# CS 5740 Homework 1 (additional component)

This is an additional component, meant **ONLY** for students enrolled in a graduate version of the course, **CS 5740**, and must be worked on individually (not in teams). If you are enrolled in CS 4740, **DO NOT** complete this section.

## Nested Named Entity Recognition

In this part of the homework assignment we hope for you to learn and think about a task of Nested Named Entity Recognition.

Frequently, named entities may have a complicated structure, making it difficult to identify tokens that need to be labeled and labels that need to be assigned. For instance, consider the sentence:

Cornell University professor of Computer Science developed a new course.

In this case, “Cornell University” is an entity, but so is “Cornell University professor of Computer Science”.

## Task

Please read the following survey paper of research done in the area of Nested Named Entity Recognition:

Wang, Yu, et al. "Nested Named Entity Recognition: A Survey." ACM Transactions on Knowledge Discovery from Data (TKDD) (2022). [PDF link](https://dl.acm.org/doi/pdf/10.1145/3522593?casa_token=6qJWjkaaQBAAAAAA:O3UlJUP-SCNLSIjW7TNmvkS9xUaisTTKTT7nALY9bPh1eVkNgJdjrNOqTHChA_whFQ6qrJXULOdFMg)

**Note**: you may skip section 4.2 of the paper.

Having read the paper, please make a local copy of this document and write up answers for questions provided further.

## Submission

This is an **individual** task (do *not* work with your partner). You will be asked to save your version of the document in **PDF** format and upload it to a grading platform (details TBA).

This component is due at the same time as the general HW1 submission: **September 27, 2022 (11:59 PM).**

### 

### **Q6.1**:

What is the Nested Named Entity Recognition task and what are its key differences from the traditional NER we have seen in class and throughout this assignment? What makes nested NER a more challenging task?

#### A6.1

... add your answers here

### **Q6.2**:

How do you think the models you trained in Parts 2 & 3 of HW1 would behave when given sentences with nested named entities? Come up with your own 2-3 examples of nested named entities, test variations of these sentences against your models, reflect on your observations, and compare to your initial hypothesis. What kinds of entities get picked up and what information is missed?

#### A6.2

... add your answers here

### **Q6.3**:

Choose any 2 approaches from the paper that are commonly used to capture nested entities and describe them.

#### A6.3

... add your answers here