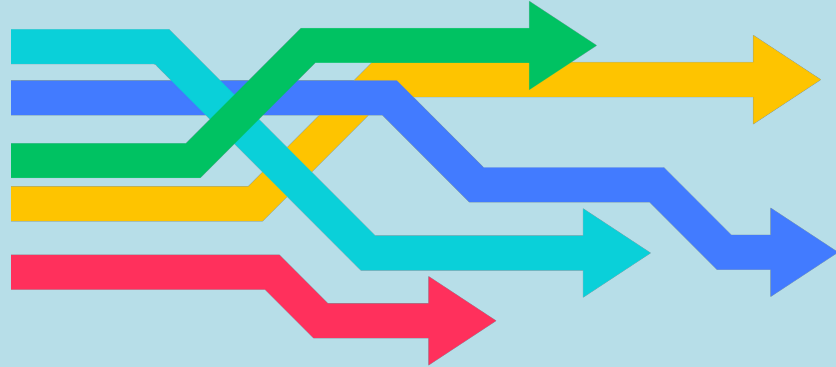
EMERSE Research Informatics Network

The future…

Incorporation of NLP features
- negation
- uncertainty
- subject (patient vs other)
- named entity recognition/mapping to ontologies
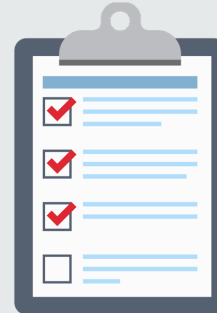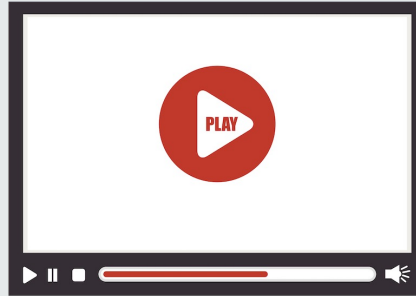
Data extraction from templated notes

(?) Integration with ChatGPT or similar tools

emerse
ELECTRONIC MEDICAL RECORD   SEARCH ENGINE

Ultrasonography displayed an inhomogeneous hypoechoic nodule measuring 20mm*17mm (Figure1). Biopsy examination of the lesion revealed scirrhous carcinoma. A chest computed tomography (CT) scan for metastases showed abnormal shadows in both upper lung fields. The patient was then referred to our department for definitive workup and treatment. She had no history of cough, sputum, or dyspnea. Our patient had no history of tobacco smoking and no exposure to any dusts associated with a high risk of lung damage. Her past history and family history were unremarkable. A chest X-ray showed slight peripheral infiltration shadows in both upper and middle lung fields (Figure2). A chest CT scan showed patchy peripheral ground-glass opacities and thickened interlobular septa in both upper lung fields (Figure3A). A peripheral blood cell count and serum and biochemical tests were normal. Autoantibody and vasculitis screening was negative. Testing for human immunodeficiency virus infection was negative. Serum carcinoembryonic antigen (CEA) and carbohydrate antigen 153 (CA15-3) were in the normal ranges (1.1ng/mL and 12.5U/mL, respectively), but granulocyte-macrophage colony-stimulating factor (GM-CSF) autoantibody was elevated (29.57?g/mL). Pulmonary function testing revealed normal lung volumes and diffusing capacity. Flexible fiberscopic bronchoscopy was then performed. The retrieved bronchoalveolar lavage fluid (BALF) was transparent; it did not have a milky appearance. However, BALF cytology showed alveolar macrophages with granular materials that stained positively with periodic acid–Schiff (PAS). Histological findings of a transbronchial lung biopsy specimen showed the alveolar spaces to be filled with PAS-positive granular materials (Figure4). As a result, this patient was diagnosed as having PAP.

From our NLP proof-of-concept system

project-emerse.org

USER GUIDE

PLAY

emerse
ELECTRONIC MEDICAL RECORD   SEARCH ENGINE

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

Twitter: @projectEMERSE

publications
software releases
announcements
webinars

emerse
ELECTRONIC MEDICAL RECORD SEARCH ENGINE

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

Interested in EMERSE?
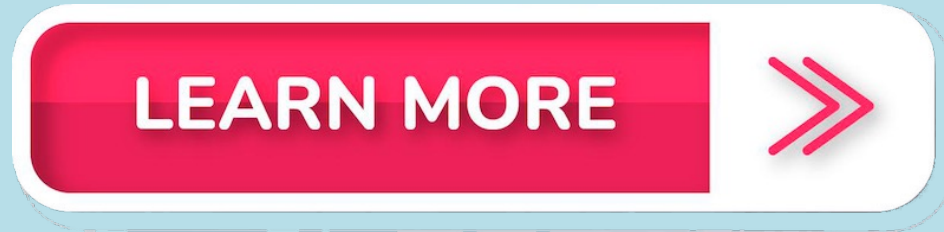
Contact us to schedule a time with your team for:

- Discussions about research strategies

- Training

- Live demonstrations

Tuesday, October 10, 2023, 1-2 PM ET

Register: bit.ly/emerse-october-2023

EMERSE-team@umich.edu

Lisa Ferguson

David Hanauer

Kellen McClain

Guan Wang

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