
SEAN FINAN
HARRY HOCHHEISER
GUERGANA SAVOVA
JEREMY WARNER

[MENTOR'S LINKEDIN URL/
HTTPS://WWW.LINKEDIN.COM/IN/SEAN-
FINAN-B028401/](https://www.linkedin.com/in/sean-finan-b028401/)
[HTTPS://WWW.LINKEDIN.COM/IN/HARRY-
HOCHHEISER-5B014B/](https://www.linkedin.com/in/harry-hochheiser-5b014b/)
[HTTPS://WWW.LINKEDIN.COM/IN/GUERGAN
A-SAVOVA-B1060611/](https://www.linkedin.com/in/guergana-savova-b1060611/)
[HTTPS://WWW.LINKEDIN.COM/IN/JEREMY-
WARNER-39001512/](https://www.linkedin.com/in/jeremy-warner-39001512/)

Required Skills:

Mobile App Development, Web App Development, Human Centered Design, Workflow/Process Optimization, Project Management, Technical Writing, Graph and Relational Databases, HL7 FHIR, familiarity with ontologic representations

Preferred Team Communications:

WEBEX, Skype or Conference call

Data Sources:

Synthetic patient data will be provided.

Other Items:

Project has time zone flexibility. Mentors and students will determine a good time for virtual meeting; students will be asked to join weekly scrum meetings if able.

DEEPPHE: A NATURAL LANGUAGE PROCESSING SYSTEM FOR EXTRACTING CANCER PHENOTYPES FROM CLINICAL RECORDS

Precise phenotype information is needed to understand the effects of genetic and epigenetic changes on tumor behavior and responsiveness. Extraction and representation of cancer phenotypes is currently mostly performed manually, making it difficult and costly to correlate phenotypic data to genomic data. In addition, genomic data are being produced at an increasingly faster pace, exacerbating the problem. The DeepPhe software enables automated extraction of detailed phenotype information from electronic medical records of cancer patients. The system implements advanced Natural Language Processing and knowledge engineering methods within a flexible modular architecture, and has been evaluated using a manually annotated dataset of the University of Pittsburgh Medical Center breast cancer patients. The resulting platform provides critical and missing computational methods for computational phenotyping. Working in tandem with advanced analysis of high-throughput sequencing, these approaches will further accelerate the transition to precision cancer treatment.

More info:

https://healthnlp.hms.harvard.edu/deepphe/wiki/index.php/Main_Page

PROJECT OBJECTIVES

1. Upgrade the final output of DeepPhe from FHIR DSTU2 to FHIR STU3 specifications.
 2. Develop and/or augment DeepPhe extensions to FHIR to represent concepts such as complex chemotherapy regimen protocols.
 3. Modify DeepPhe output to create a post-rules patient-level summary as well as document-level summaries
-

SUCCESSFUL PROJECT

1. A solution would be a functional FHIR server that can store and expose DeepPhe data, and potentially an app that can take advantage of such data structures for one or more use cases, such as organizing various cancer treatments into a clear timeline.
-

Intellectual Property: Project involves a government agency so the resulting project is made available to the public. Students do not own IP. Students will be recognized as contributors