

26.10.21

Natural Language Processing with Transformers

EINFÜHRUNG

VORSTELLUNG

- **Organisatorisches**
- **Aufbau des Kurses**
- **Definition von NLP**
- **Rolle von NLP in der KI**
- **Grundlegende Merkmale von Transformern**

EDU.OPENCAMPUS.SH

CHAT

Top Left: User profile for **sose21 @steffen** with a yellow arrow pointing to it.


Channel Header: **C_Einführung in Data Science & maschinelles Lernen** (46 members, 1 star). Description: **Dienstags, 18-20 Uhr: Zoom; Kurshandbuch**.

Left Sidebar: Search bar "Find channel", navigation icons, and a list of channels including "00 - Announcements", "01 - Questions", "C_Advanced Machine Lear...", "C_Deep Learning from Scr...", "C_Einführung in Data Scie..." (highlighted), "C_Machine Learning für di...", "C_Machine Learning With ...", and "Kursleitungen". Below is a "DIRECT MESSAGES" section.

Central Chat Area: Header "Beginning of C_Einführung in Data Science & maschinelles Lernen" with a description: "This is the start of the C_Einführung in Data Science & maschinelles Lernen channel, created by Luca Palmieri on March 25, 2021. Any member can join and read this channel." Below are options to "Invite others to this channel" and "Set a Header". A date separator for "March 25" is shown. A pinned message from **Steffen Brandt** (22:58) reads: "Willkommen im Kurs 'Einführung in Data Science und maschinelles Lernen'! Ich werde versuchen Euch in diesem Semester die Themen Data Science und maschinelles Lernen etwas näher zu bringen. Dabei werde ich vor allem versuchen, Euch viel praktisches Durchführungswissen mit an die Hand zu geben. Ein wichtiger Teil für die Arbeit im Bereich Data Science ist aber vor allem auch die Vernetzung und die Arbeit im Team. Deswegen ist es auch ein wichtiges Ziel des Kurses, dass Ihr Euch untereinander kennenlernt und ein Projekt im Team durchführt. Ich möchte ich Euch daher".

Right Sidebar: Section "Pinned Posts" for "C_Einführung in Data Scie...". It shows the same pinned message from **Steffen Brandt** (22:58) with a "Jump" link.

KURSHANDBUCH

 **opencampus.sh Machine Learning Program**

EDU-Plattform Chat

🔍 Search...

opencampus.sh Machine Learning Program

Course Kick-Off

How do I choose a course?

FAQ

COURSES

Einführung in Data Science und maschinelles Lernen ▾

- Bedingungen für ein Leistungszertifikat oder ECTS
- Vorbereitung
- Woche 1 - Einführung in Data Science
- Woche 2 - Grafische Darstellung von Daten


Woche 1 - Einführung in Data Science


Diese Woche werdet Ihr...

eine Einführung zu den folgenden Themen bekommen:

- Was ist Data Science?
- R vs. Python vs. SPSS vs. ...
- Wozu RStudio?
- Datenstrukturen in R

Lernressourcen

 **Präsentation zu Woche 1 - Einführung** 201114_Einführung.pdf - 5MB

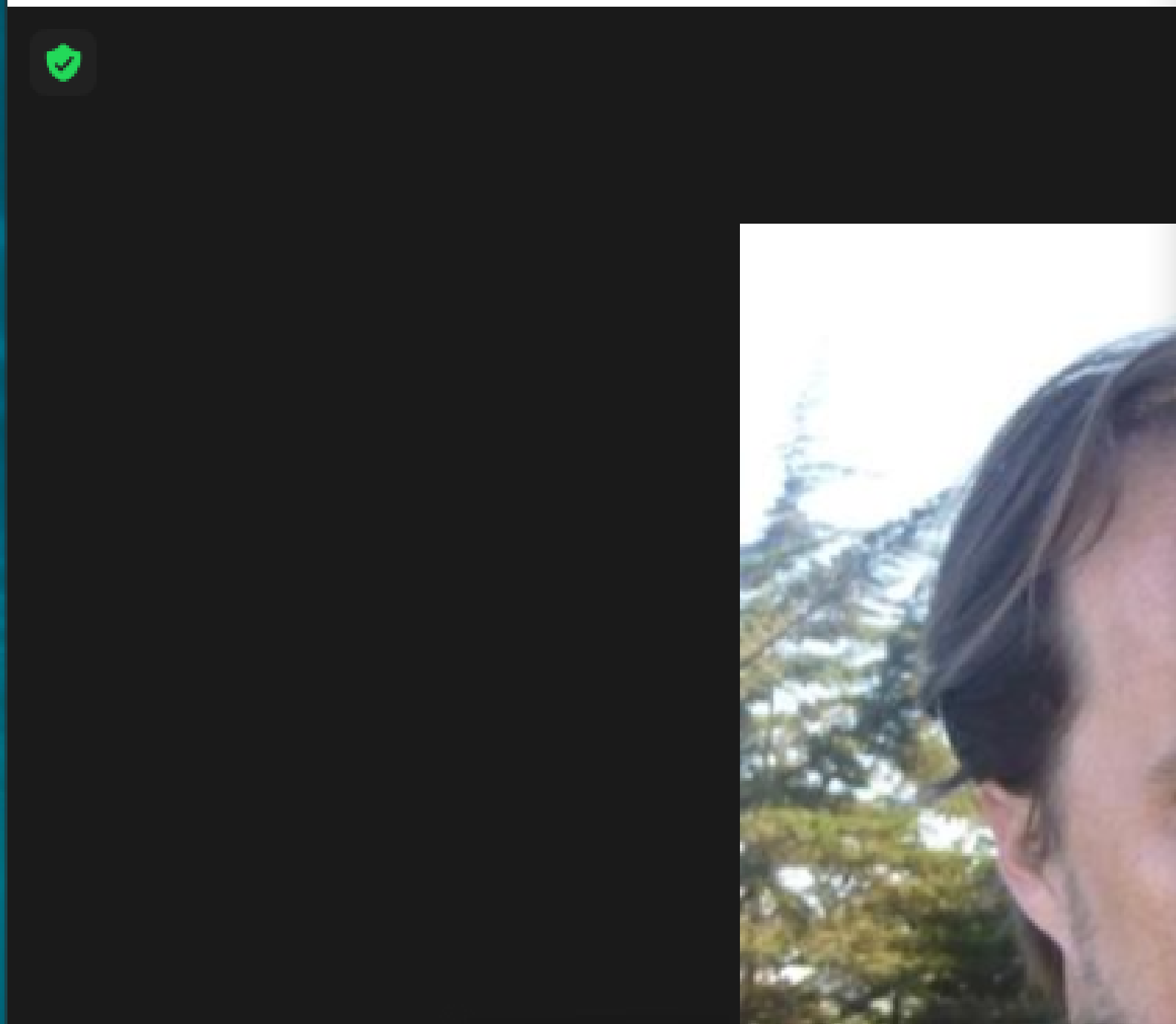
 CONTENTS

- Diese Woche werdet Ihr...
- Lernressourcen
- Bis zur nächsten Woche sollt...

ZOOM

- **Versucht die unterschiedlichen View Modes:**
 - **Gallery View/ Active Speaker**
 - **Split Screen/ Full Screen Mode**
- **Schaut Euch ggf. dazu dieses Video an:**
<https://www.youtube.com/watch?v=v3IPAbpVjd4>

Zoom Meeting



Select a Camera (Alt+N to switch)

NewTek NDI Video

✓ Integrated Camera

Choose Virtual Background...

Choose Video Filter...

Video Settings...

Steffen Brandt



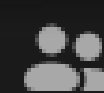
Mute



Start Video



Security



Participants 1

Settings

General

Video

Audio

Share Screen

Chat

Background & Filters

Recording

Profile

Statistics

Keyboard Shortcuts

Accessibility

Window size when screen sharing:

☐ Fullscreen mode

☐ Maximize window

☒ Maintain current size

☒ Scale to fit shared content to Zoom window

☐ Show my Zoom Windows to other participants when I am screen sharing

☐ Enable the remote control of all applications

☒ Side-by-side mode

☒ Silence system notifications when sharing desktop

When I share my screen in a meeting

☐ Automatically share desktop

☒ Show all sharing options

When I share directly to a Zoom Room

☒ Automatically share desktop

☐ Show all sharing options

Advanced

ORGANISATORISCHES

- **Benutzt Euren vollständigen Namen in den Zoom-Meetings!**
- **Scannt den QR-Code, wenn Ihr in Präsenz teilnehmt**
- **Vervollständigt Euer Profil im Mattermost-Chat mit Eurem Namen und einem Foto.**
- **Bitte schreibt mir, wenn Ihr den Kurs abbrecht!**

27. 10. 16: 00- 17: 45	GENERAL INTRODUCTION INTO NATURAL LANGUAGE PROCESSING Zoom + Starterkitchen, Kuhnkestraße 6, Wissenschaftspark
03. 11. 16: 00- 17: 45	TRANSFORMERS IN APPLICATIONS Zoom + Starterkitchen, Kuhnkestraße 6, Wissenschaftspark
10. 11. 16: 00- 17: 45	INTRODUCTION TO TRANSFORMER MODELS Zoom + Starterkitchen, Kuhnkestraße 6, Wissenschaftspark
17. 11. 16: 00- 17: 45	USING TRANSFORMERS Zoom + Starterkitchen, Kuhnkestraße 6, Wissenschaftspark
08. 12. 16: 00- 17: 45	FINE-TUNING A PRETRAINED MODEL Zoom only!
05. 01. 16: 00- 17: 45	SELECTING THE RIGHT MODEL Zoom + Starterkitchen, Kuhnkestraße 6, Wissenschaftspark
12. 01. 16: 00- 17: 45	SPECIAL ISSUES CONSIDERING YOUR FINAL PROJECTS Zoom + Starterkitchen, Kuhnkestraße 6, Wissenschaftspark
19. 01. 16: 00- 17: 45	PRESENTATION OF THE FINAL PROJECTS Zoom + Starterkitchen, Kuhnkestraße 6, Wissenschaftspark

ERSTER BREAKOUT

- **Welche KI hat Euch in den letzten Monaten am meisten beeindruckt?**
- **Kennt Ihr bereits Transformer-Modelle – wenn ja, welche?**

DEFINITION OF NLP

- **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.**

ROLLE DES NLP IN DER KI

Touring Test

- **Test der Fähigkeit einer Maschine, intelligentes Verhalten zu zeigen, das dem eines Menschen entspricht oder von diesem nicht zu unterscheiden ist.**
- **Ein:e menschliche:r Bewerter:in beurteilt Gespräche in natürlicher Sprache zwischen einem Menschen und einer Maschine und muss entscheiden, wer der Mensch ist.**

AGI-FORSCHUNG



ANTHROPIC

CHALLENGES IN NLP

- **Speech recognition**
- **Natural language understanding**
- **Natural language generation.**

TRADITIONAL TASKS IN NLP

- **Text and speech processing**
- **Morphological analysis**
- **Syntactic analysis**
- **Lexical semantics (of individual words in context)**
- **Relational semantics (semantics of individual sentences)**
- **Discourse (semantics beyond individual sentences)**
- **Higher-level NLP applications**

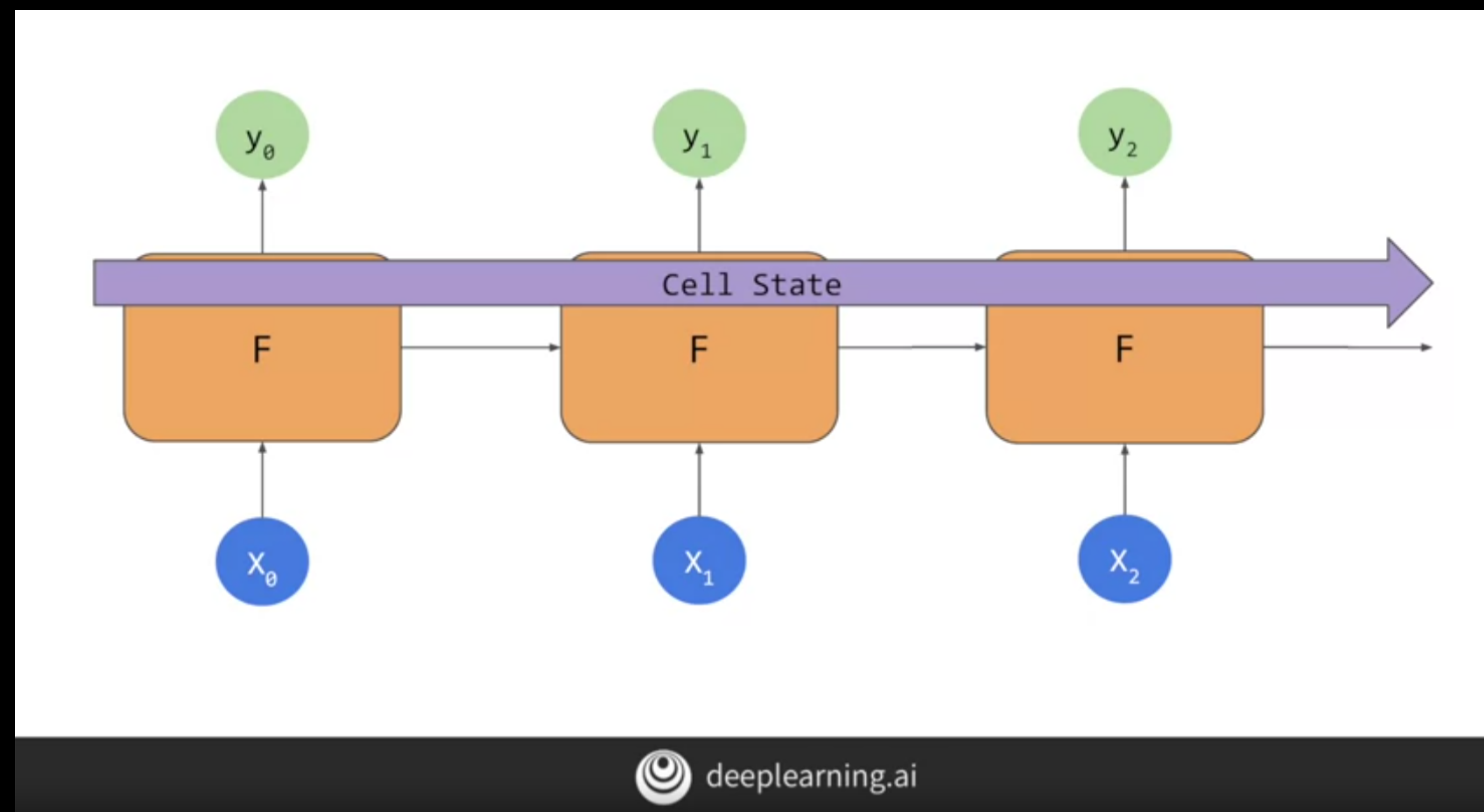
HISTORY

- **Symbolic NLP**
- **(Theory-/ feature-based) Statistical NLP**
- **NLP using Neural Networks**

LONG SHORT-TERM MEMORY (LSTM)

(Hochreiter, S., & Schmidhuber, J. (1997). Long Short-term Memory. Neural Computation, 9, 1735–1780.)

Seit 2015 und bis 2019/20 stark genutzt von Google, Facebook, Amazon, and Apple.



BASIC TRANSFORMER CHARACTERISTICS

- **Large Models**
- **Large Training Data Sets**
- **Unsupervised Learning Methods**
- **No Vanishing Gradient Problem**
- **Multimodality**

TODOS BIS ZUR NÄCHSTEN WOCHE

- **Recherchieren guter Quellen:**
 - **Newsletter, Blogs, Personen, denen man in Social Media folgen kann**
 - **Postet diese gerne bereits bis zum nächsten Termin in Mattermost**
- **Mindestens drei aktuelle Anwendungen von Transformer-Modellen heraussuchen (versucht auch eine multimodale zu finden).**