

# SIT330-770: Natural Language Processing

## Week 0 - Course Overview

**Dr. Mohamed Reda Bouadjenek**

School of Information Technology, Faculty of  
Sci Eng & Built Env

[reda.bouadjenek@deakin.edu.au](mailto:reda.bouadjenek@deakin.edu.au)



**DEAKIN**  
UNIVERSITY



- Introduction to SIT<sub>330-770</sub>
  - Unit team
  - Unit objectives and Introduction to NLP
  - Unit structure
  - Teaching and Assessment
  - Unit logistics





- Unit chair: *Mohamed Reda Bouadjenek*
  - Email: [reda.bouadjenek@deakin.edu.au](mailto:reda.bouadjenek@deakin.edu.au)



- Burwood Campus Coordinator: *Bahadorreza Ofoghi*
  - Email: [b.ofoghi@deakin.edu.au](mailto:b.ofoghi@deakin.edu.au)



- Teaching Assistants: *Aymen Khouas*
  - Email: [s223766761@deakin.edu.au](mailto:s223766761@deakin.edu.au)



- Dr. Reda Bouadjenek <reda.bouadjenek@deakin.edu.au>
  - Unit chair, Waurin Ponds
- Senior Lecturer of Applied AI at Deakin University since November 2019
  - Interested in information Retrieval, Natural Language Processing, social media analysis, recommender systems
- Previously:
  - **2017-2019**: Research fellow at Toronto University, Canada
  - **2015-2017**: Research fellow at Melbourne University
  - **2014-2015**: Research fellow at inria, France
  - **2009-2013**: PhD student at the University of Paris-Saclay, France



# Understanding principals behind Natural Language Processing



# What is natural language processing?

The study of language and linguistic interactions from a computational perspective, enabling the development of algorithms and models capable of (a) **natural language understanding (NLU)** and (b) **natural language generation (NLG)**





# Go beyond the keyword matching



- Identify the **structure** and **meaning** of **words**, **sentences**, **texts** and **conversations**
- **Deep** understanding of **broad** language
- NLP is all around us



# Machine translation

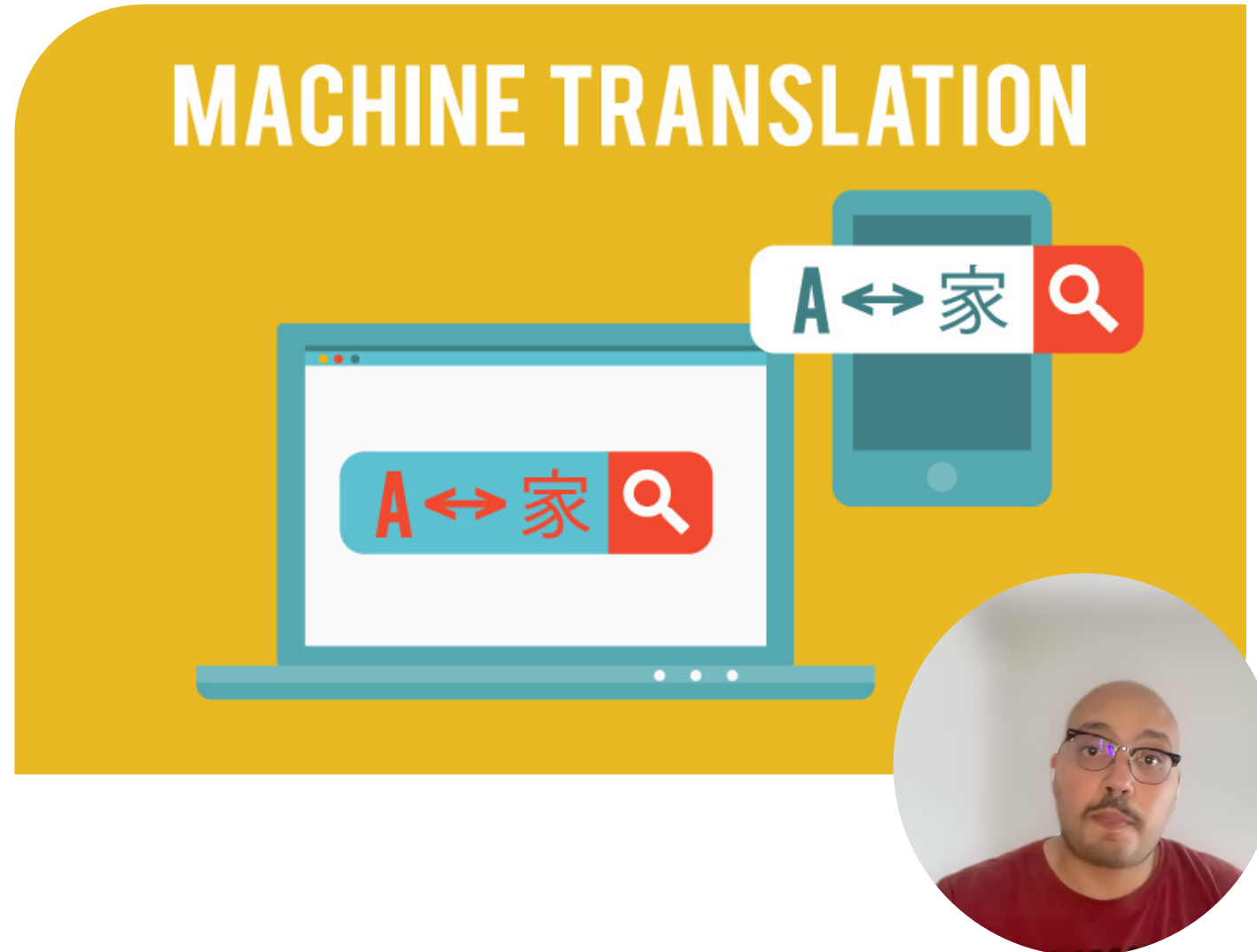
 **Alessandro Del Piero**   
June 7, 2023 · 

Ho incontrato Luis Figo per selezionare una ALLSTAR Team di calcio a 5 basata sulle caratteristiche dei Transformers [#TransformersIlRisveglio](#), dal 7 giugno al cinema [#ad](#)

I met Luis Figo to select an ALLSTAR 5-a-side football team based on the characteristics of Transformers [#TransformersTheAwakening](#), in theaters June 7 [#ad](#)

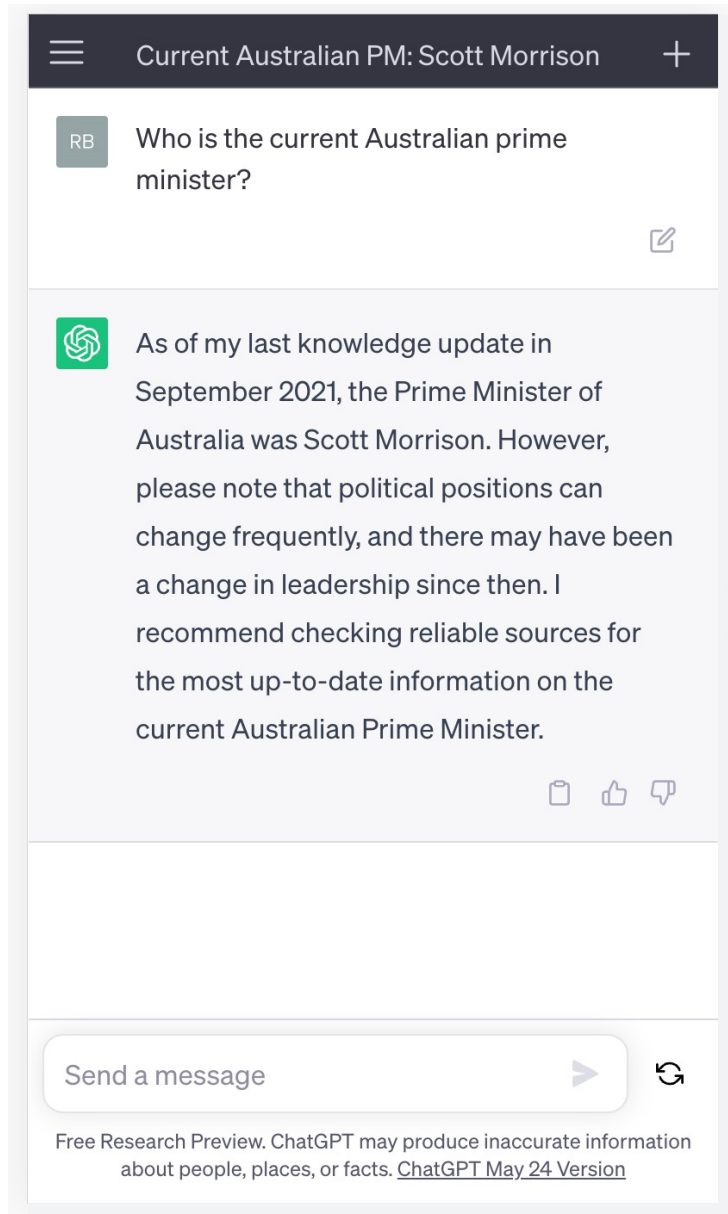
 · [Hide Translation](#) · [Rate this translation](#)







# Conversational System



# ChatGPT



# Sentiment/Opinion Analysis



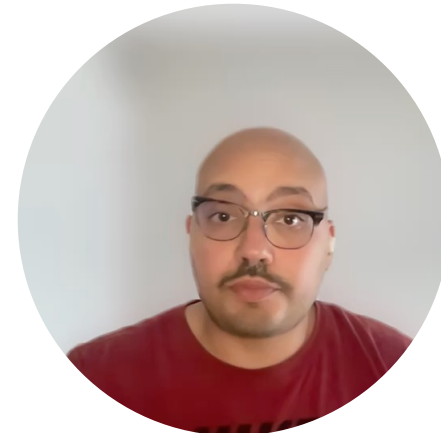
Negative



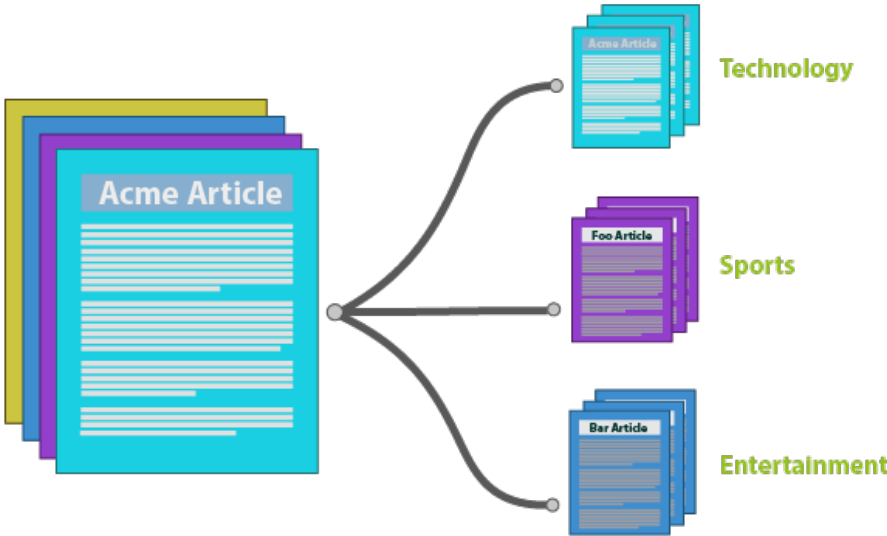
Neutral



Positive



# Text Classification



Primary

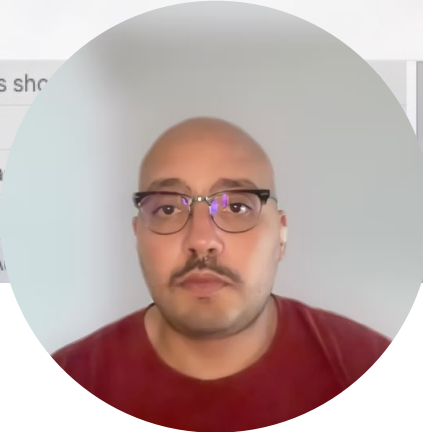
Promotions

Social

Updates

Forums

<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Short paper	SIGIR'24 submission 8282 update - Dear authors, we acknowledge that we received new files for your SIGIR'24 submission. The information about this update is sh
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Short paper 2	SIGIR'24 submission 8300 update - paper, version 2 (474368 bytes)
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Short paper 2	SIGIR 2024 short papers: authorship, COIs and anonymity - Dear Mohamed Reda Bouadjenek, We look forward to your SIGIR 2024 short paper submission, dea
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Resource paper & Re.	SIGIR'24 submission 3233 - Dear authors, We received your submission to SIGIR'24 (The 47th International ACM SIGIR Conference on Research and Developme
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	IJCNN 2024 4	[IJCNN 2024] Information about paper #1570994543 (A Mask-Based Prior Knowledge Dissemination Method for Deep Neural Networks) has been changed - A



# Natural language instruction

Will it rain tomorrow?

Set an alarm at 8am

Play music by Bruno Mars

How many teaspoons  
are in a tablespoon?

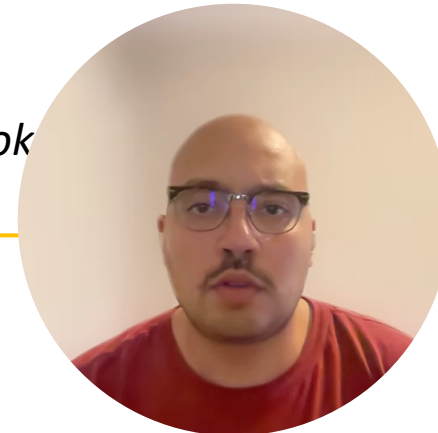
Add soap to my  
shopping list

Play my \*dinner party\* playlist

When is Australia Day?

What is the weather in  
Melbourne this Weekend?

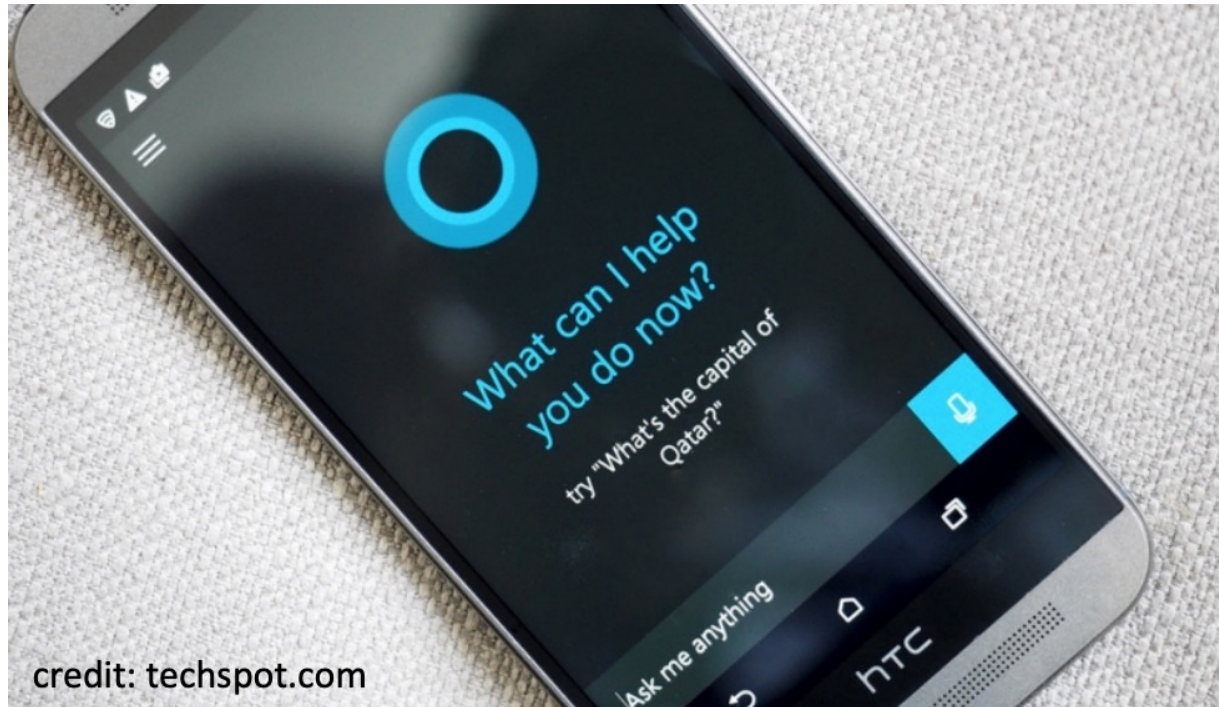
Add \*make book  
my todo list





# Digital personal assistant

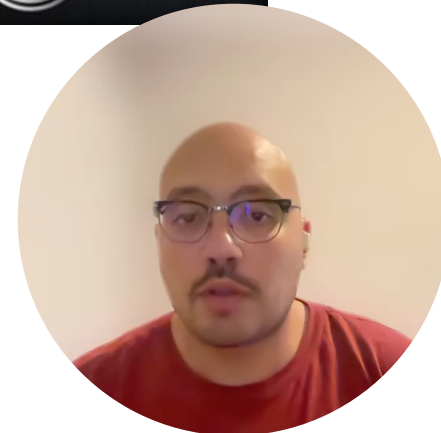
More on natural language instruction



credit: techspot.com



- Semantic parsing – understand tasks
- Entity linking – “my wife” = “Kellie” in the phone book

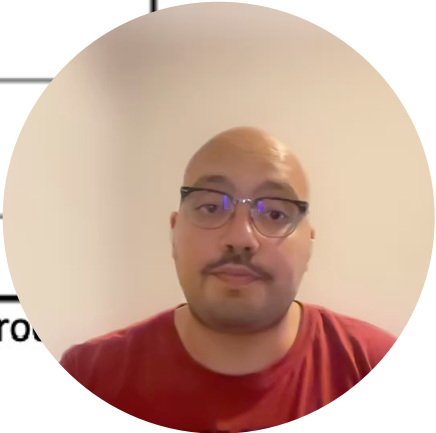


- Unstructured text to database entries

New York Times Co. named Russell T. Lewis, 45, president and general manager of its flagship New York Times newspaper, responsible for all business-side activities. He was executive vice president and deputy general manager. He succeeds Lance R. Primis, who in September was named president and chief operating officer of the parent.

Person	Company	Post	State
Russell T. Lewis	New York Times newspaper	president and general manager	start
Russell T. Lewis	New York Times newspaper	executive vice president	end
Lance R. Primis	New York Times Co.	president and CEO	start

Yoav Artzi: Natural language processing





## First part

- **Week 1:** Information Retrieval Part 1
- **Week 2:** Information Retrieval Part 2
- **Week 3:** Text processing
- **Week 4:** N-gram Language Models
- **Intra-trimester break**

## Second part

- **Week 5:** Naïve Bayes and Sentiment Classification
- **Week 6:** Vector Embeddings
- **Week 7:** Neural Networks and Neural LMs
- **Week 8:** Sequence Labeling
- **Week 9:** RNNs and LSTMs
- **Week 10:** Transformers and Pretrained LMs
- **Week 11:** Wrap-up



# Teaching: Engaging Learning Experience

## What to Expect:

- **Active Learning:** Get ready for an interactive and hands-on learning experience.
- **Flipped Classroom Model:** We're flipping the traditional classroom! Pre-class materials will be provided for self-study before our sessions.
- **Your Role:** Come prepared to discuss, question, and apply concepts during class time.
- **Collaborative Environment:** Engage with your peers, share ideas, and actively participate.



- OnTrack learning portfolio 100% <https://ontrack.deakin.edu.au/>
  - An introduction video is available on DeakinCloud
- You set your target grade (P/C/D/HD)
  - Several tasks (tailored to your target grade) need to be completed regularly
  - Tasks have different due dates
- Weekly submissions are to get feedback
  - **Maximum two feedbacks. Please review carefully your assignment before submitting it!**
- Portfolio submission at the end of the trimester is **very important!**
  - **Assessments will be published gradually during the trimester!**
- **Please do not put assessments on public repositories (GitHub)!**



- **SIT330**

- Session 1:
  - Time: Monday 15:00-17:50
  - Location : Burwood B1.29
- Session 2:
  - Time: Monday 18:00-19:50
  - Location : Online on MS Teams

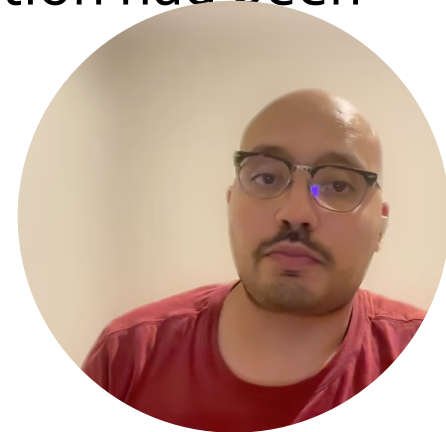
- **SIT770**

- Session 1:
  - Time: Monday 9:00-11:50
  - Location : Burwood HE2.007
- Session 2:
  - Time: Wednesday 17:00-18:50
  - Location : Waurin Pond KA4.412
- Session 3:
  - Time: Friday 18:00-19:50
  - Location : Online on MS Teams

**Do not create a calendar meeting!**  
**Weekly material will be gradually released!**



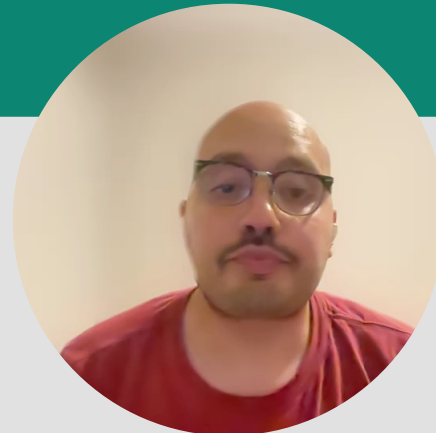
- Please ask questions on the discussion forum to share with others and avoid redundant inquiries
- **Please use email only if necessary**
  - You must start your email subject with “[SIT330] or [SIT770]”
  - I will normally answer your email within 48 hrs (working days)
  - If you don’t receive a reply within 48 hrs, this often means that your question had been answered somewhere else before (e.g., in the class, forum, etc.)
  - If in doubt, check with me again during allocated contact time



This course requires you to write substantial code in Python to achieve a **C, D, or HD grade!**



The unit does not teach Python!





- Graduate Learning Outcomes (GLOs)
- Unit Learning Outcomes (ULOs)
- Teaching team contact details
- Learning activities
  - Class and workshop
- Reference resources
- Trimester plan
- Key dates

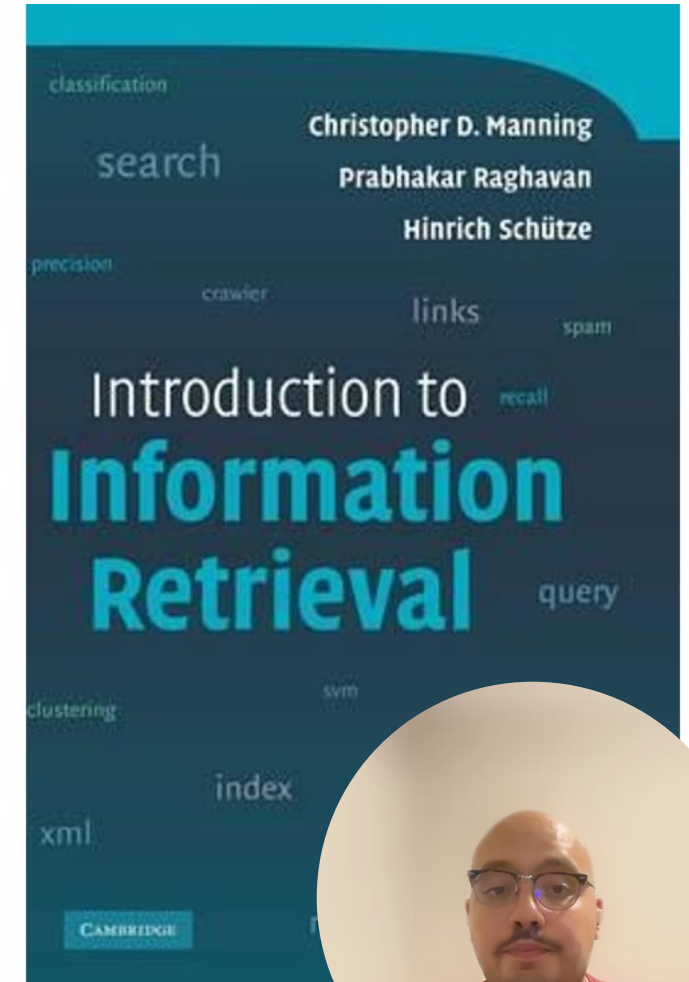


- **Introduction to Information Retrieval**

- *Christopher D. Manning, Prabhakar Raghavan, and Hinrich Schütze*
- Draft chapters, April 1, 2009



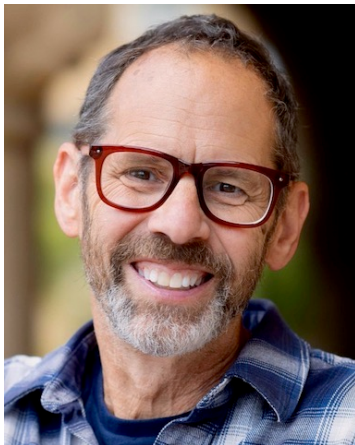
- <https://nlp.stanford.edu/IR-book/pdf/irbookonlinereading.pdf>



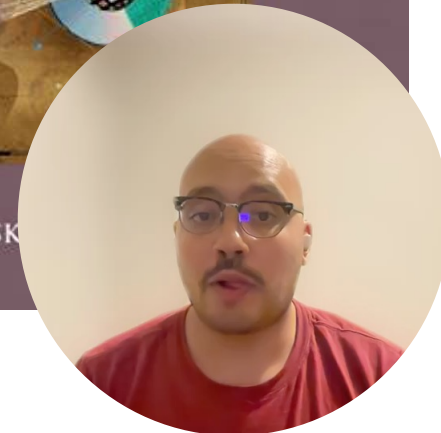
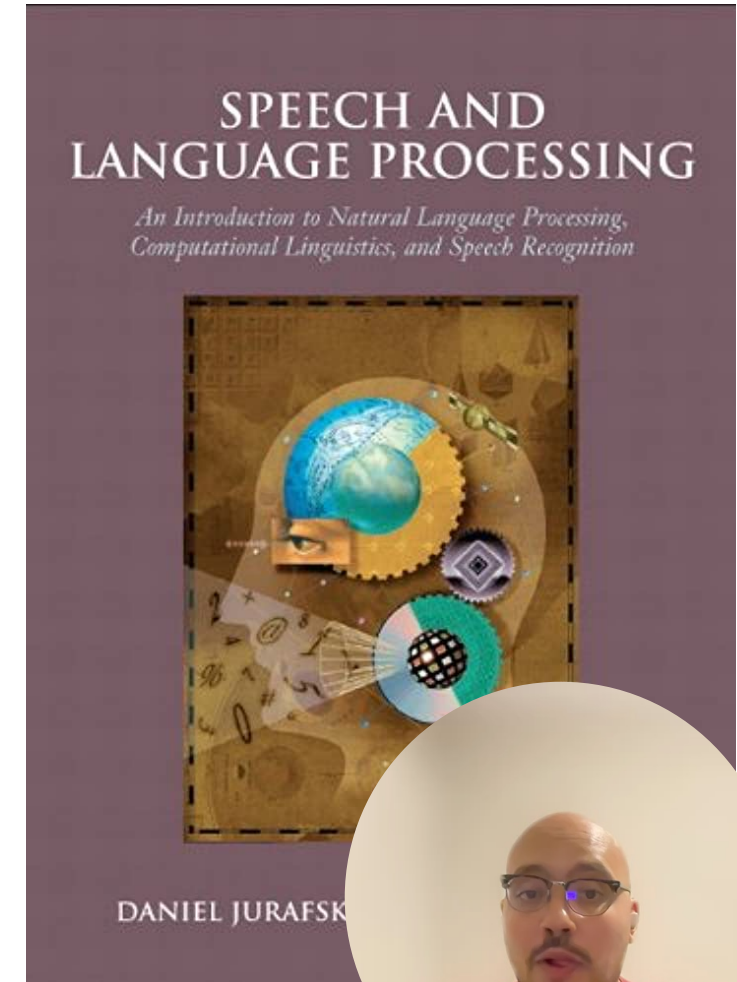
- **Speech and Language Processing** (3rd ed. draft)

- *Dan Jurafsky and James H. Martin*

- Draft chapters in progress, January 7, 2023



- <https://web.stanford.edu/~jurafsky/slp3/>



- SIT744- Deep Learning
  - Explain deep learning and its role in data science and AI





