

Natural Language Processing

Introduction to natural language processing

Marco Kuhlmann

Department of Computer and Information Science

What is natural language processing?

- **Natural language processing (NLP)** enables computers to work with human language.
- Some well-known example applications are web search, machine translation, and chatbots.
- These diverse applications are based on a common set of ideas from algorithms, machine learning, and other disciplines.

Question answering

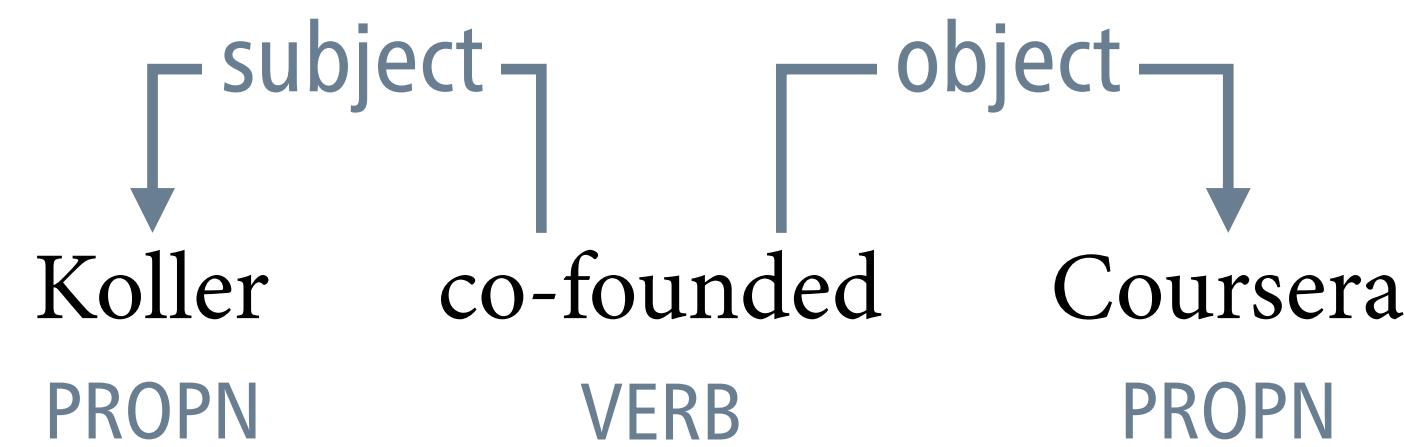
Which Stanford University alumna co-founded the educational technology company Coursera?



[SPARQL query against DBpedia](#)

```
SELECT DISTINCT ?x WHERE {  
  ?x dbp:education dbr:Stanford_University.  
  dbr:Coursera dbp:founder ?x.  
}
```

Linguistic representations



dbr:Coursera dbo:founder dbr:Daphne_Koller

End-to-end neural networks

MA

You

Which Stanford University alumna co-founded the educational technology company Coursera?



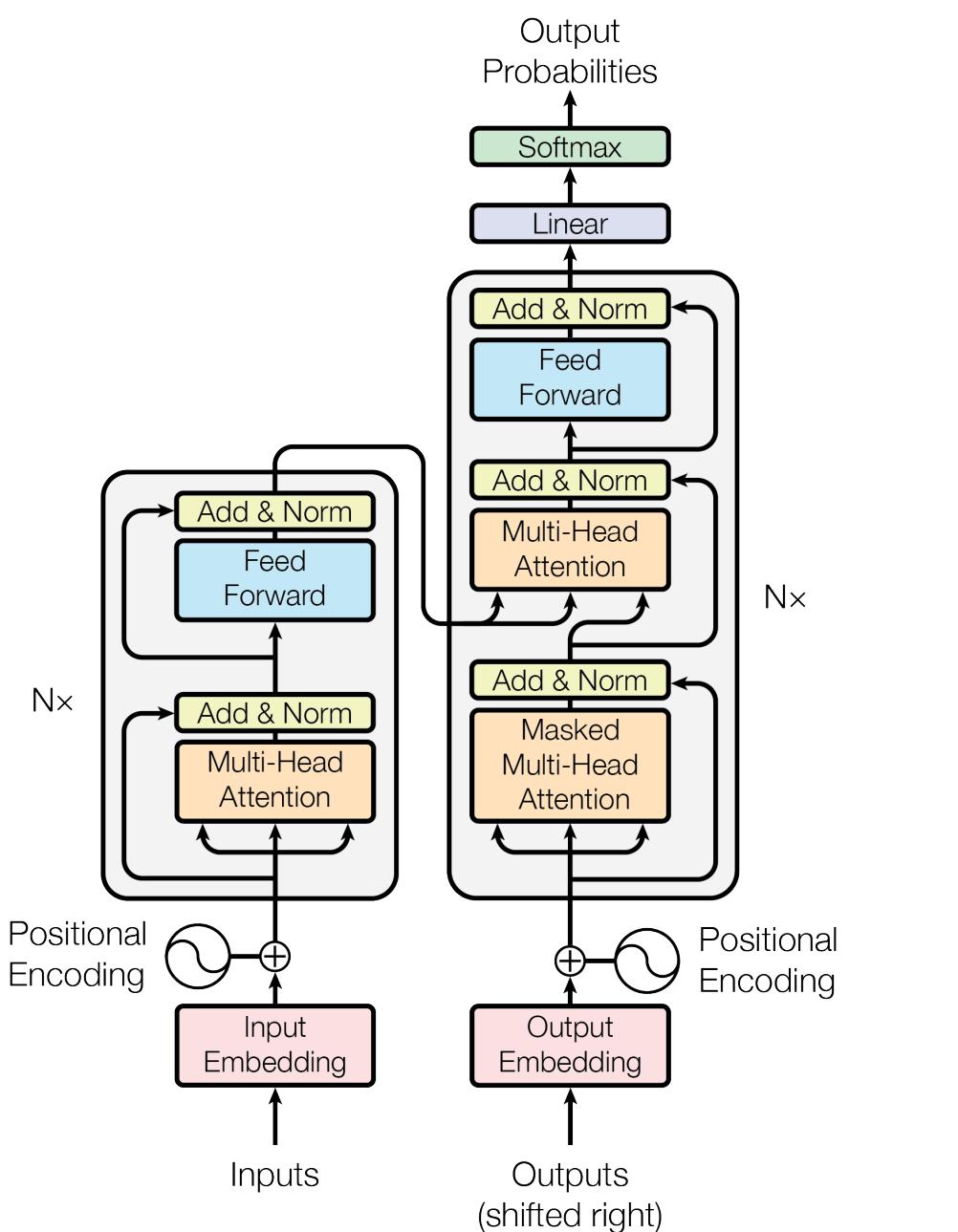
ChatGPT

The Stanford University alumna who co-founded the educational technology company Coursera is Daphne Koller. She co-founded Coursera in 2012 alongside Andrew Ng, another Stanford professor, with the goal of providing universal access to world-class education.

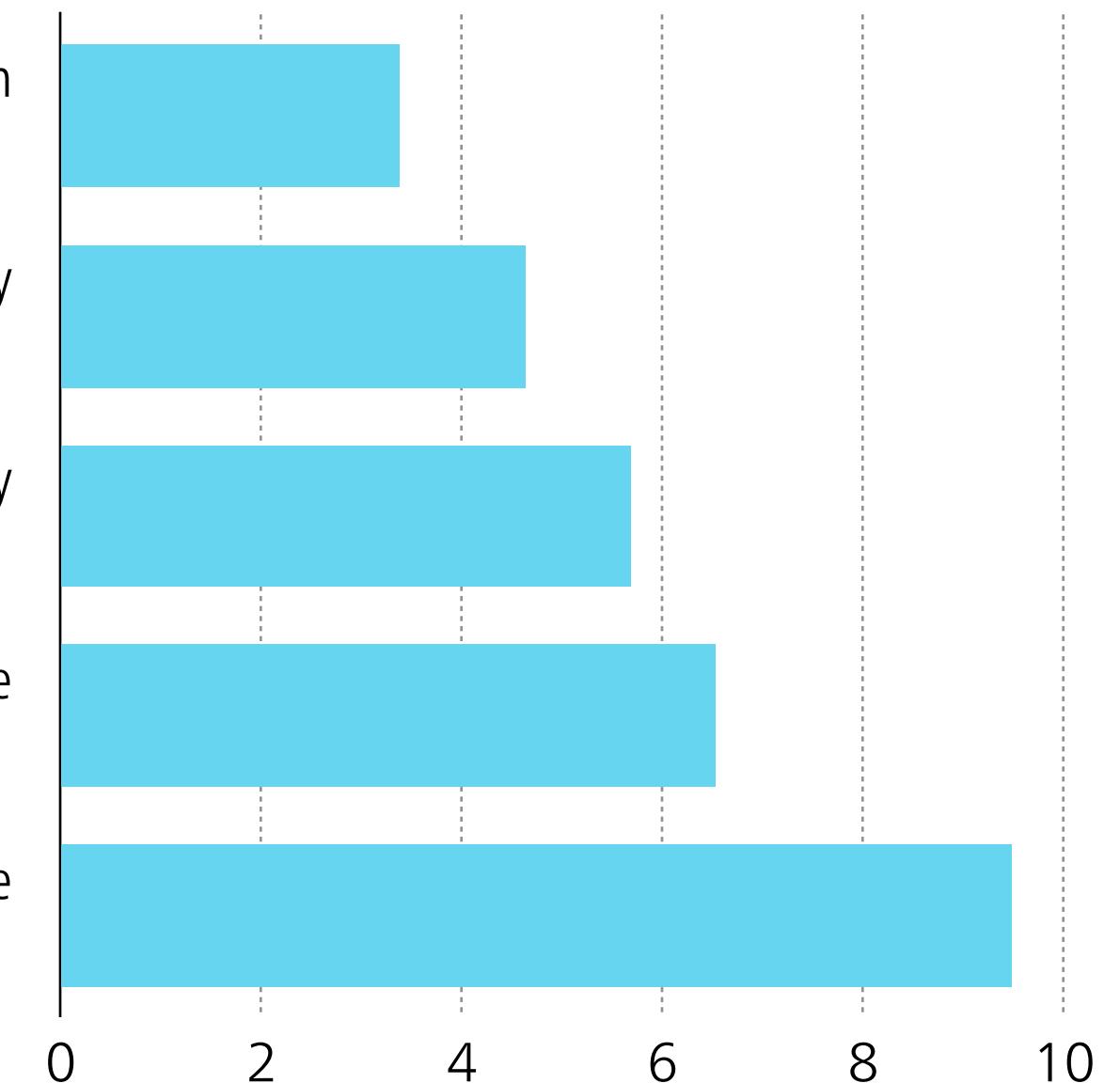


ChatGPT 4, 2023-12-31

“Natural language processing from scratch”



Part-of-speech
tagging
Named entity
recognition
Dependency
parsing
Semantic role
labelling
Coreference
resolution



[Vaswani et al. \(2017\)](#)

[Tenney et al. \(2019\)](#)

MA

You

Convert the following question into a SPARQL query against DBpedia: "Which Stanford University alumna co-founded the educational technology company Coursera?"



ChatGPT

To convert your question into a SPARQL query for querying DBpedia, we need to construct a query that looks for a person who is an alumna of Stanford University and a co-founder of the educational technology company Coursera. Here's how the query would look:

sparql

Copy code

```
PREFIX dbo: <http://dbpedia.org/ontology/>
PREFIX dbr: <http://dbpedia.org/resource/>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>

SELECT ?person WHERE {
  ?person dbo:almaMater dbr:Stanford_University .
  ?person dbo:occupation dbr:Entrepreneur .
  ?person dbo:knownFor dbr:Coursera .
  ?person a dbo:Person .
  ?person dbo:gender dbr:Female .
} LIMIT 10
```

Two paradigms

- **Linguistic knowledge**

Build pipelines of modular components that produce general-purpose representations grounded in linguistic knowledge.

morphemes, parts-of-speech, dependency trees, meaning representations

- **Deep learning**

Train end-to-end neural networks that directly transmute raw text into whatever form the desired application requires.

text, code, images

Widespread adoption of chatbots

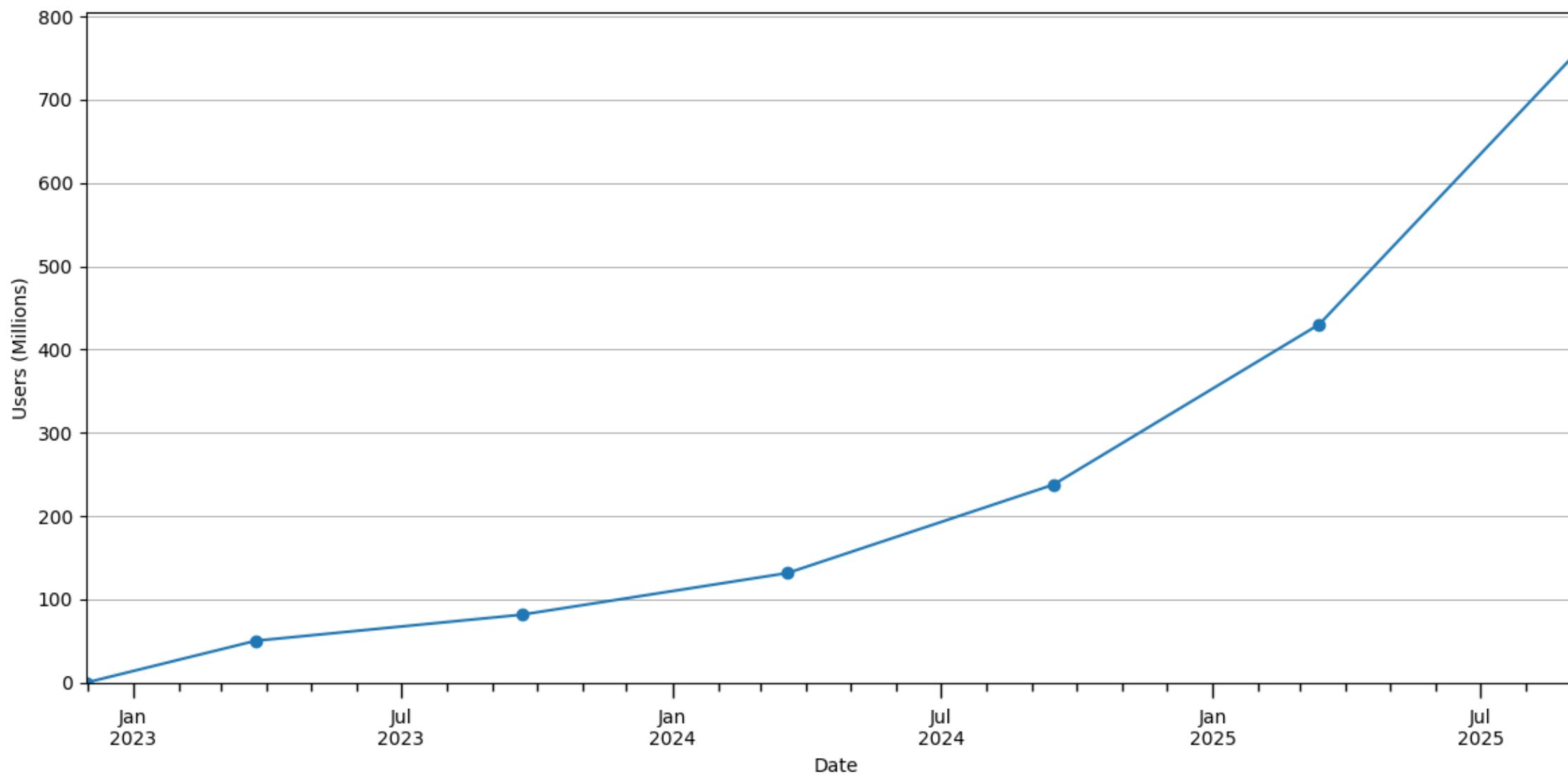


Figure 3: Weekly active ChatGPT users on consumer plans (Free, Plus, Pro), shown as point-in-time snapshots every six months, November 2022–September 2025.

[Chatterji et al. \(2025\)](#)

Brittleness of chatbots

