

## Empower Data-based Clinical Trial Protocols for Patient-to-Trial Matching via Artificial Intelligence (AI)

Are you an engineer or data scientist with expertise in Al/natural language processing? Interested in helping patients find clinical trials?

## Goal: Use machine learning to find patterns of similar eligibility protocol criteria for clinical trials and guide the development of data-based clinical trial protocols for patient-to-trial matching.

Less than 5% of patients eligible for clinical trials register, resulting in higher drug development costs, drugs based on limited/overrepresented populations, and longer wait times. Making clinical trial search more effective can potentially decrease barriers to enrollment.

The writing of clinical trial protocols -- the registration information describing a trial, its goals, the plan for carrying out the study, and who is eligible – historically has been written as free text independently for each trial. As a result, this text is not easily read by a computer. Searching across trials or linking data (like linking patients to eligible trials, or trials to eligible patients) becomes extremely problematic and error prone.

One can make clinical trial searching and matching more successful by helping machines better understand the information that describes who is eligible for a clinical trial.

## The Hackathon

The hackathon takes a first step in searching for patterns across trials that can be used to facilitate databased clinical trial protocol standards and templates. In this hackathon, the focus will be on leveraging data from the Experimental Therapeutics Clinical Trials Network and a corresponding set of cancer trial eligibility criteria that have been curated based on a couple approaches. In addition, free text will be available on the cancer trials via NCI's Clinical Trials Reporting Program (CTRP) Application Programming Interface (API). Through these resources, clinical trial eligibility criteria can be clustered via unsupervised learning to find patterns and areas of consensus in text, logical expressions, and semantics across trial criteria. This clustering will enable for templates and consensus clusters to naturally emerge as potential evidence for standardizing eligibility sections of clinical trial protocols in the future, with the goal of improving AI/NLP performance on matching patients with trials (and vice versa).

National Institutes of Health Data Hackathon

July 23-25, 2018 (9-5p) Board of Regents Room. Building 38, NIH Campus Bethesda, MD

https://github.com/NCBI-Hackathons/CTEligible

Note this event is co-located, but also separate from, the following event on related topic also on July 23, 2018: <u>https://www.nih.gov/news-events/news-releases/nih-host-workshop-using-artificial-intelligence-machine-learning-advance-biomedical-research</u>