EMERSE (electronic medical record search engine)

The problem
• Unique phenotypic data are ‘trapped’ in clinical documents (needed for cohorts, eligibility, identifying adverse events, data abstraction, and more)
• NLP tools require NLP expertise

Our solution
• Information retrieval system, tailored for medical documents
• Users are already familiar with searching online

EMERSE
• Built on top of Apache Lucene/Solr
• Developed at U of Michigan for 12+ years
• Easy to use, intuitive, minimal training
• Made for “regular” researchers, not informaticians
• http://project-emerse.org (new site sneak peak: https://goo.gl/EYDiZn)
• Supported > 1,000 studies
• 150 known publications
• Used heavily by our Cancer Center Clinical Trials Office, other cancer researchers (example PMIDs: 22907512, 21841163, 18832652, 24836548, 22665938)
• Very fast, supports API access if needed
• Fits well into research infrastructure with tools such as REDCap, i2b2, etc.
EMERSE (electronic medical record search engine)
EMERSE (electronic medical record search engine)

Planned work supported by ITCR
• Disseminate to other institutions
• Network sites together, at least for cohort discovery
• Develop an adoption ‘roadmap’
• Add new features
• Staggered implementations with partners, learn and update

Our partners
• UNC, Columbia, Cincinnati, Case Western, Kentucky

Challenges
• System must be implemented centrally
• Integration within clinical environment, EHR
• Users must convince high level people (‘C-suite’) to adopt it

Immediate Next Steps
• Updating website, help videos, documentation
• Code/security reviews
• Looking for ITCR collaborators

Contact
• David Hanauer, hanauer@umich.edu