

Build a Knowledge Base for the Human Body

The proposal period for 2022 internships is now closed
The proposal period for 2023 internships will open in November 2022

This is new project, more information coming soon. If you are interested in this project contact [Lorraine Chapman](#).

Find out about the [HPCC Systems Summer Internship Program](#).

Project Description

Create a knowledge base using NLP++ and the conceptual grammar for the human body. This includes a "part-of" hierarchy for the human body, an accompanying vocabulary. This will allow for processing of medical texts.

If you are interested in this project, please contact [Add email link to mentor](#).

Completion of this project involves:

- Become familiar with NLP++ and VisualText
- Research and find medical data that can be used to create a KB and dictionary
- Implement NLP++ parsers to ingest and create a KB and dictionary for the human body
- Run the KB and dictionary on medical texts using the NLP++ ECL plugin
- Create a VisualText repository with the Human Body KB and dictionary

By the mid term review we would expect you to have:

- <What must be completed to pass the evaluation and continue on to complete the project>

Mentor	David de Hilster david.dehilster@lexisnexisrisk.com Backup Mentor: TBD
Skills needed	<ul style="list-style-type: none">• Keen interest in natural language• Ability to do research on the internet• Ability to learn and program in NLP++• Ability to write test code in ECL using the NLP++ plugin to test the KB and dictionary
Deliverables	Midterm <ul style="list-style-type: none">• <Deliverable(s) to be achieved> End of project <ul style="list-style-type: none">• <Deliverables expected by the end of the internship>
Other resources	<ul style="list-style-type: none">• HPCC Systems website• JIRA issue for this project:• Blog: Understanding Natural Language Processing• Github Repository• Video: Deploying Digital Human Readers Leveraging HPCC Systems <p>Video: NLP++ ECL Plugin</p> <ul style="list-style-type: none">• Visual Text Open Source Website• NLP++ Language Extension• Formal language description• Learning ECL documentation and on-line training courses.