## January 19, 2026

# Health Care Applications with Natural Language Processing

Unstructured documents often come with embedded structured data. Representing valuable and structured information as tables is popular in health, financial, and many domains. However, manual extraction of structured information from documents typically costs tremendous time and labor, motivating the need for a system for automating the process. After such tables have been extracted, the data can be used for a wide variety of tasks such as question answering and various "down-stream" analytics tasks. In this talk, we will discuss how to leverage ground breaking pre-trained language models (e.g., BERT, ChatGPT) to develop tools for automated table extraction from various types of documents. We will present different applications from cancer registry reporting, cancer care, and psychiatry hospitalization prediction.

**Presenter**	**Raymond Ng**

Director, Data Science Institute and Professor, Department of Medicine and Department of Computer Science

C2E2 Spotlight is held monthly on Mondays from 12:00 pm to 1:00 pm. To attend in-person or virtually please contact pamela.lee@ubc.ca.