

# **Sequence Learning**

## **Introduction**

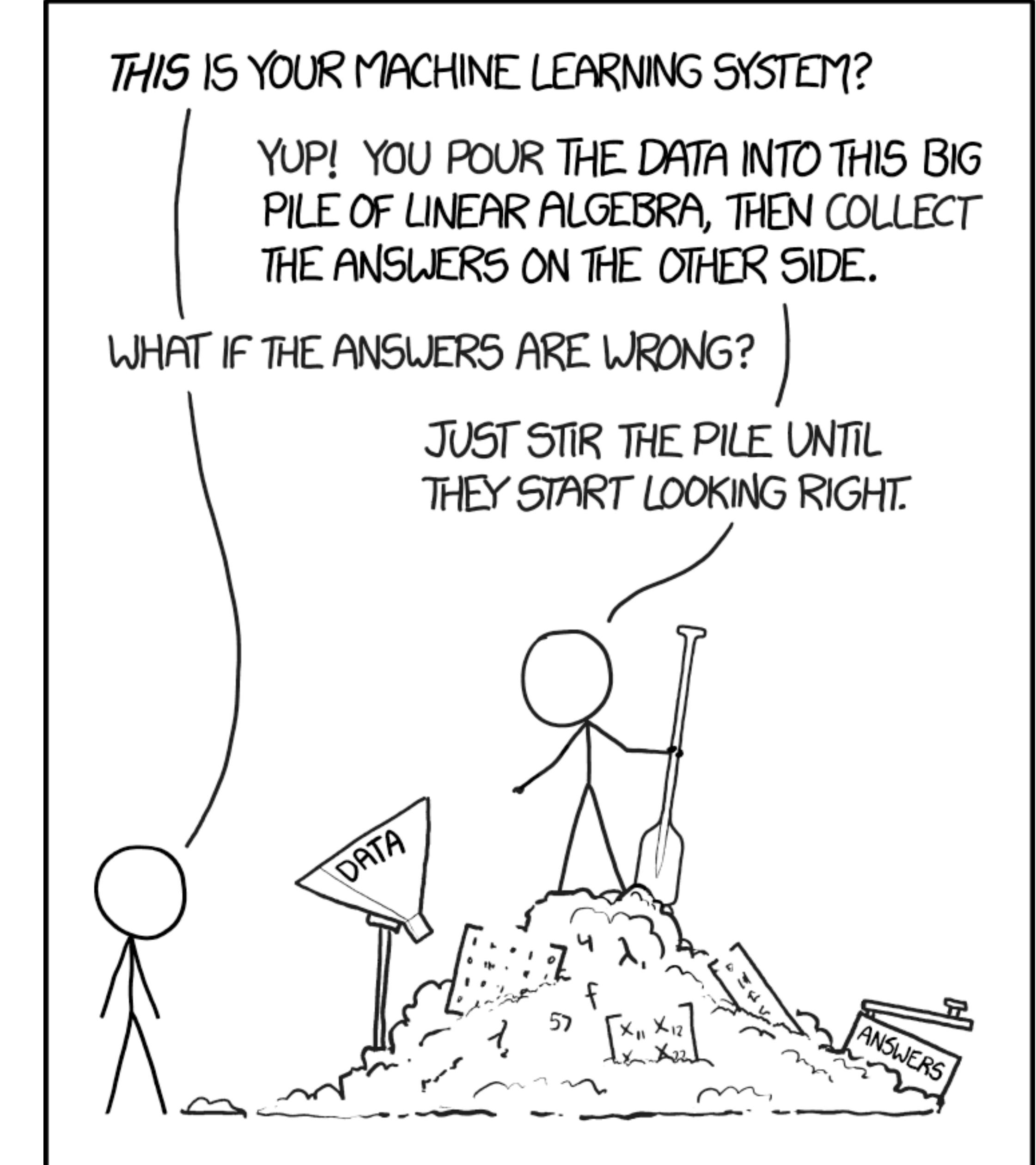
**Korbinian Riedhammer**



TECHNISCHE HOCHSCHULE NÜRNBERG  
GEORG SIMON OHM

# Today

- Logistics
- Why I teach this class
- Why you should take this class
- Motivation
- Syllabus
- What you should bring to this class



# Logistics

- Mondays at 9.45a on Zoom, discussion on Teams (Code: 4fbxju8)
- Materials: <https://seqlrn.github.io> (continuously updated...)
- New/changed this summer:
  - mandatory assignments in python (pair-programming ok)
  - 20' oral exam in the last week of lecture period (July 5)

# Why I teach this class

- Industry background in speech recognition/indexing ([mod9.io](#))
- Research focus
  - Speech processing for medical applications (eg. stuttering, dementia)
  - Speech recognition for indexing/search
  - Sequence learning for industrial applications (mostly anomaly detection)

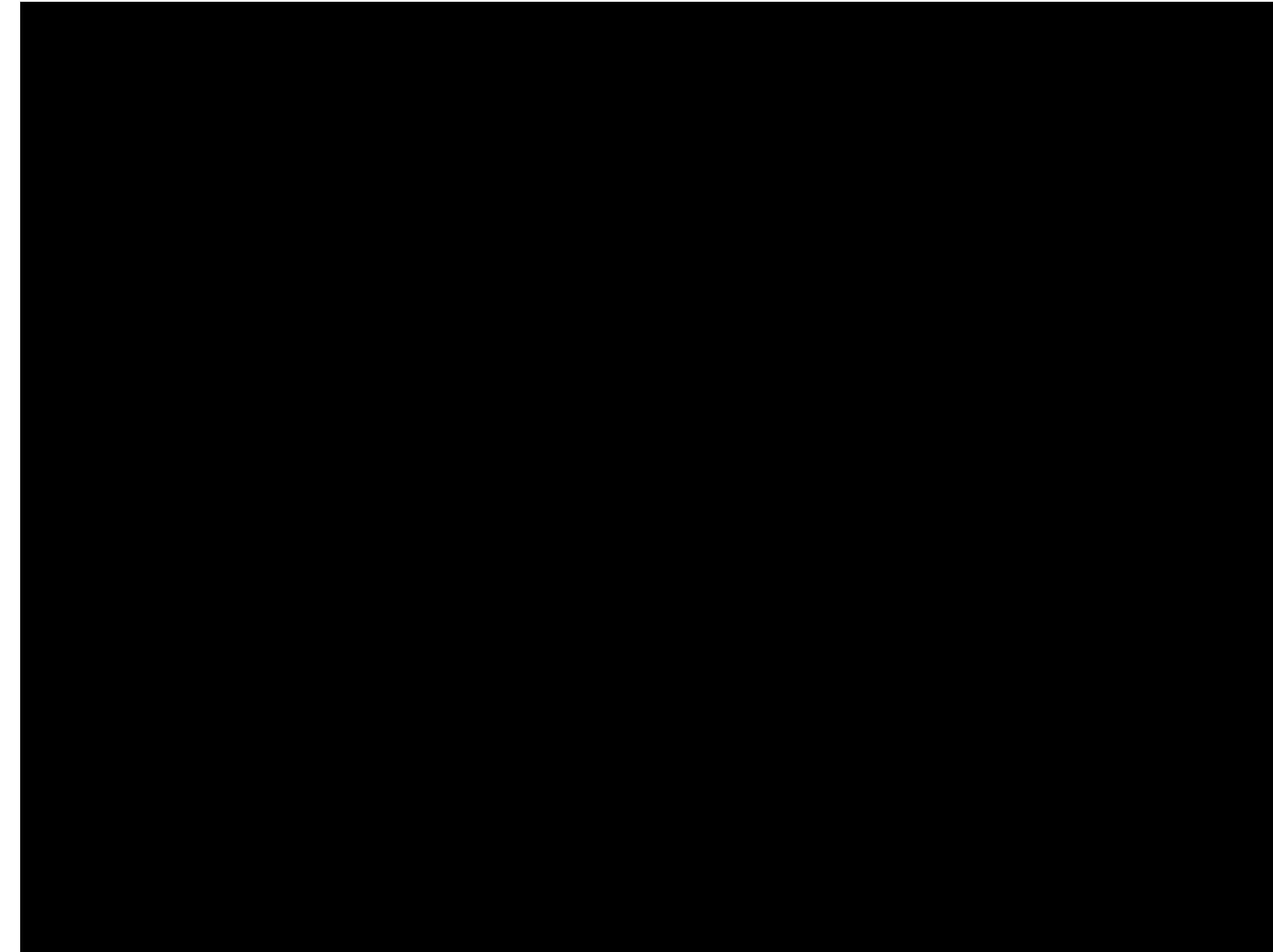
# Why you should take this class

- Machine learning is the future\*
- Many applications are to sequences, not single observations
- Understand the foundations of sequence classification

\*or at least a very well paid part of it

# **Flashback: Verbmobil**

## **Research project 1993-2000 (!)**



<https://www.youtube.com/watch?v=DcG9-KWx0Fg>

# Motivation: AI in Fiction

# **Star Trek & Star Trek TNG**

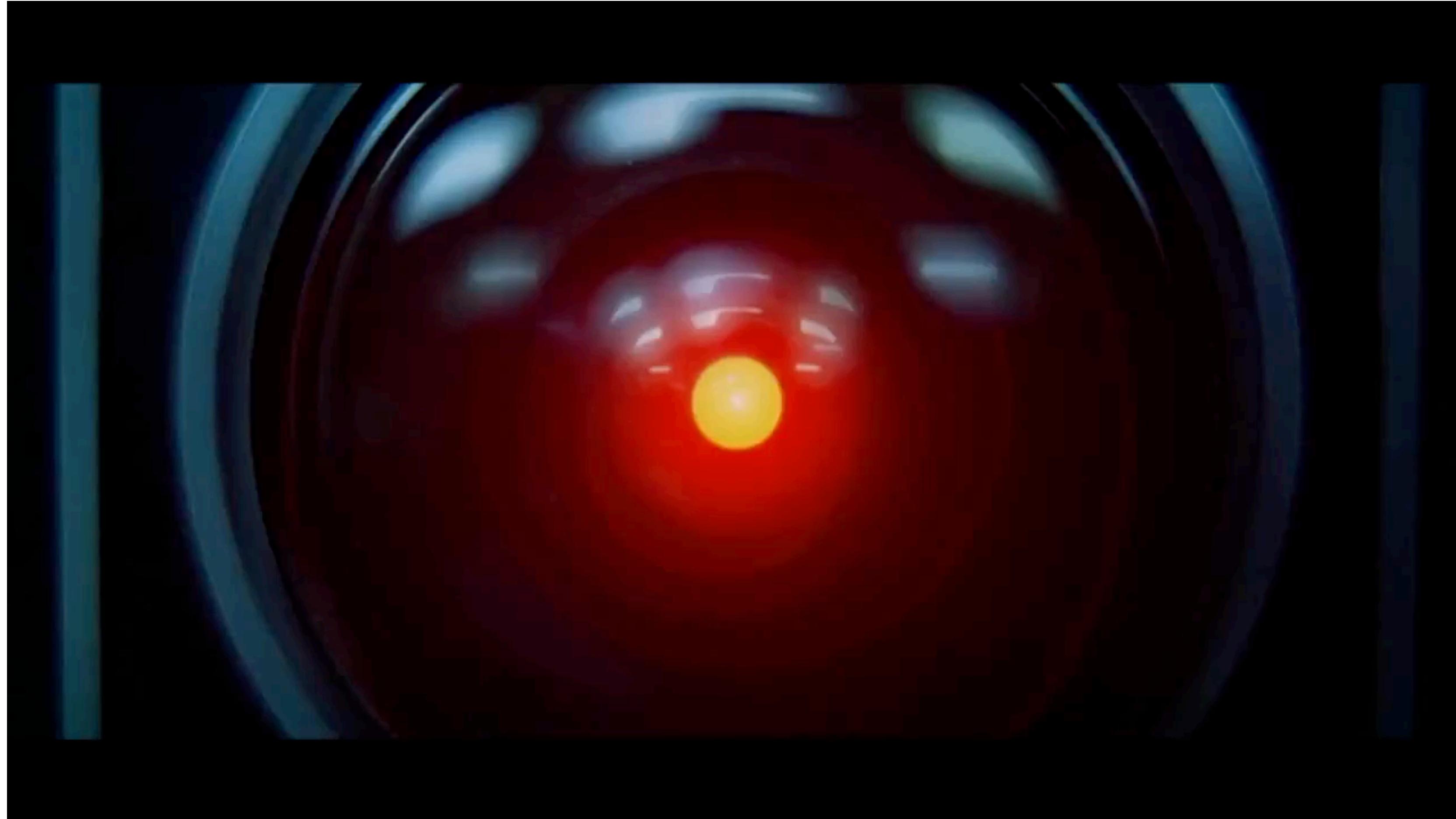
**1966, 1987**



<https://www.youtube.com/watch?v=tpKx7Oj0oeM>

# **2001: Space Odyssee**

**1968**



# Star Wars: A New Hope

1977



<https://www.youtube.com/watch?v=ElZfE1AVDPQ>

# Knight Rider

1982



[https://www.youtube.com/watch?v=OY\\_nGtN4ARA](https://www.youtube.com/watch?v=OY_nGtN4ARA)

# Charlie's Angels

2000



<https://www.youtube.com/watch?v=rRHRa80wVq8>

# **24 (S02E23)**

**2003**



"Cypress Recording" wurde durch Sound Engine mit angelernter Sprachsynthese erstellt.

# Wild Hogs

## 2007



[https://www.youtube.com/watch?v=qVZtE3rL\\_sQ](https://www.youtube.com/watch?v=qVZtE3rL_sQ)

Her  
2013



<https://www.youtube.com/watch?v=ne6p6MfLBxc>

# Blade Runner 2049

2018



[https://www.youtube.com/watch?v=RL0gX1\\_NWTk](https://www.youtube.com/watch?v=RL0gX1_NWTk)

# Motivation: AI in Products

# Radio Rex

1920



“Classic” signal processing: triggers on 500Hz (“reks”)

# **Worlds of Wonder's Julie Doll**

**1987**



<https://www.youtube.com/watch?v=UkU9Sblictc>

# **PenPoint OS**

## **1991**



<https://www.youtube.com/watch?v=x0XE08BjQDQ>

# Graffiti (Palm OS)

1997



<https://www.youtube.com/watch?v=iL0YLuClysY>

# Microsoft Speech Recognition

2008



<https://www.youtube.com/watch?v=-0kDcUEDfmY>

# **BMW Voice Control**

**2009**



<https://www.youtube.com/watch?v=xJo9pK42VRs>

# Apple Siri

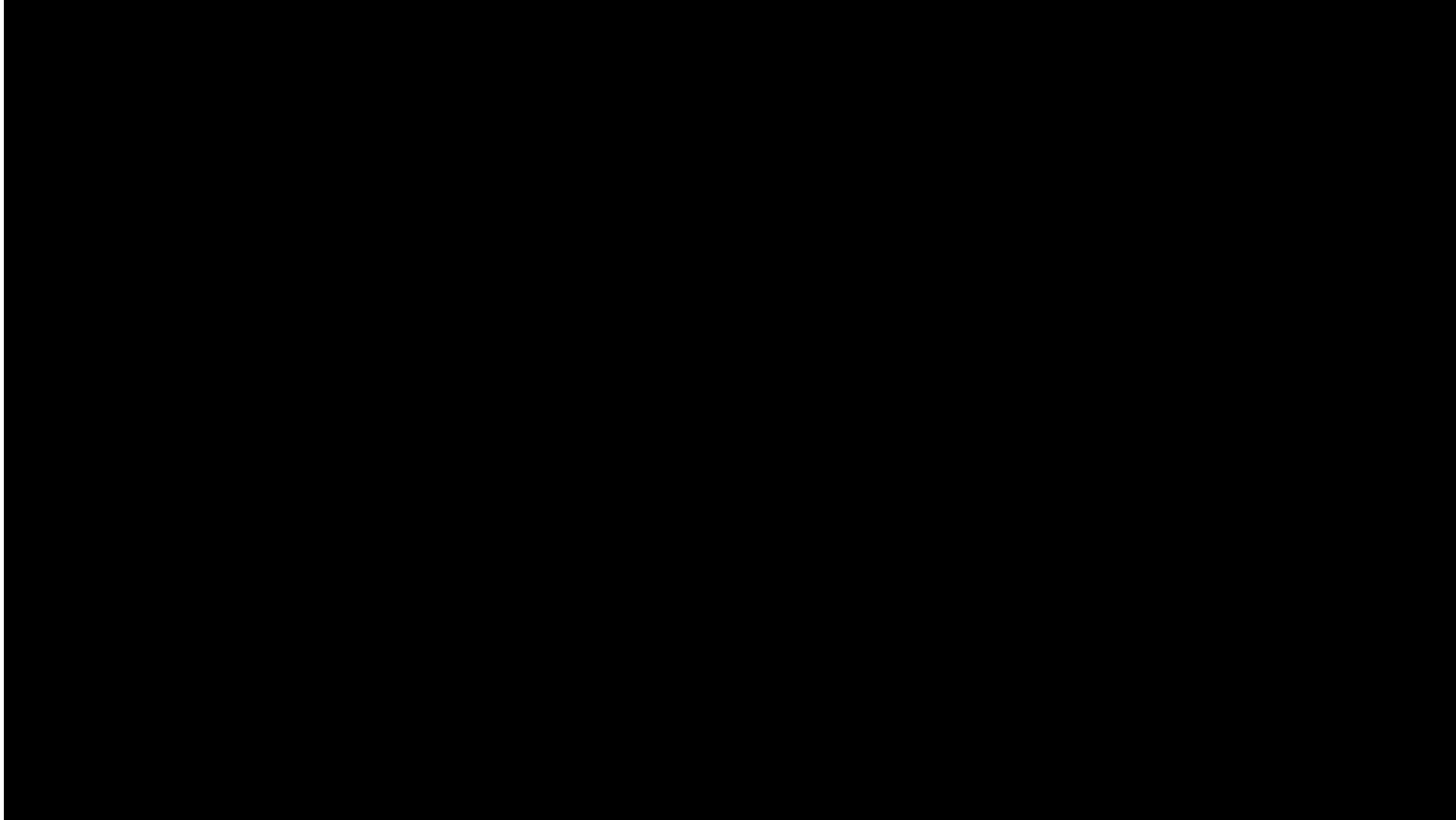
## 2011



<https://www.youtube.com/watch?v=agzItTz35QQ>

# Amazon Alexa

2014



[https://www.youtube.com/watch?v=YvT\\_gqs5ETk](https://www.youtube.com/watch?v=YvT_gqs5ETk)

# Microsoft Cortana

## 2015

The screenshot shows a Microsoft Cortana interface running on a Windows 10 desktop. The taskbar at the bottom includes icons for File Explorer, Edge browser, Task View, and Cortana. The Cortana search bar displays the query "Ask me anything". The main window is a Salesforce dashboard for Gerald Frazer, dated Wednesday September 16, 2015. The dashboard features two main sections: "Current Quarter Sales Predictions" and "Predicted Risk by Team".

**Current Quarter Sales Predictions**

Category	Value	Trend
Sales Won	\$3.3MM	Up
Team Estimate	\$6.3MM	Up
Predicted Sales	\$5MM	Down
Predicted Upside	\$1.3MM	Up
Predicted Risk	\$2.2MM	Up

**Predicted Risk by Team**

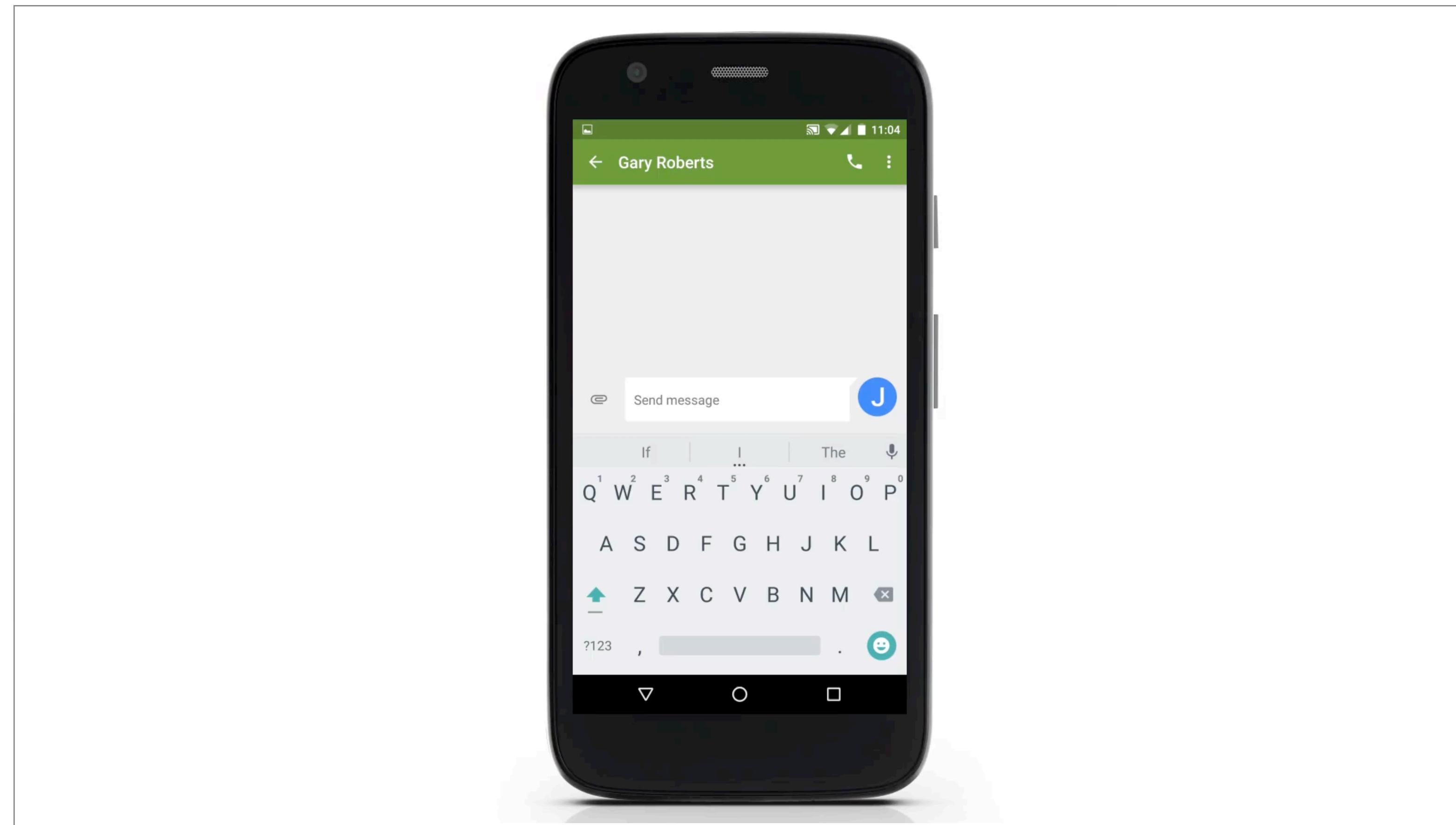
Team	Risk Value
Tom Nettell's Team "Total"	\$450k
Buz Weas's Team "Total"	\$448k
Anna Andover's Team "Total"	\$392k
Jennifer Smith	\$344k
Berji Miller	\$225k
Gerald Frazer	\$148k
Jana Johnson	\$23.2k

At the bottom left, it says "Powered by AlpineMetrics | Microsoft Cortana Analytics Suite". The status bar at the bottom right shows the date as 9/16/2015 and the time as 5:42 PM.

<https://www.youtube.com/watch?v=DDqrfCmIPxI>

# Google GBoard

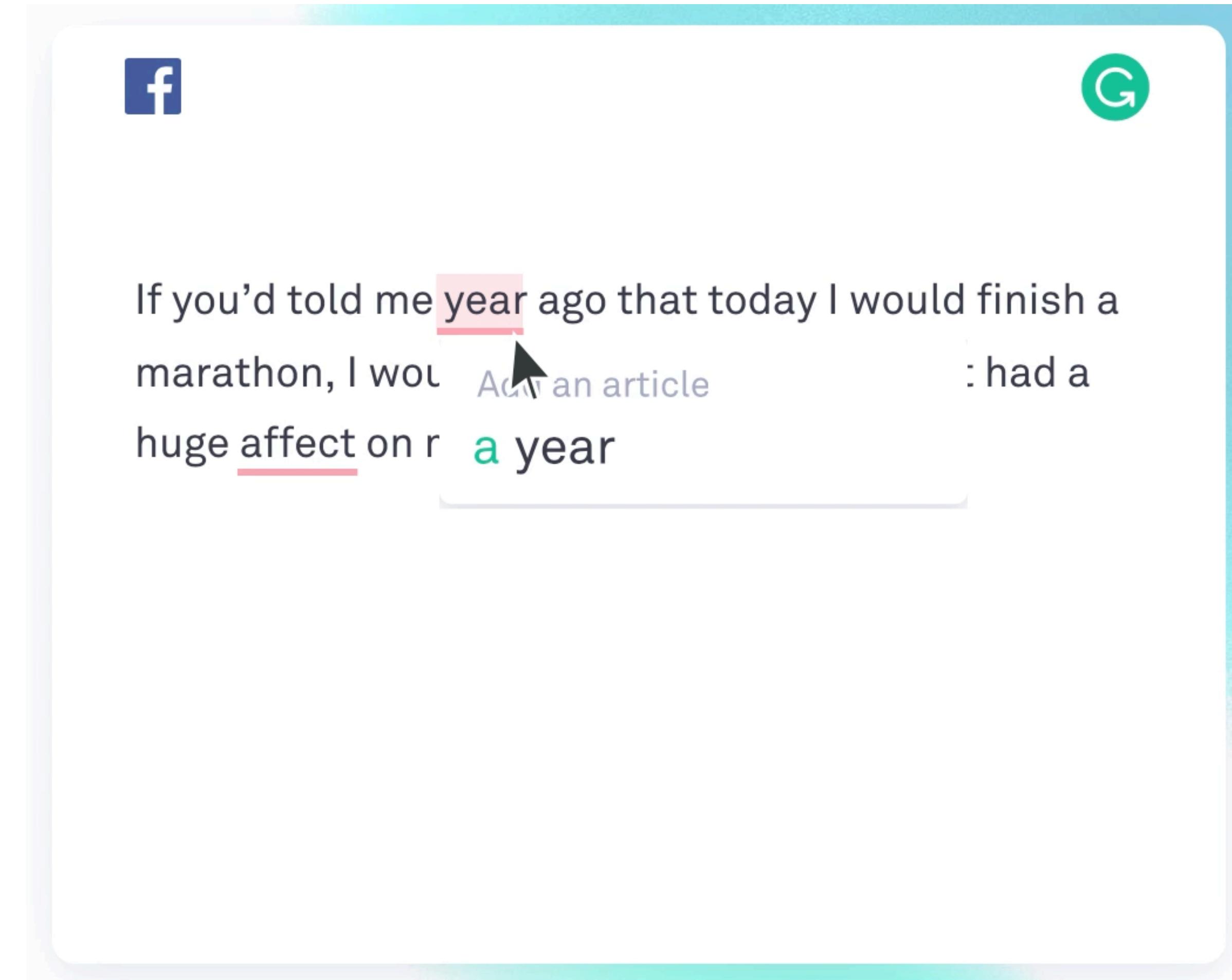
## Predictive Text, Auto-Correct, Glide Typing



<https://www.youtube.com/watch?v=5DSfFDdybzg>

# Grammarly

## Spelling and grammar Correction



- Machine translation
- Automatic summarization
- Stock market prediction
- Anomaly detection
- Controller automation (eg. Marl/O)
- Music composition
- Human-machine co-creation

# Drop Jump Classification

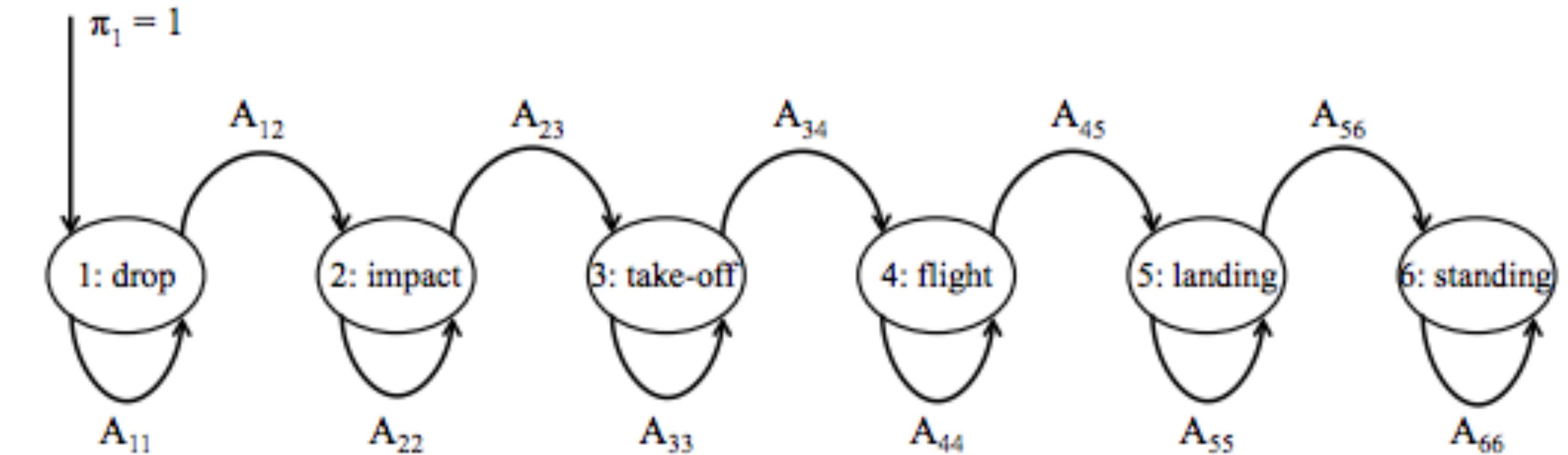


Figure 1. Left-right state transitions of the HMM that was used to analyze a drop jump sequence.

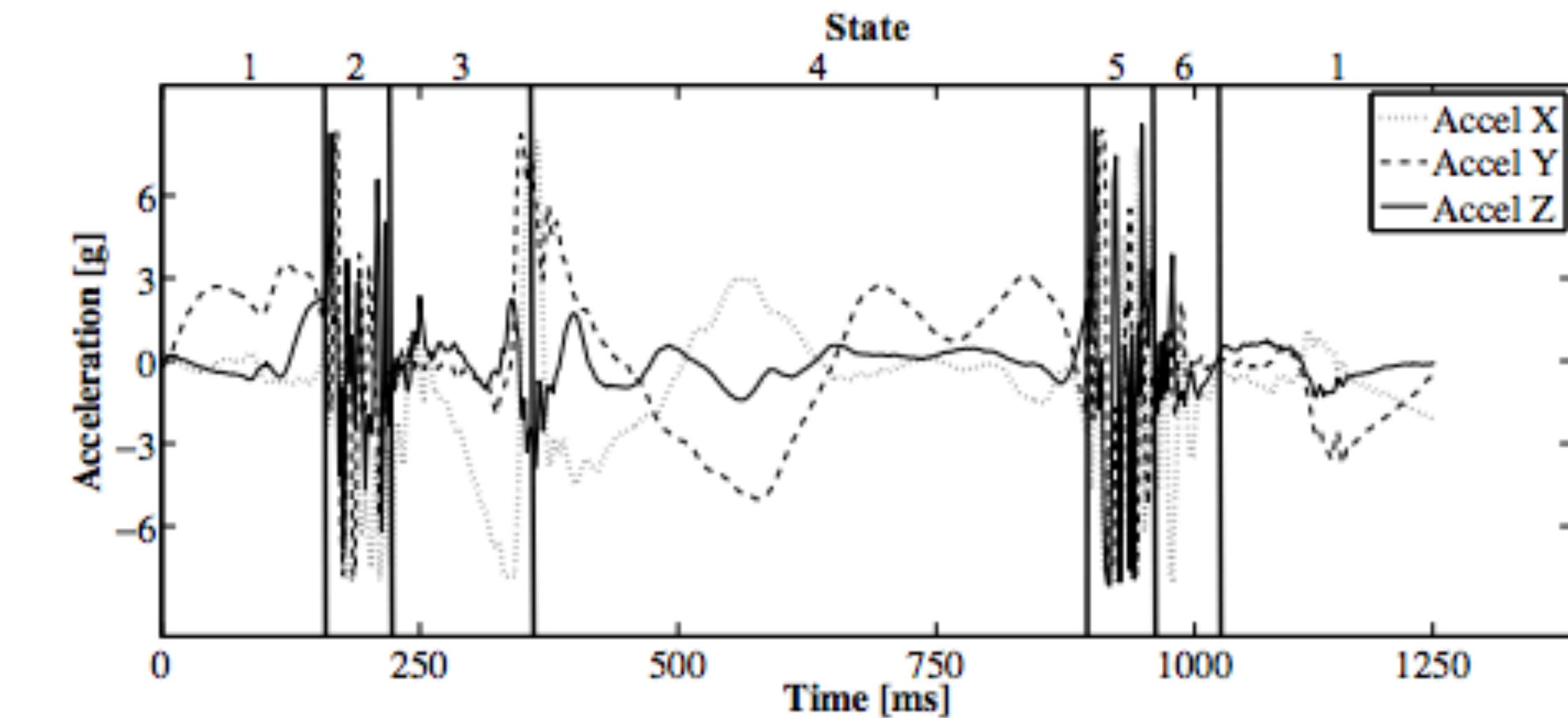
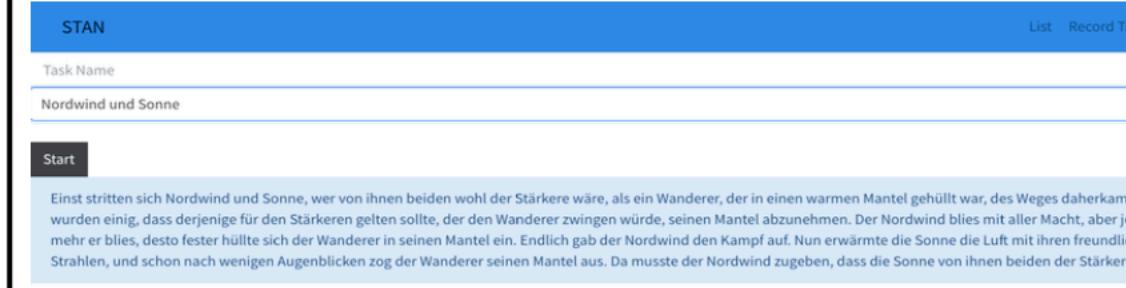
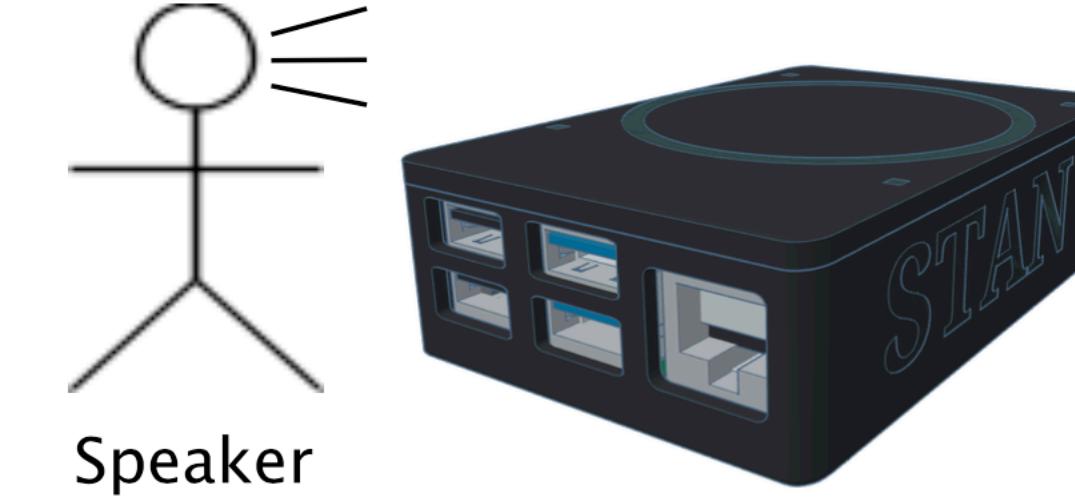
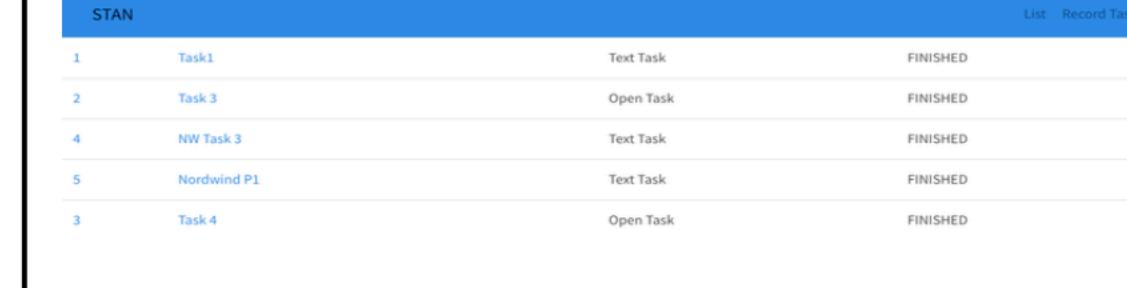


Figure 2. Left foot accelerometer signal of one drop jump and corresponding state segmentation.

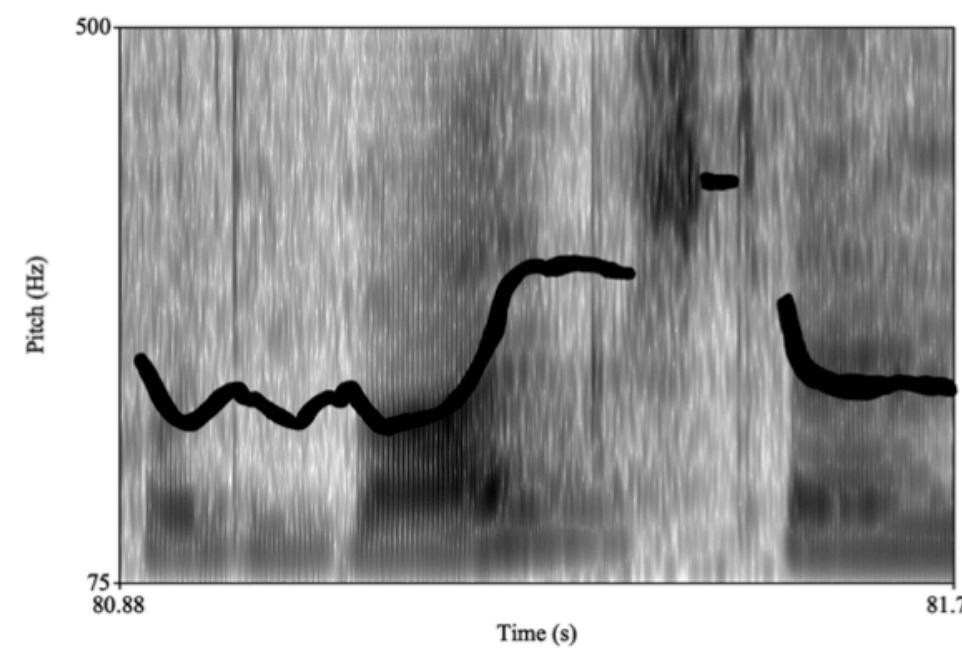
# Stuttering Detection

Ongoing research project (w/ Kasseler Stottertherapie)

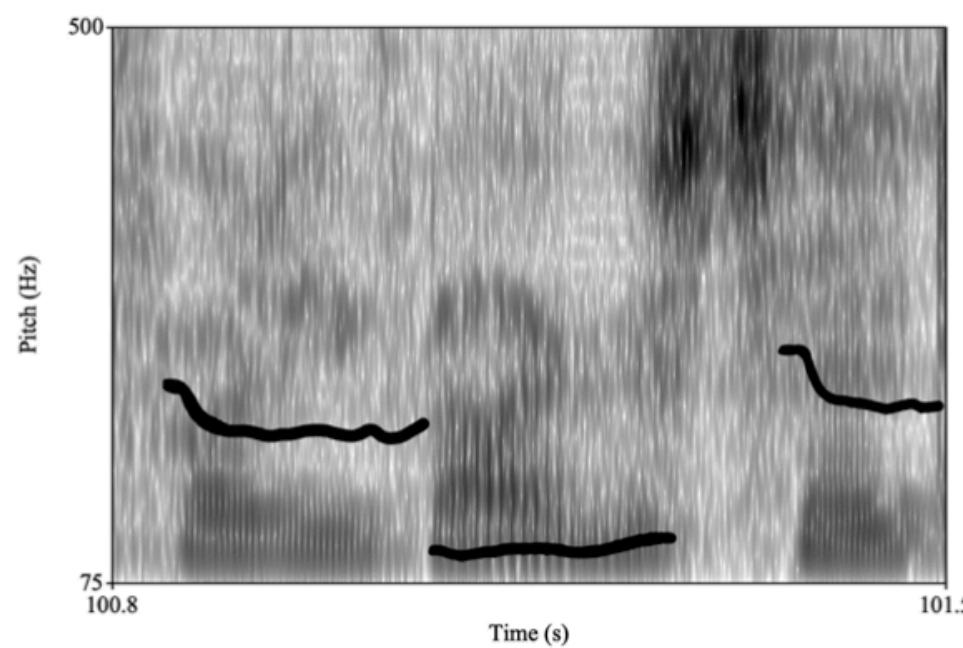
Data Collection	Starting and Stopping of tasks	Record audio	Task List
Analysis	Full details with audio player	Posterogram detail	Spectrogram details
		 Speaker	

# Emotional Carriers

Ongoing research project (w/ U Trento)



(a) emotion carrier



(b) non emotion carrier

Figure 2: Spectrograms with f0-contour (a) showing the bold part of the phrase: “*Dann haben sie mich sozusagen vor die Wahl gestellt, ja, entweder du kannst studieren gehen oder du hast ein Pferd, weil beides kann man nicht finanzieren.*”<sup>1</sup>, which was marked as an EC. (b) was taken from the same recording session, but was not marked as EC.

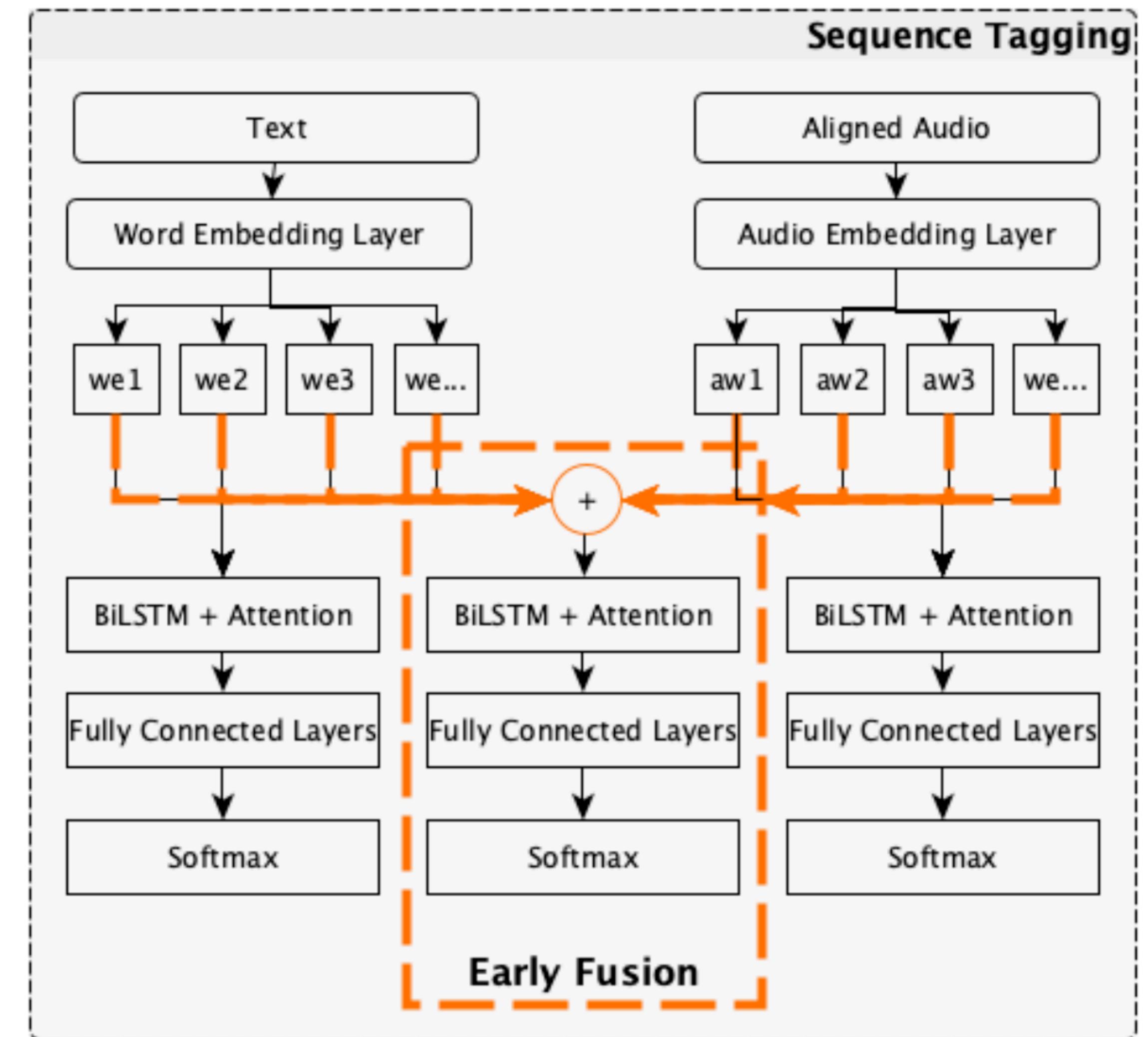
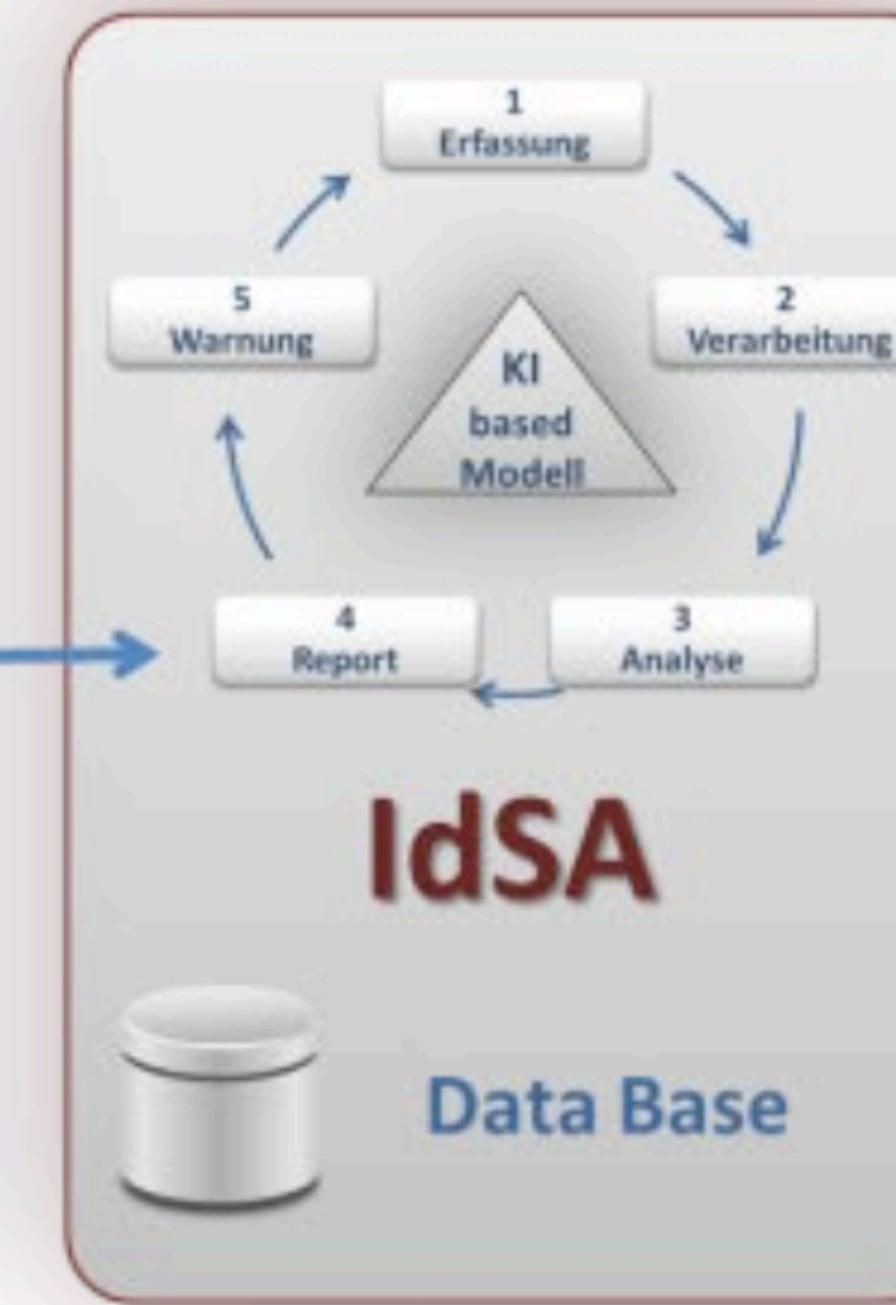
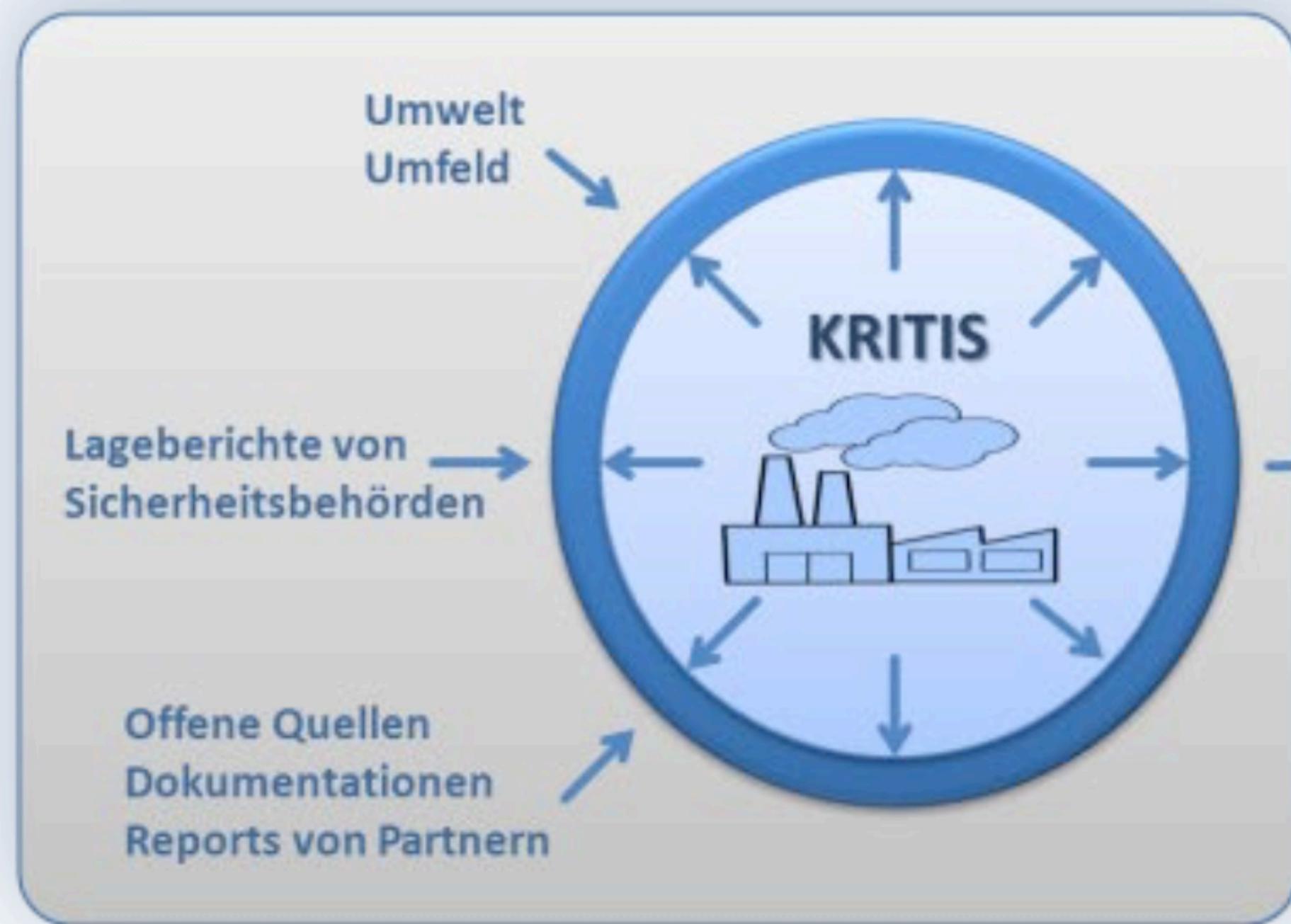


Figure 3: Sequence tagging (ST) architecture that can operate using either Word Embeddings (WE), Audio Word Embeddings (AWE) or a combination of WE and AWE in an early fusion (EF) approach. The embedding layer of the ResNet classifier is used to extract AWE for each word.

# Anomaly Detection

Ongoing research project (w/ ESTW)



Anomalien sind alle Abweichungen von Erwartungen und technischen Vorsorgen.

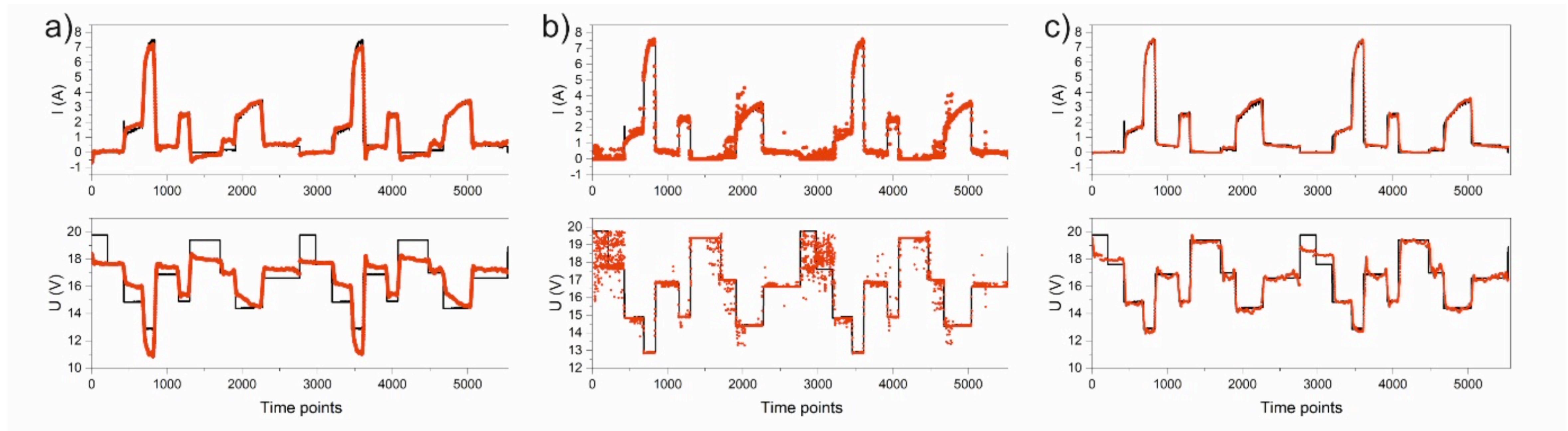
→ Assistenz  
→ Report  
→ Warnung

IdSA generiert Modelle zum Normalbetrieb mittels Training aus umfassenden Daten und Informationen.

IdSA soll Anlagenführer bei der frühzeitigen Erkennung kritischer Situationen unterstützen und helfen, Schaden zu vermeiden.

# Instrumentation and Control Engineering

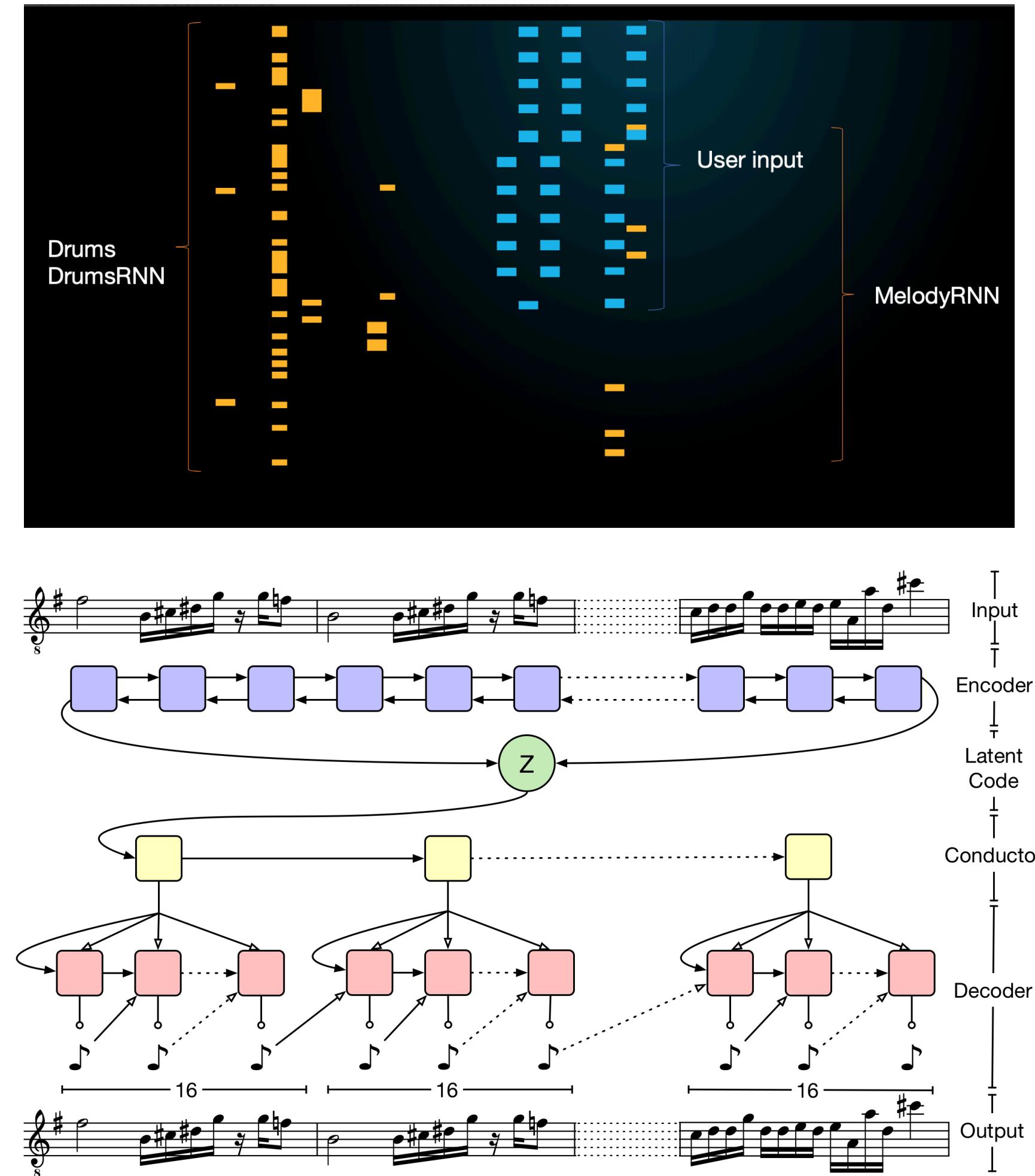
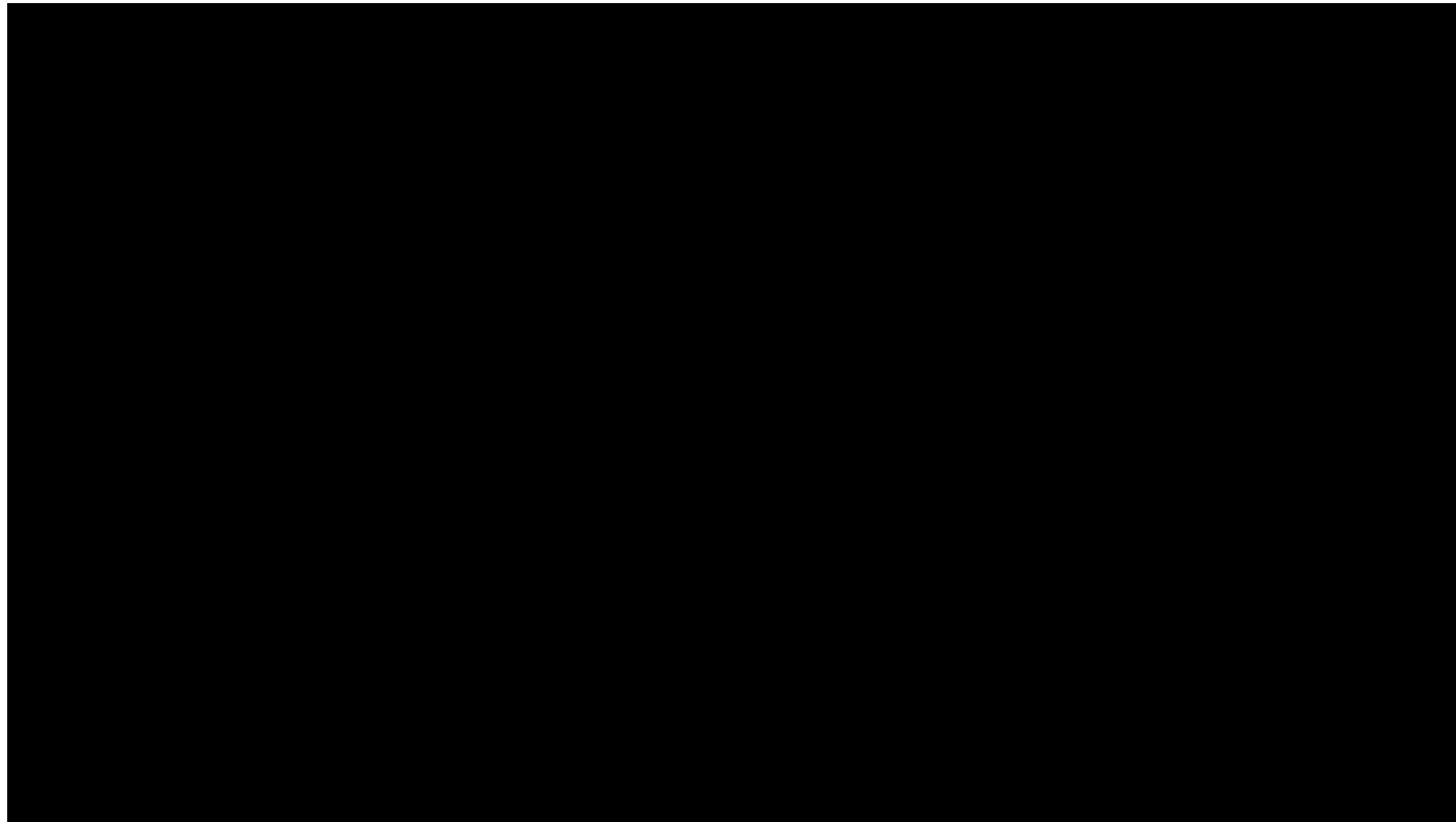
Ongoing research project (w/ AC)



*Figure 2. Calculated (red points) and measured (black lines) output parameters stack current  $I$  and voltage  $U$  of a 100 W fuel cell stack. The calculated values were obtained by training and testing MLR (a), RFR (b) and DNN (c) models with static performance profiles as shown in Figure 1a) using stack temperature and  $H_2$  flow as input parameters. In a statistical error analysis root mean square errors of 0.837 (MLR), 0.616 (RFR) and 0.537 (DNN) were obtained for the calculated output parameters.*

# Spirio Sessions: Human-Machine Co-Creation

*Ongoing research project (w/ HfM Nürnberg)*



<https://magenta.tensorflow.org/music-vae>

# Data Sources

## Analog signals (discretized)

- Microphones
- Vibrations
- Conductivity
- Ambient: pressure, temperature, humidity, ...
- Positional: GPS, gyro, distances
- User input: key-press, gestures, pressure, swipe, ...

# Data Sources

## Digital or “Big Data” signals

- Log streams
- Network traffic
- Events (IoT, MQTT, ...)
- User-generated content (Twitter, blogs, ...)

# Toolkits

- Python3
- Unittest, code cells (Visual Studio Code)
- numpy/scipy
- PyTorch (basics), Keras (transfer learning)

# Syllabus

- Basic algorithms
  - Matching and comparing (discrete) sequences
  - Dynamic programming
- Statistical modeling
  - Markov chains, hidden Markov models
  - Maximum likelihood, expectation maximisation
- Neural networks
  - Feed-forward and recurrent networks
  - Attention and transformers
  - Transfer learning

# Assignments

- Assignments due...
  - April 11: Dynamic Programming
  - April 25: Markov Chains
  - May 9: Hidden Markov Models
  - May 30: RNN
  - June 13: Attention
  - June 27: Transformers

# What you should bring to this class

- A little bit of probability theory
- A little bit of optimization theory
- Algorithms and programming
- Curiosity and perseverance: understanding is hard, implementing sometimes even harder...