26.10.22

OPENCAMPUS.sh

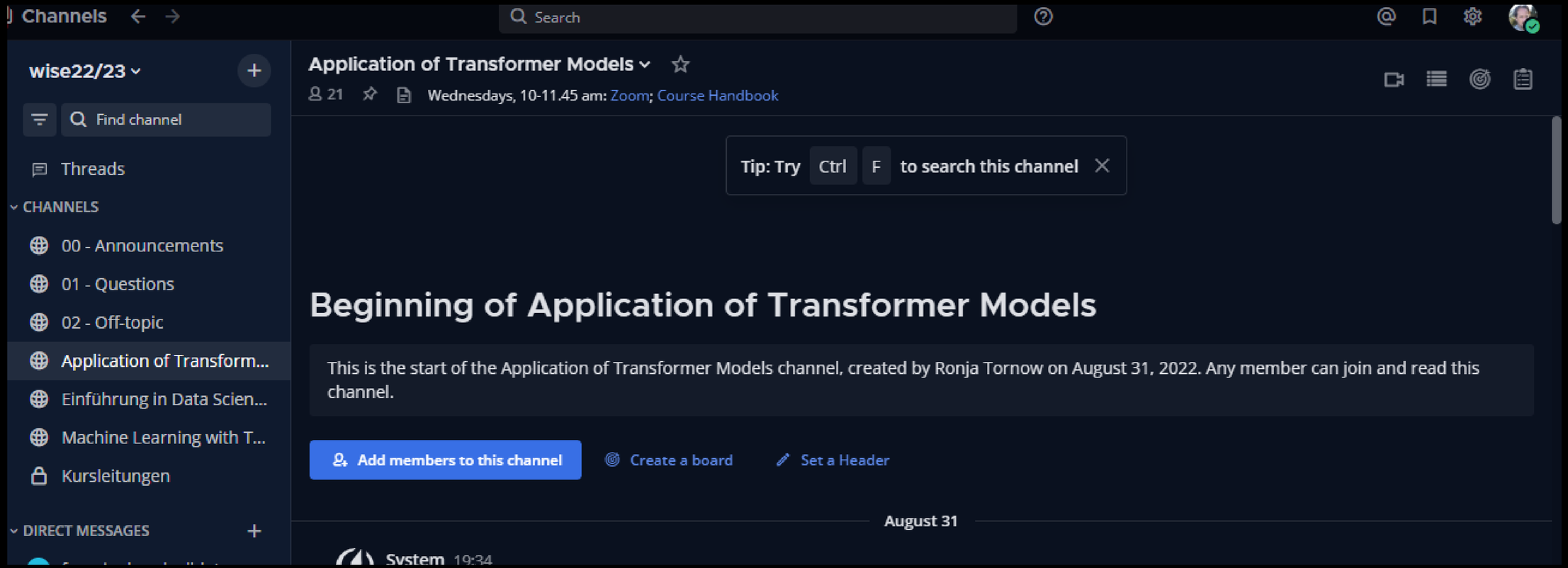
Application of Transformer Models

GENERAL INTRODUCTION

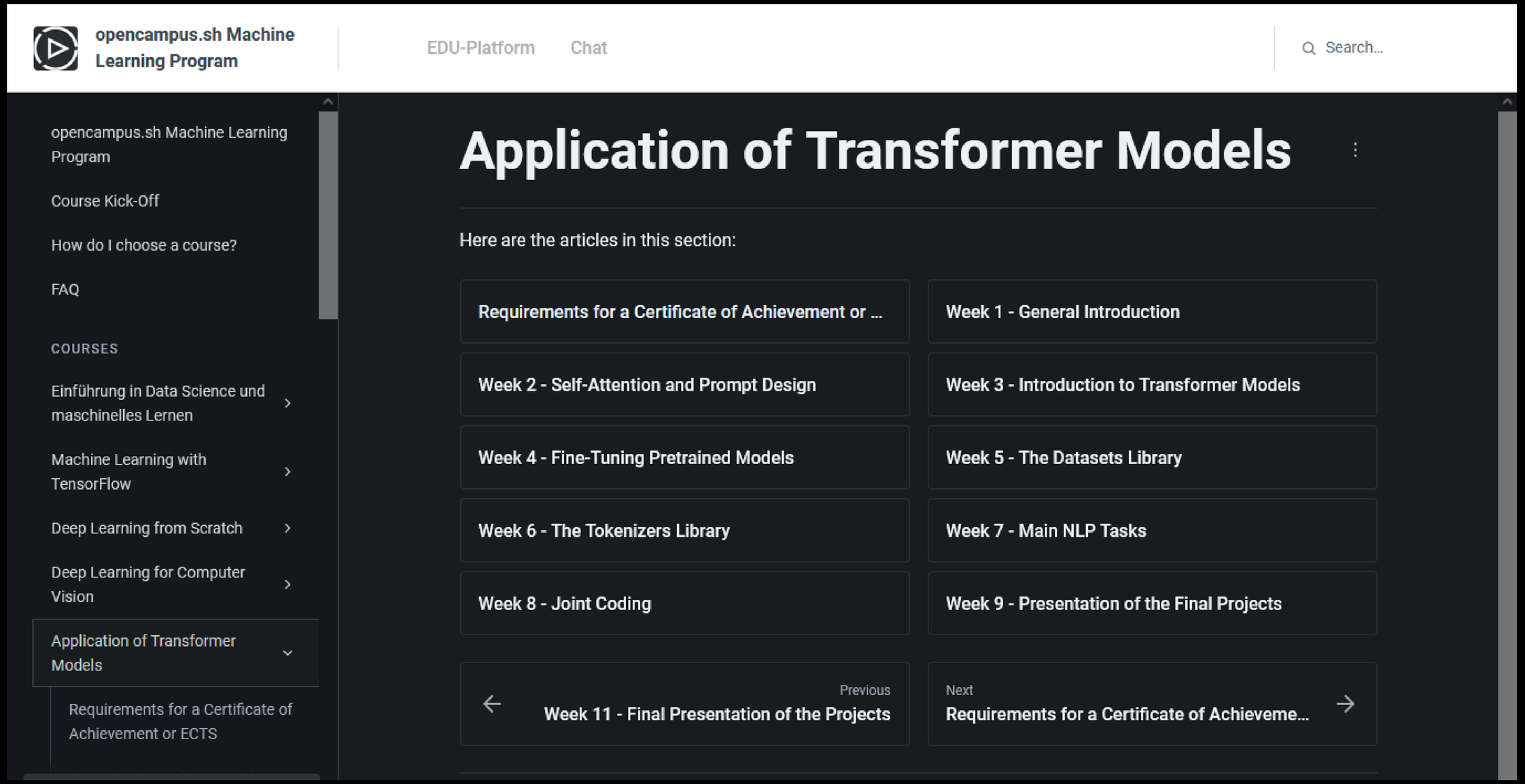
- Introduction Round
- Organizational Matters
- Structure of the Course
- The Role of NLP in AI
- Basic Introduction to Transformers
- How to Stay Up-To-Date

INTRODUCTION ROUND

CHAT


Channels
Search
wise22/23
Find channel
Threads
CHANNELS
00 - Announcements
01 - Questions
02 - Off-topic
Application of Transform...
Einführung in Data Scien...
Machine Learning with T...
Kursleitungen
DIRECT MESSAGES
Application of Transformer Models
21
Wednesdays, 10-11.45 am: Zoom; Course Handbook
Tip: Try Ctrl F to search this channel
Beginning of Application of Transformer Models
This is the start of the Application of Transformer Models channel, created by Ronja Tornow on August 31, 2022. Any member can join and read this channel.
Add members to this channel
Create a board
Set a Header
August 31
System 19:34

COURSE HANDBOOK

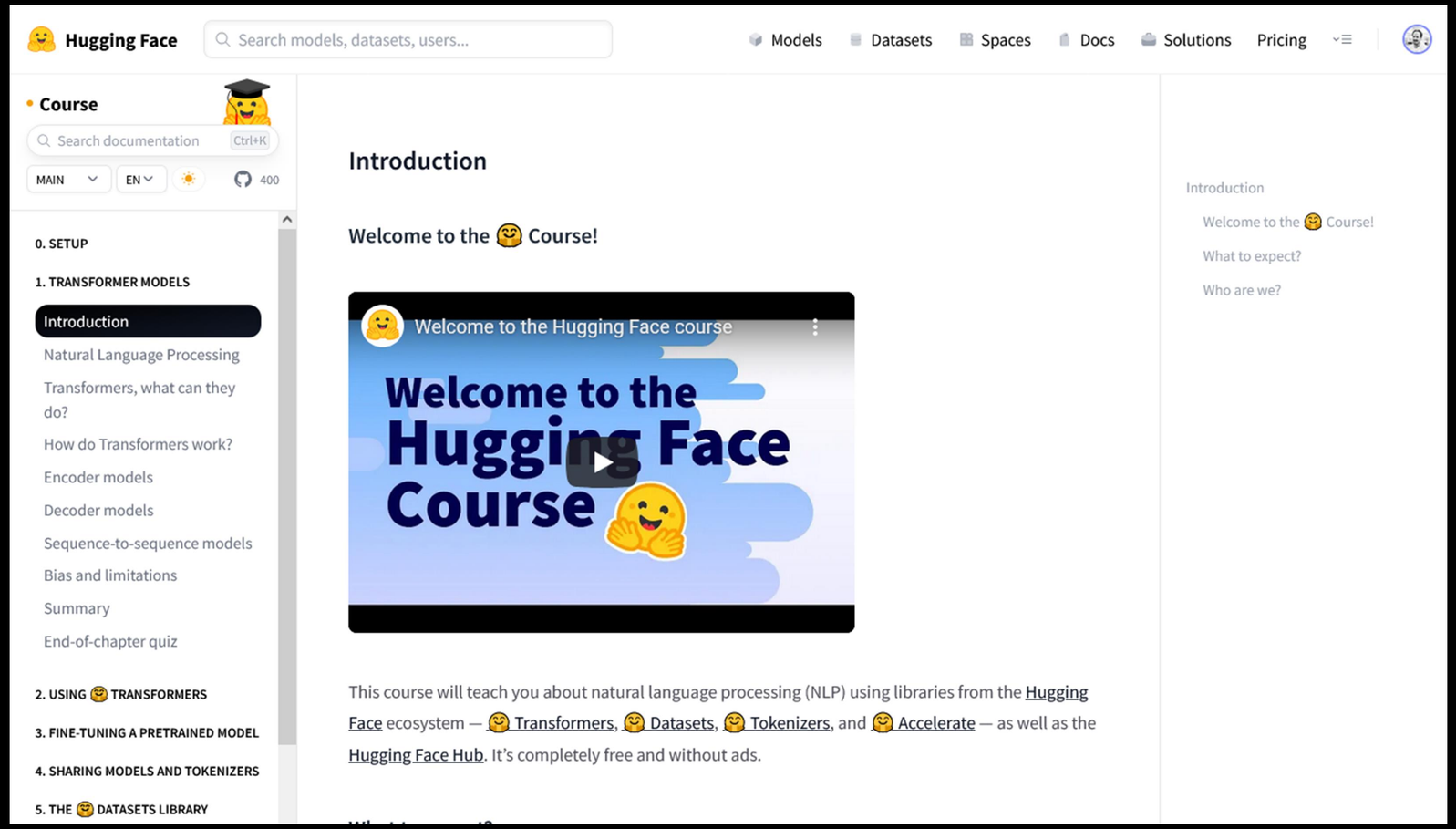


ORGANIZATIONAL MATTERS

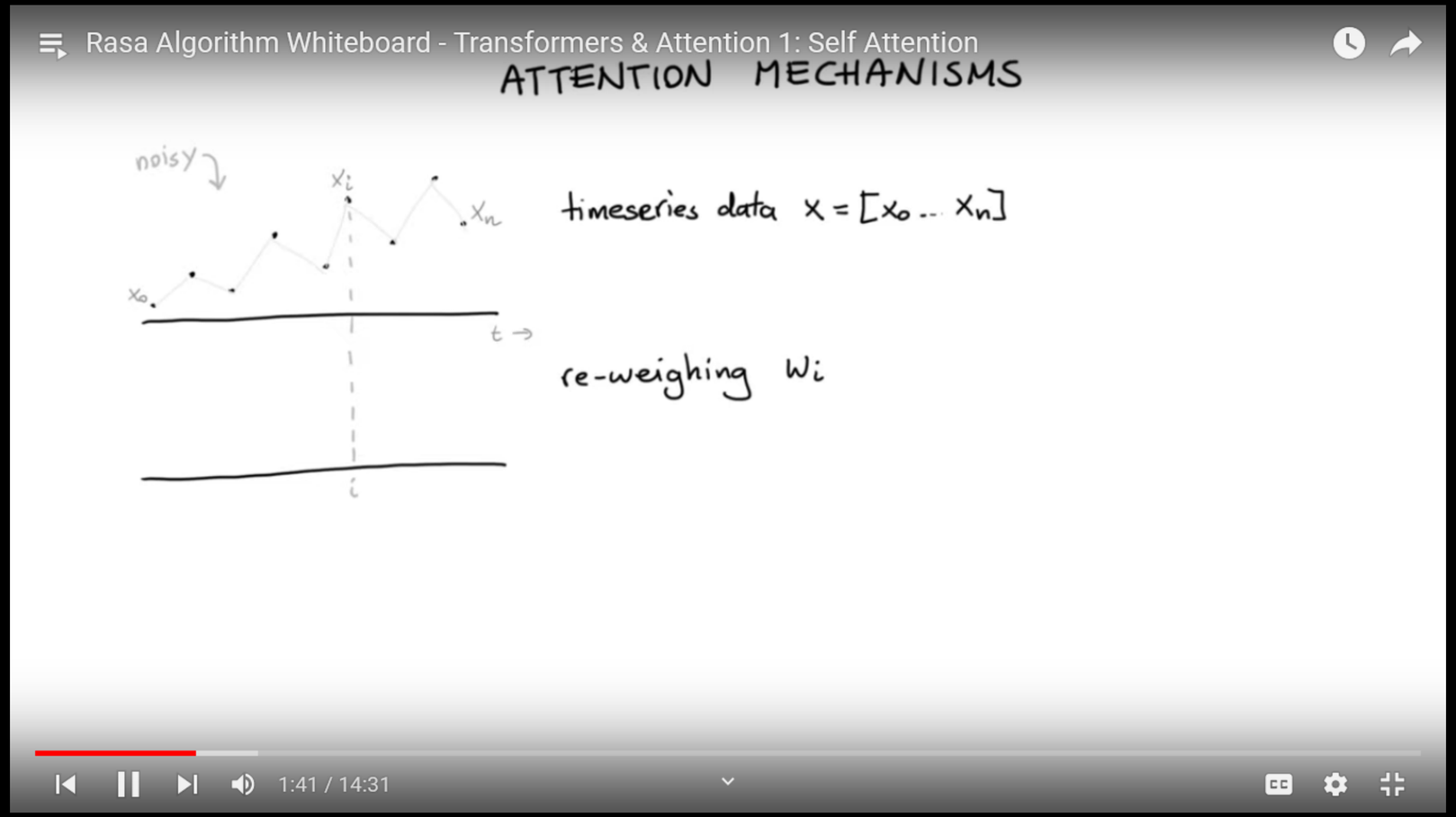
- Use your full names in the zoom meetings!
- Scan the QR-Code if you participate in presence.
- Complete your profile in the Mattermost chat with your full name and a photo.
- Please write me if you will not go on with the course!

26. 10. **10: 00 - 12: 00**	GENERAL INTRODUCTION Remote via Zoom
02. 11. **10: 00 - 12: 00**	SELF-ATTENTION AND PROMPT DESIGN Remote via Zoom
09. 11. **10: 00 - 12: 00**	INTRODUCTION TO TRANSFORMERS Remote via Zoom
16. 11. **10: 00 - 12: 00**	FINE-TUNING PRETRAINED MODELS Remote via Zoom
23. 11. **10: 00 - 12: 00**	THE DATASETS LIBRARY Remote via Zoom
30. 11. **10: 00 - 12: 00**	THE TOKENIZERS LIBRARY Remote via Zoom
04. 01. **10: 00 - 12: 00**	TRAINING MODELS FROM SCRATCH Remote via Zoom
11. 01. **10: 00 - 12: 00**	PRESENTATION OF THE FINAL PROJECTS Remote via Zoom

HUGGING FACE TRANSFORMERS COURSE



RASA ALGORITHM WHITEBOARD



BREAKOUT

- Which AI impressed you the most in the last few month?
- Do you know any transformer models?

APPLICATION EXAMPLES

- Whisper by OpenAI
- Stable Diffusion (Image Generation)
- Jukebox (Music Generation)
- Alpha Tensor

APPLICATION EXAMPLES

- Code Generation: Co-Pilot in VS-Code from GitHub and OpenAI (GPT model)
- Search: Google (Bert model)
- Prediction of Protein folding: Alpha Fold from Deepmind
- Image Generation: DALL-E 2 by OpenAI

AGI RESEARCH

DeepMind

OpenAI

ANTHROP\C

THE TRADITIONAL ROLE OF NLP IN AI

- A subfield of linguistics, computer science, and artificial intelligence
- Dealing with the interaction between computers and human language
- The goal is a computer capable of "understanding" text contents, including contextual nuances
- The technology can then accurately extract information and insights contained in the documents as well as categorize and organize the documents themselves.

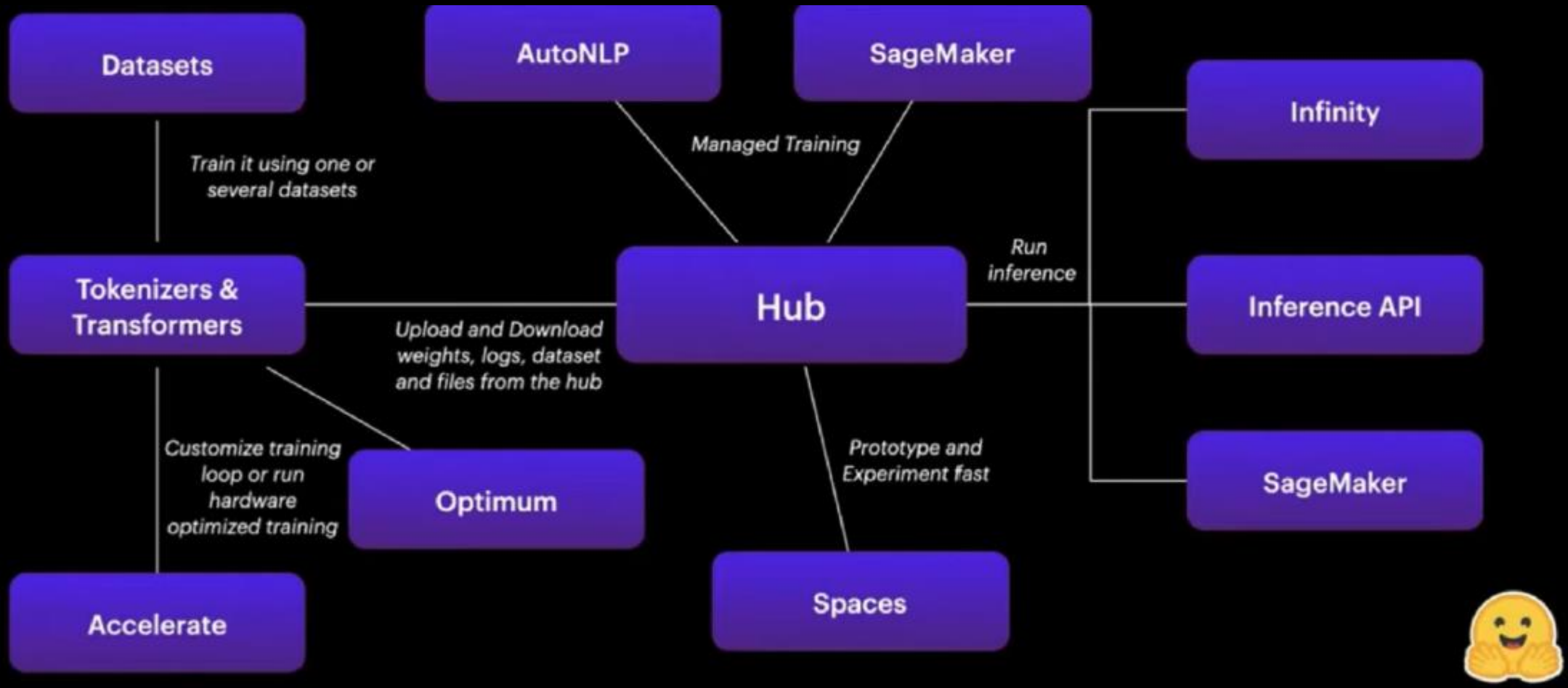
Natural language processing. (2021). In *Wikipedia*. https://en.wikipedia.org/w/index.php?title=Natural_language_processing&oldid=1048621730

THE ROLE OF NLP FOR ARTIFICIAL GENERAL INTELLIGENCE (AGI)

Touring Test:

- **Test for intelligence in a computer, requiring that a human being should be unable to distinguish the machine from another human being by using the replies to questions put to both.**

HUGGING FACE ECO SYSTEM



WHAT IS A TRANSFORMER?

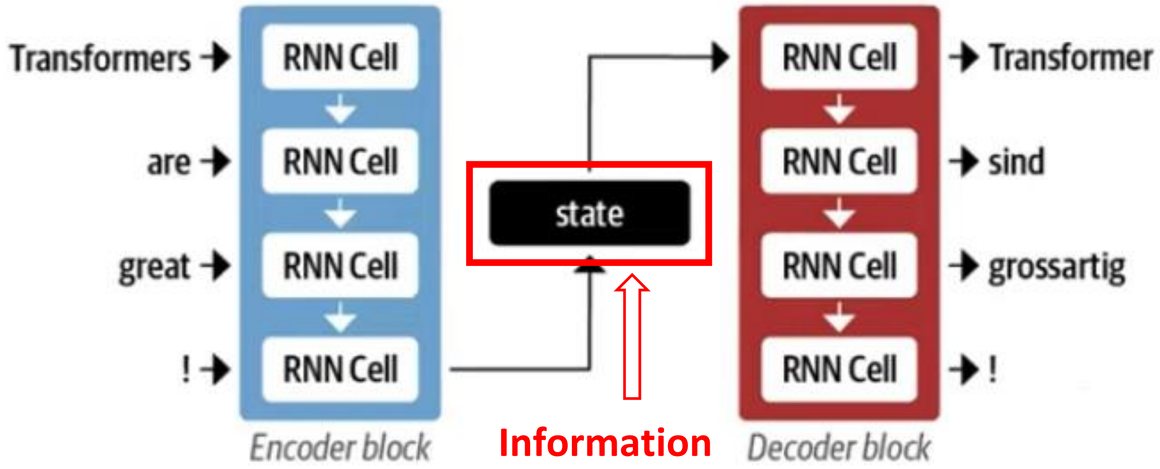
Main ingredients

Attention mechanisms

Self-supervised learning (Pretraining)

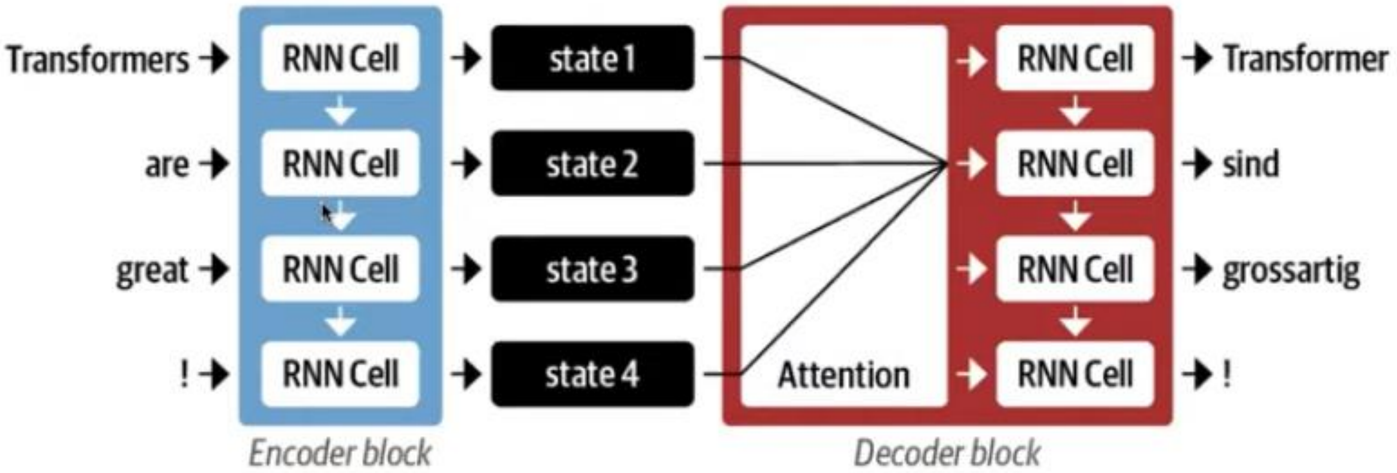
Transfer learning (Fine-tuning)

CLASSICAL SEQUENCE TO SEQUENCE APPROACH



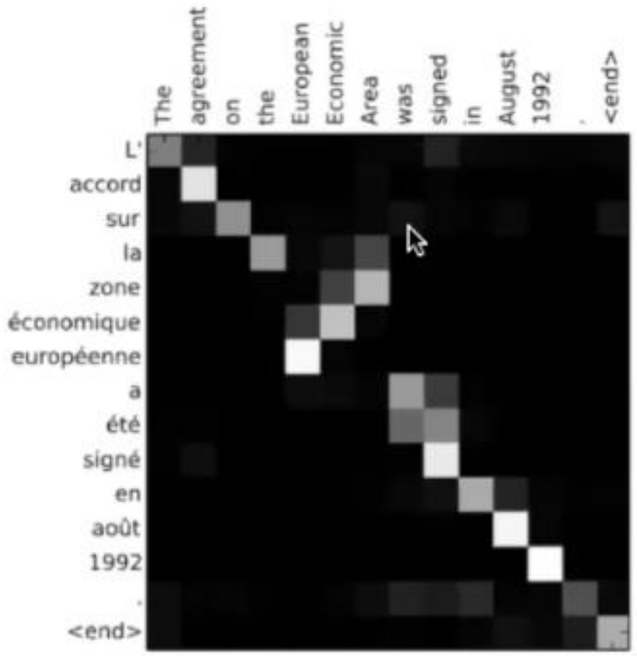
Originally developed for recurrent neural networks

ATTENTION MECHANISM



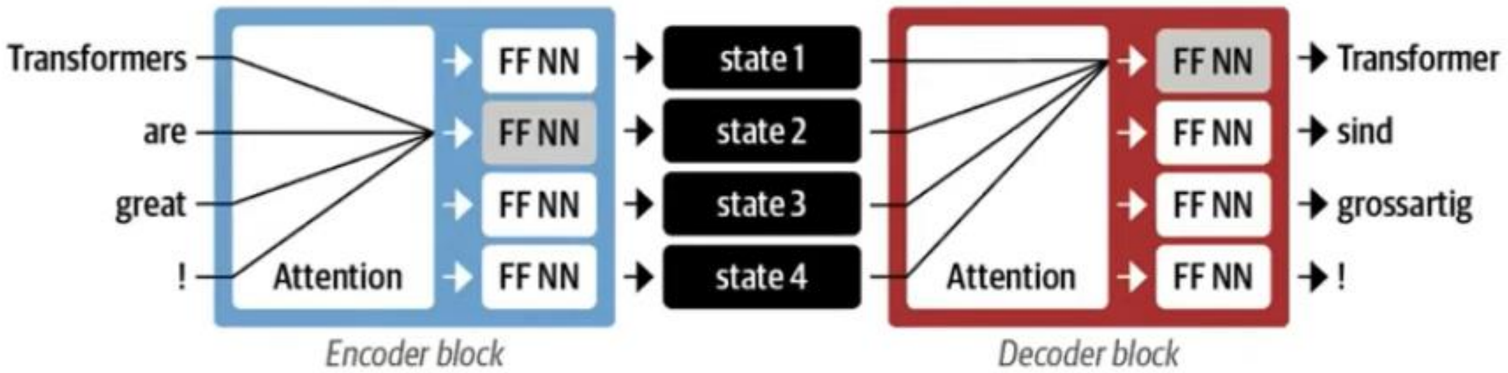
Assign a weight or "pay attention" to specific states

INTERPRETATION



Attention gives better modelling of word order

ATTENTION IS ALL YOU NEED



Transformers much easier to scale with compute & data

BENEFITS

- Solving the bottleneck problem in sequence-to-sequence tasks
- Provides some interpretability
- No Vanishing Gradient Problem
- Multimodality

HUGGING FACE MODEL HUB

- https://huggingface.co/models

- Check out models for one or two tasks, you are not that familiar with yet.

STAYING UP-TO-DATE

STAYING UP-TO-DATE

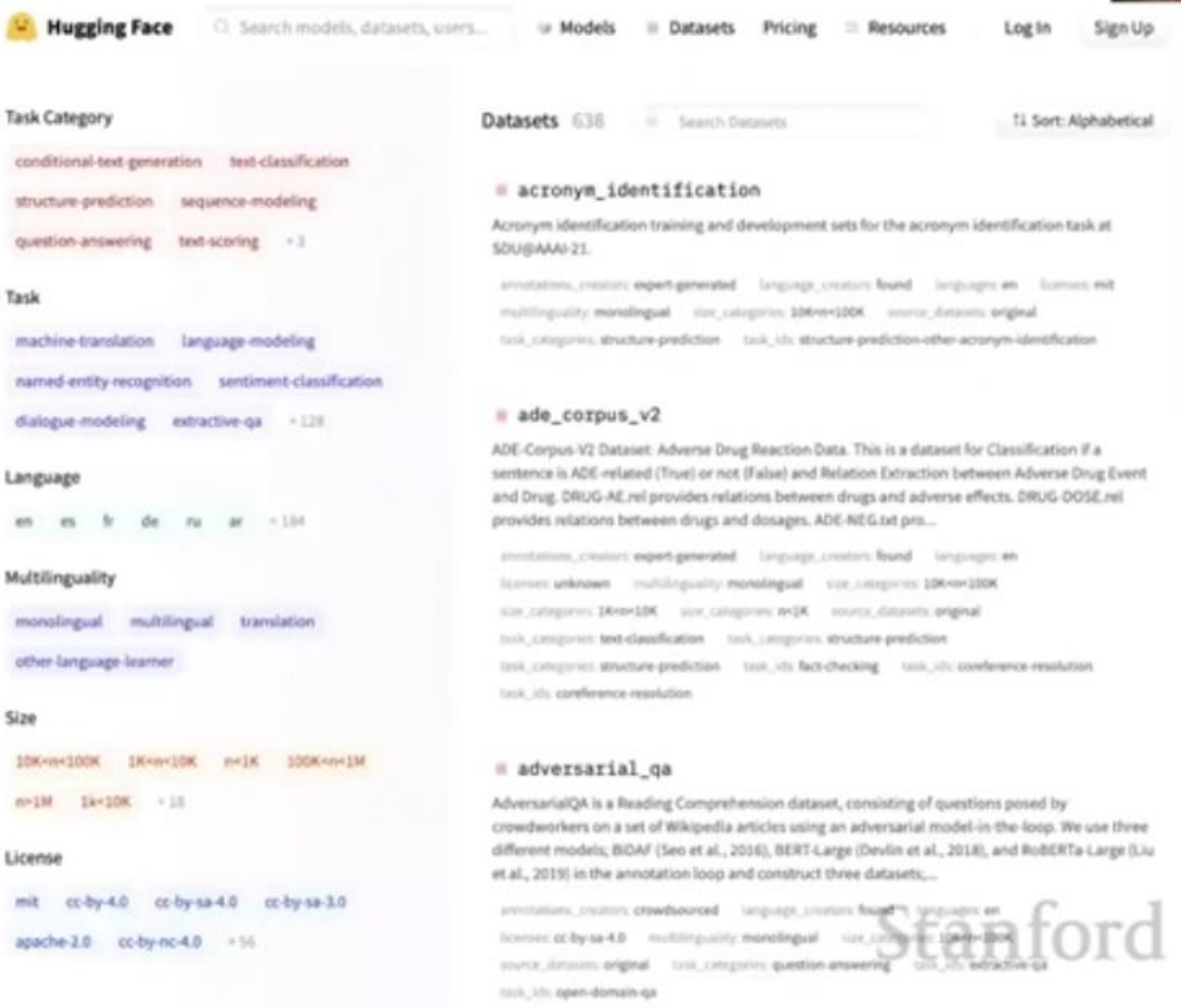
- Newsletter
  - Medium
  - Reddit
  - The Batch
  - …
- Twitter
  - Thomas Wolf
  - Nvidia
  - …
- LinkedIn
  - Steve Nouri
  - Hugging Face
  - …
- ArXive / ArXive Sanity

TODOS FOR NEXT WEEK

- Watch video 1 (Self-Attention) and video 2 (Keys, Values, Queries) of the Rasa Series on Transformers & Attention.
- Note at least one question on the each of the videos!
- Do chapter 1 of the Hugging Face course.

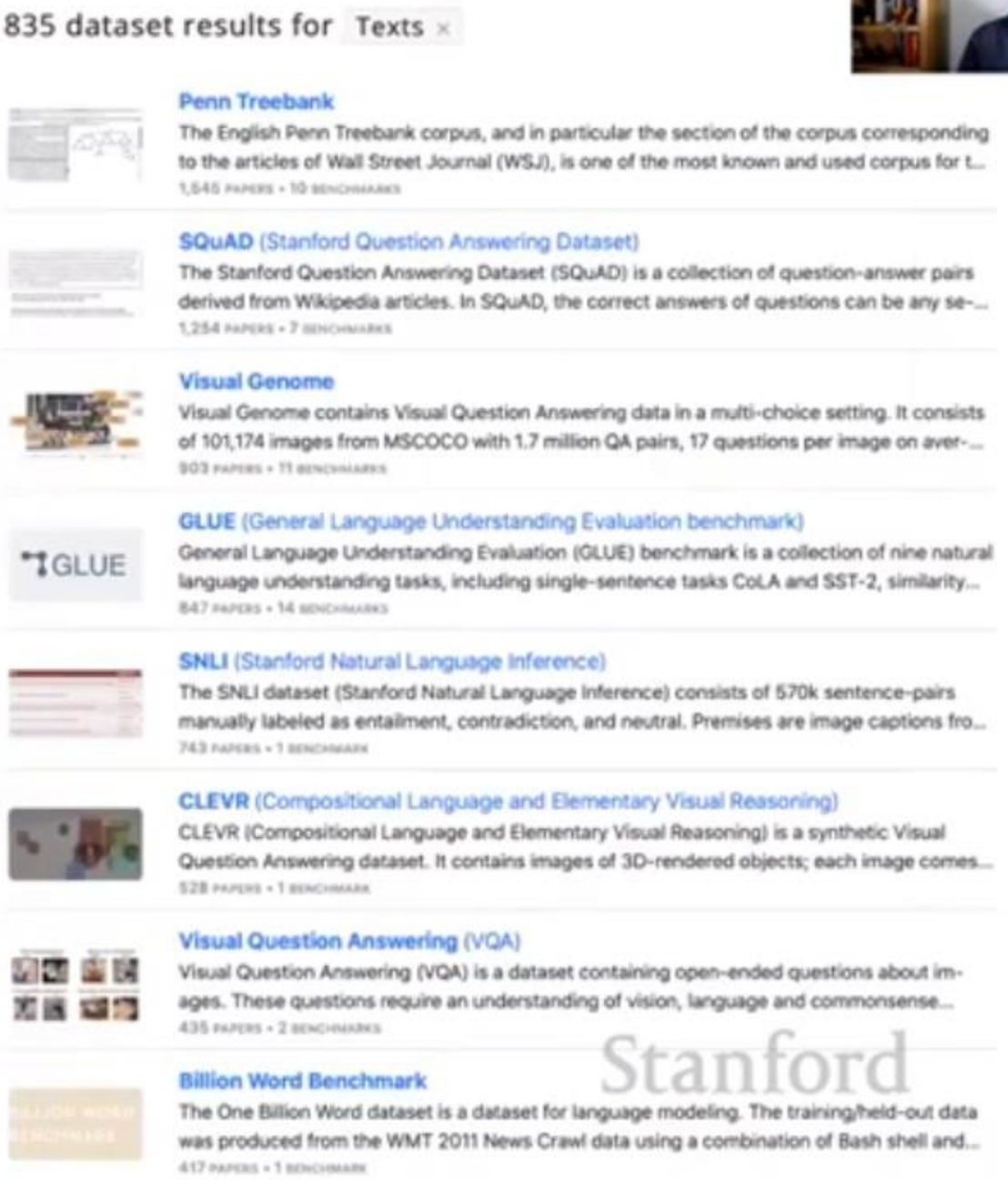
Huggingface Datasets

- https://huggingface.co/datasets



Paperswithcode Datasets

- https://www.paperswithcode.com/datasets?mod=texts&page=1



Linguistic Data Consortium

- https://catalog.ldc.upenn.edu/
- Stanford licenses data; you can get access by signing up at: https://linguistics.stanford.edu/resources/resources-corpora
- Treebanks, named entities, coreference data, lots of clean newswire text, lots of speech with transcription, parallel MT data, etc.
  - Look at their catalog
  - Don't use for non-Stanford purposes!

Dependency parsing: Universal Dependencies

- https://universaldependencies.org

Universal Dependencies

Universal Dependencies (UD) is a framework for cross-linguistically consistent grammatical annotation and an open community effort with over 200 contributors producing more than 100 treebanks in over 70 languages.

- Short introduction to UD
- UD annotation guidelines
- More information on UD:
  - How to contribute to UD
  - Tools for working with UD
  - Discussion on UD
  - UD-related events
- Query UD treebanks online:
  - SETS treebank search maintained by the University of Turku
  - PML Tree Query maintained by the Charles University in Prague
  - Kontext maintained by the Charles University in Prague
  - Grew-match maintained by Inria in Nancy
- Download UD treebanks

If you want to receive news about Universal Dependencies, you can subscribe to the UD mailing list. If you want to discuss individual annotation questions, use the Github issue tracker.