

## 1. Elegir temática

- (a) Enfermedad/Área  
Pharma
- (b) Tarea  
Name entity recognition
- (c) Obtención de datos  
Pubmed

## 2. Elegir modelo

- (a) Lista de modelos
  - BERT
  - GPT
  - LLAMA
  - CNN
  - LSTM
- (b) Pros y contras

## 3. Preprocesamiento

- (a) Tratar el corpus
  - Limpiar
- (b) Etiquetar entidades manualmente (muy costoso)
  - Bio encoder (or alternatives)

## 4. Construcción (sólo idea)

## 5. Conclusiones