

Assignment 3: Question 3

Software Innovation Institute, ANU

Named Entity Recognition (NER)

Identify text strings that correspond to named entities. Such as PERSON, or GPE (geopolitical entity).

Donald Knuth PERSON was born in America GPE , to Ervin Henry Knuth PERSON and Louise Marie Bohning PERSON .

This is a challenging task but can be done reasonably well with existing methods.

Try out NER: <https://explosion.ai/demos/displacy-ent>

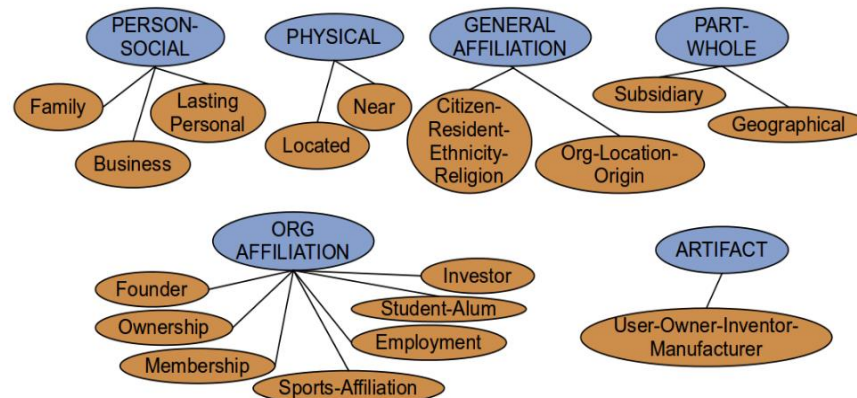
Relations

brother(Wilbur Wright, Orville Wright)

Relations

Entities

born-in(Wilbur Wright, 1867)



Relation Extraction

Extract a relation from the text:

Donald Knuth PERSON was born in America GPE , to Ervin Henry Knuth PERSON and Louise Marie Bohning PERSON .



nationality(Donald Knuth, America)

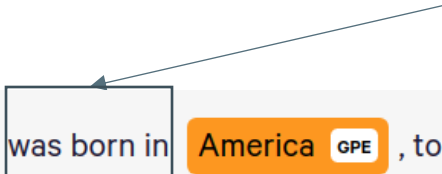
Relation Extraction: Classification

- Identify entities in text
- Decide what relation (if any) holds between the two entities
 - Can use a standard classifier (logistic regression, MLP)

Relation extraction features

Use the words surrounding the named entities to identify the relation.

For example, you could extract this and use it to construct features.



Donald Knuth PERSON was born in America GPE , to Ervin Henry Knuth PERSON and Louise Marie Bohning PERSON .

Relation extraction features

If you want it to work on new entities **don't use the entities themselves** as features.

You can extract features from the entity. For example: what cluster the word occurs in, how long the word is, capitalization etc..

Relation extraction features

More advanced features

Dependency parse features: path
between the named entities

