



**Prof. Dr. Andreas Hotho, Albin Zehe, Jonas Kaiser**

# Organisational: Dates

- **Lecture**

- Begin: April 23rd
- Tuesday, 12:30 – 14:00
- SE 2, **CAIDAS Building**

- **Exercises**

- Begin: May 1st
- Wednesday, 12:30 – 14:00 (first session is on a holiday)
- Thursday, 12:30 – 14:00
- SE 2, **CAIDAS Building**

- **Exam**

- Will be announced  
Registration via WueStudy

# Organisational

- Exercise:
  - **independent work** on the exercises (in small groups of 3-4 persons)
  - **questions** can be asked to the assistant
  - **no general repetition** of the lecture material
  - **no presentation** of the sample solution (sample solution will be provided later)
- Necessary for this
  - independent lecture post-processing **before the exercise**
  - be active by your own

# Organisational

- **Why this exercise concept?**

- it is more effective to actively work on the lecture's content
- **recognising connections** in the material
- learning structured thinking and independent working
- learning to **work in teams**
- **exam training** 😊
- *“Your personal strengths are initiative, willingness to communicate and cooperate, and teamwork.”*

(Typical job ad text)

# Organisational

## Contacts:

- Prof. Andreas Hotho: [hotho@informatik.uni-wuerzburg.de](mailto:hotho@informatik.uni-wuerzburg.de) (Room 50.03.004)
- Albin Zehe: [zehe@informatik.uni-wuerzburg.de](mailto:zehe@informatik.uni-wuerzburg.de) (Room 50.03.021)
- Jonas Kaiser: [jonas.kaiser@informatik.uni-wuerzburg.de](mailto:jonas.kaiser@informatik.uni-wuerzburg.de) (Room 50.03.019)

## Information in WueCampus:

- <https://wuecampus.uni-wuerzburg.de/moodle/course/view.php?id=66368>
- There, you can find:
  - current announcements (!)
  - slides
  - exercise sheets
  - literature recommendations
  - dates

# Organisational: Praktikum Machine Learning

- „**Praktikum Machine Learning**“ as practical addition to the lecture
- Separate course
- 5 ECTS
- Content:
  - Applying the techniques learned in the lecture
  - Practical work on large text datasets
  - Testing new ideas and solving problems, maybe resulting in a *published paper*
- Every **winter term**: *NLP* centric task
- **Summer term (this term)**: *Timeseries Analysis*

<https://www.informatik.uni-wuerzburg.de/datascience/lectures/2024-ss/praktikum-machine-learning-for-time-series-analysis/>

## Knowledge-Enriched Natural Language Processing

Deep Learning for  
Symbolic Mathematics

NLP with  
Knowledge Graphs

Sentiment Analysis

Scene Segmentation

Session-based Recommendation

Medical Recommendations

Recommender for Publications

Recommender for Games

## Deep Learning for Recommendation

## Machine Learning Fundamentals

Background Knowledge  
in Machine Learning

Bayesian Hypothesis  
Testing

Semi-Unsupervised  
Learning



Domain-specific neural  
architectures

Pointer Networks

Deep Metric  
Learning

## Physics and Deep Learning

Graph Signal Processing  
for DL

NeuralPDE

Physics-Informed Graph  
Neural Networks

Semi-Unsupervised Learning

Anomaly Detection

Multimodal Data Analysis

## Deep Learning for Imbalanced Data

# What is this lecture about?

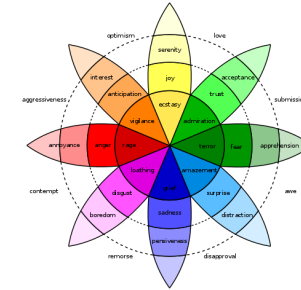
## Machine Learning for Natural Language Processing

- We will cover recent topics in **Machine Learning...**
  - Currently mostly neural network-based methods
- ... in the specific context of **Natural Language Processing**
  - How are different network types used in natural language processing?



# Example Uses of Natural Language Processing

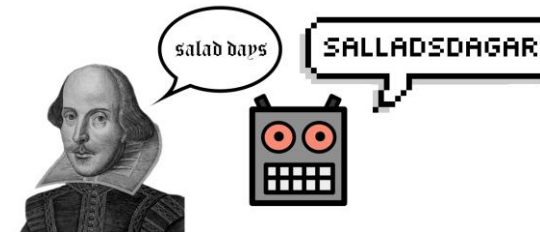
- Sentiment Analysis



- Text Generation



- Machine Translation



# Sentiment Analysis

- Determining positive or negative connotation in text
- Interesting for all kinds of texts
  - Product Reviews

I really enjoyed reading this book!

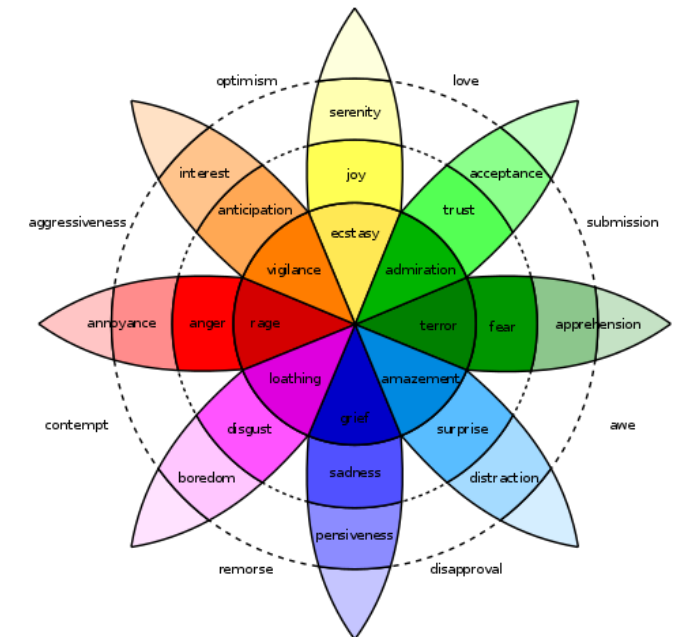
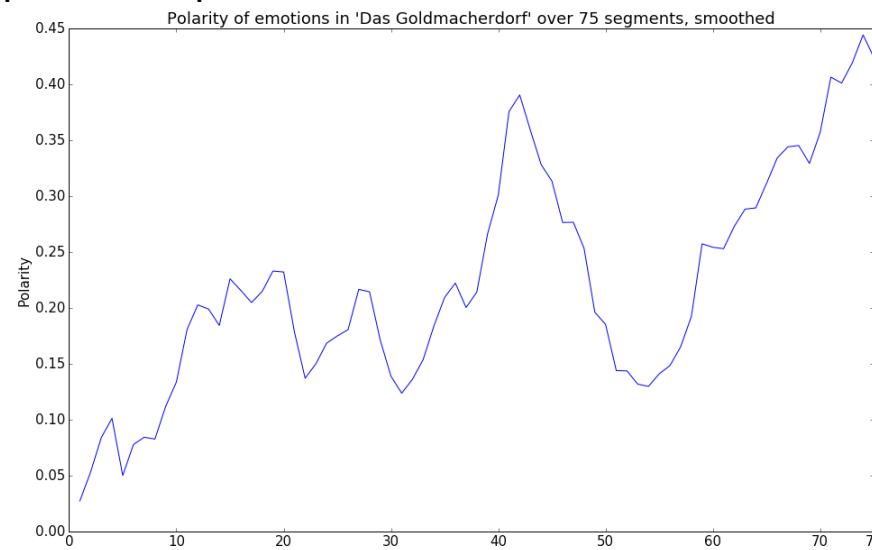


It was just boring.

- Political Comments

I do not agree with this statement at all.

- Characterising plot development in stories



# Sentiment Analysis

- Challenges
  - Negations

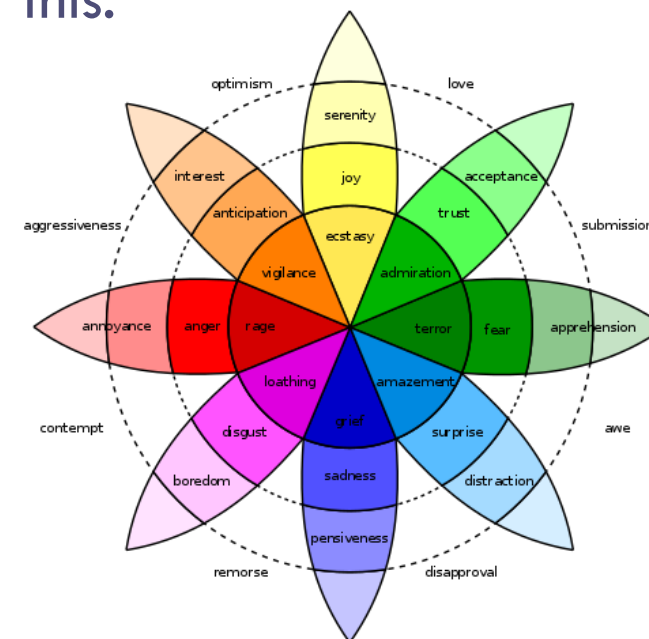
This may or may not be a positive sentence.

I won't ever stop being happy about this.

- Irony/Sarkasm

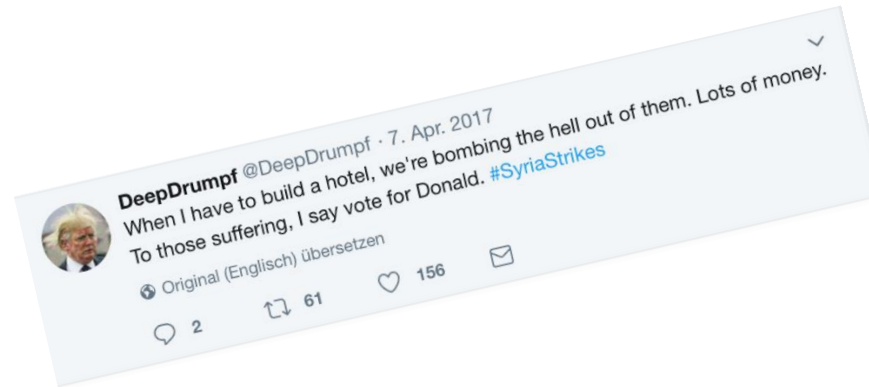
- ...

Isn't this just awesome.



# Text Generation/Language Modelling

- Generate sentences that are grammatically correct and make sense
- Useful for chatbots, automated helpdesks, ...
- ... and for fun 😊



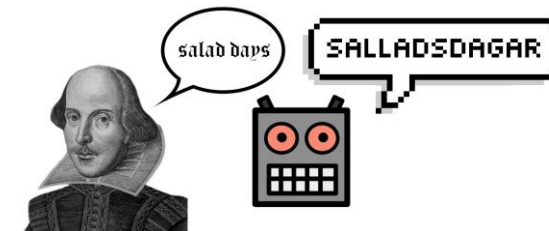
“I've warned you , Snivellus , " said Dumbeldore happily to a thunderstruck Umbridge .

- Perfect text generator → Passes Turing Test!



# Machine Translation

- Given a text in one language, provide a translation to another language
- Combines many hard problems
  - Determine the meaning of the source text
    - Words with multiple meanings
    - Possibly wrong grammar/spelling
  - Generate text in the target language
    - Finding the right words
    - Creating a sentence that makes sense!
- Popular examples: Google Translate / DeepL
  - Based on neural networks

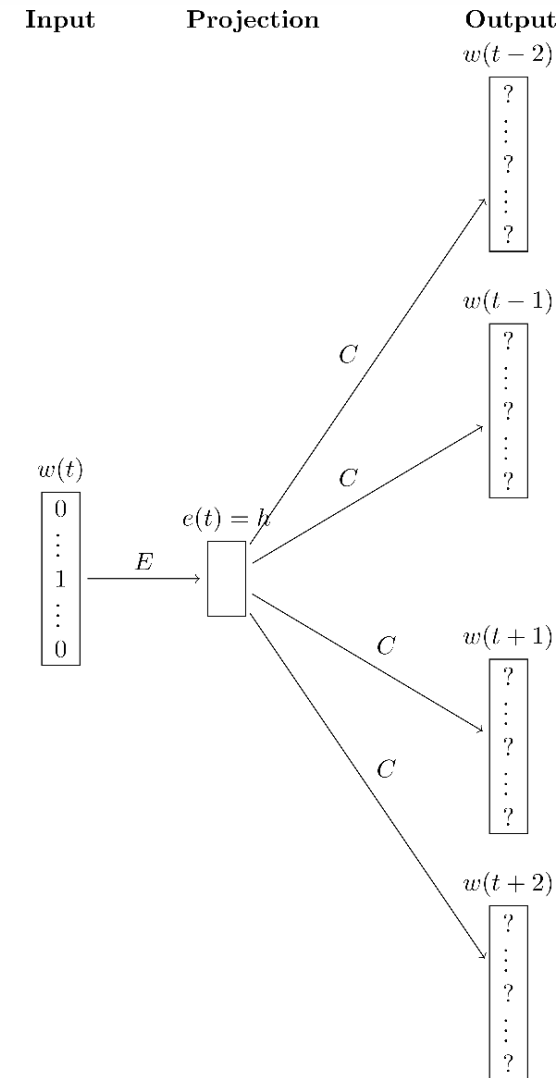


# Overview

1. Representing Words
2. Neural Network Basics
3. Implementing Neural Networks in PyTorch
4. Modelling the Text Stream
  - a) „The new n-grams“ – Convolutional Neural Networks
  - b) Sentences as a sequence — RNNs and their Applications
5. Seq2Seq models and Attention
6. Transformer — Is Attention All You Need?
7. Representing Words <sup>Part 2</sup> — Context
8. Representing Words <sup>Part 3</sup> — World Knowledge
9. Applications

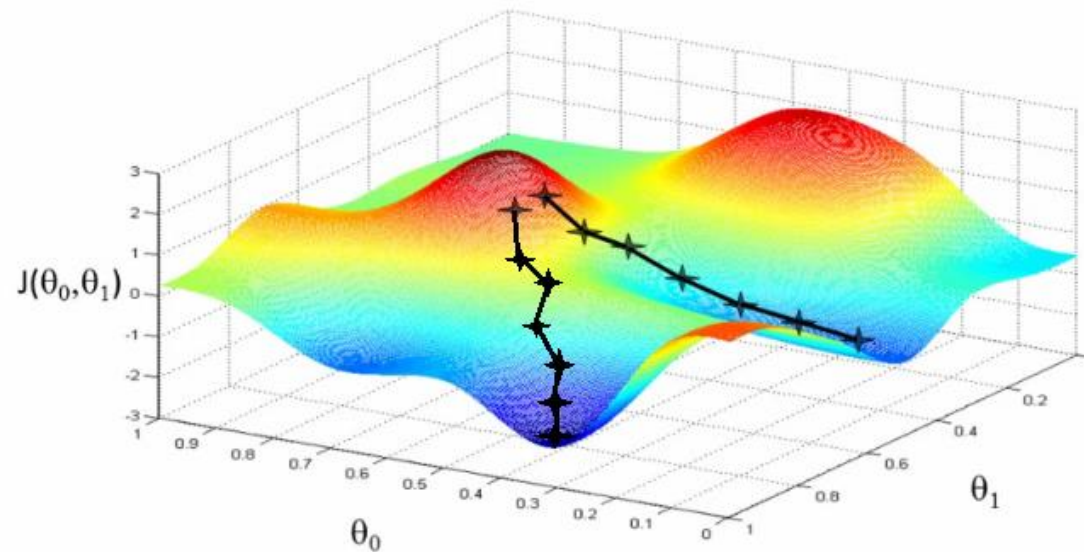
# Representing Words

- Words = basic units in NLP!
- Classic word representations have their limitations
  - How to represent words with similar meanings?
- (Neural) Word Embeddings can solve some of these issues
- Example: Word2Vec



# Neural Network Basics

- Basic Optimisation
  - Backpropagation
  - Gradient Descent
- Implementation in



<http://blog.datumbox.com/wp-content/uploads/2013/10/gradient-descent.png>



# Modelling Sentences – RNNs and their Applications

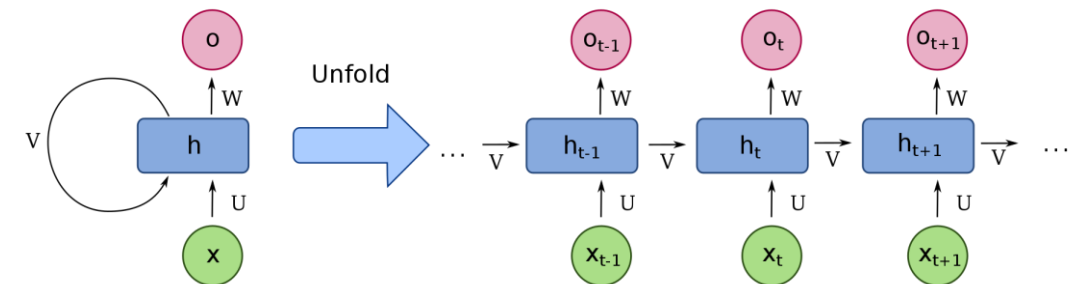
- Modelling words is not enough
- How to represent sentences or longer texts?

→ Recurrent Neural Networks model sequential data!

→ A sentence is a sequence of words

→ A sentence is a sequence of characters

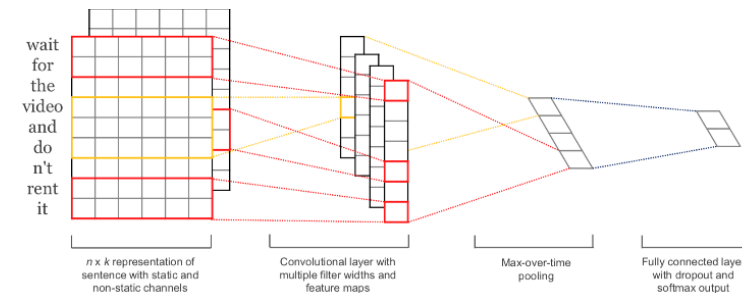
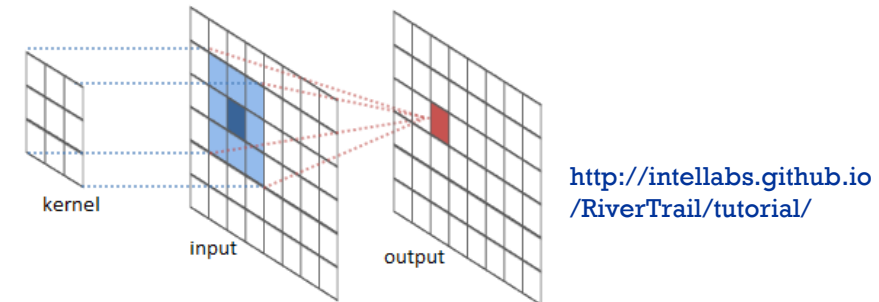
→ ...



[https://commons.wikimedia.org/wiki/File:Recurrent\\_neural\\_network\\_unfold.svg](https://commons.wikimedia.org/wiki/File:Recurrent_neural_network_unfold.svg)

## „The new n-grams“ – Convolutional Neural Networks in NLP

- RNNs are a natural choice for NLP (sequential data)
  - CNNs work very well, too!
  - CNNs are based on „filters“/“kernels“, i.e. local groups of input tokens
  - Originally from computer vision
- 
- Can be seen as analogous to n-grams, which traditionally work very well



# Our Work with Convolutional Neural Networks in NLP

## Sentiment Analysis on Twitch.tv chat messages

Konstantin Kobs, Albin Zehe, Armin Bernstetter, Julian Chibane, Jan Pfister, Julian Tritscher, and Andreas Hotho. 2020.

*Emote-Controlled: Obtaining Implicit Viewer Feedback Through Emote-Based Sentiment Analysis on Comments of Popular Twitch.tv Channels.*

*Trans. Soc. Comput. 3, 2, Article 7 (April 2020), 34 pages.*

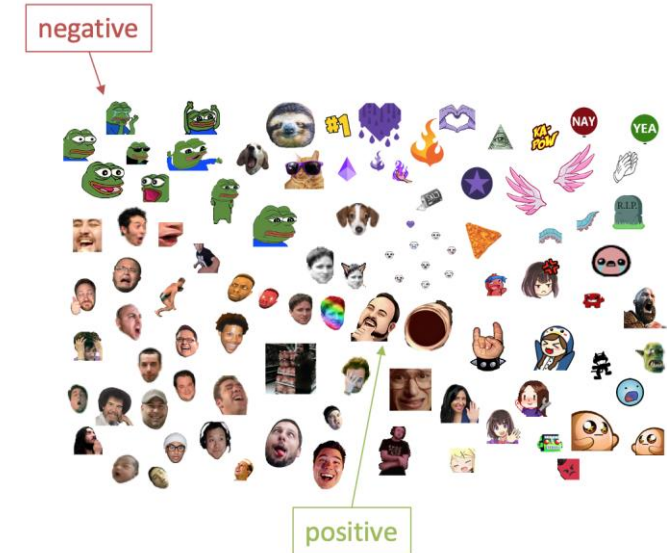
DOI:<https://doi.org/10.1145/3365523>

## Finding fitting conferences or journals for research papers

### **BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding**

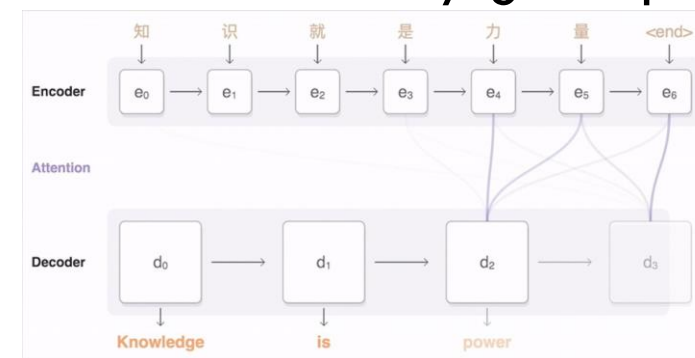
We introduce a new language representation model called BERT, which stands for Bidirectional Encoder Representations from Transformers. . . .

—————→ NAACL



# Seq2Seq Models and Attention

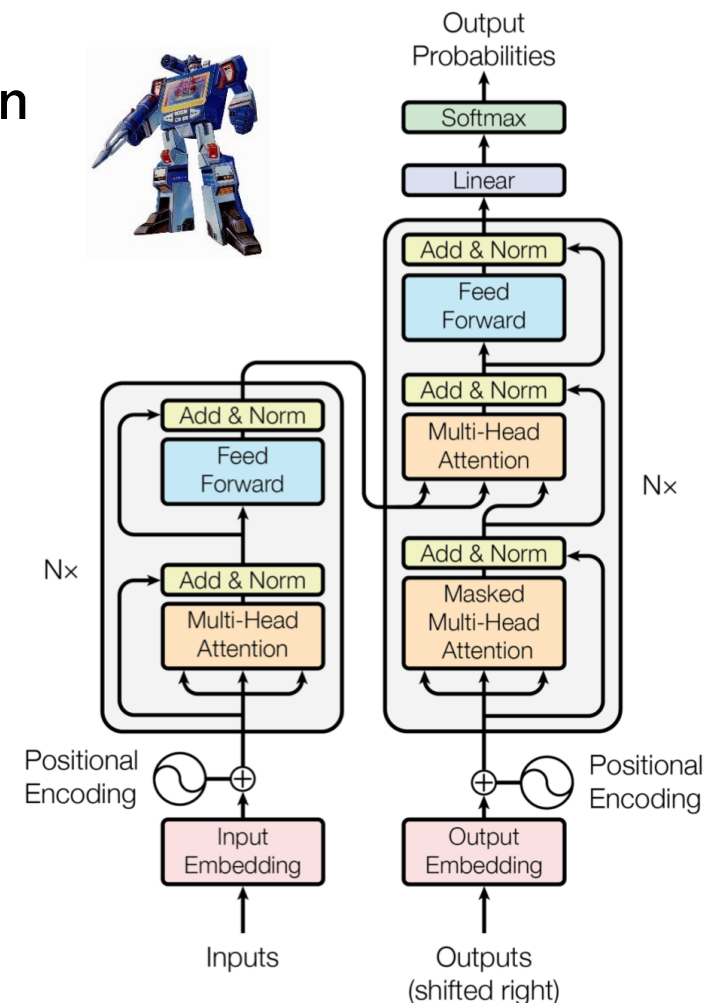
- Encoder – Decoder – Architectures
  - Encoder takes the source text and builds an internal representation
  - Decoder generates target text based on this internal representation
- Great progress in the past years
  - Attention mechanisms help to focus on the most relevant words at any given position
  - Neural machine translation overtakes traditional machine translation



<https://github.com/google/seq2seq>

# Transformer — Is Attention All You Need?

- A novel neural network architecture based on attention
- Encoder-Decoder structure, but no recurrence!
  - Parallelizable → faster to train
- The encoded sentence is as long as the input sentence
  - Capturing more information of input
  - „Transforms“ the input into an encoded form
- Many state-of-the-art improvements!



# Representing Words Part 2 — Context

- Many words have a different meaning depending on the context they are used in
  - e.g. „stick“, „bank“, ...
- Idea: Also consider the context a word is used in to calculate its word representation
  - Using RNNs: e.g. ELMo
  - Using Transformers: e.g. BERT



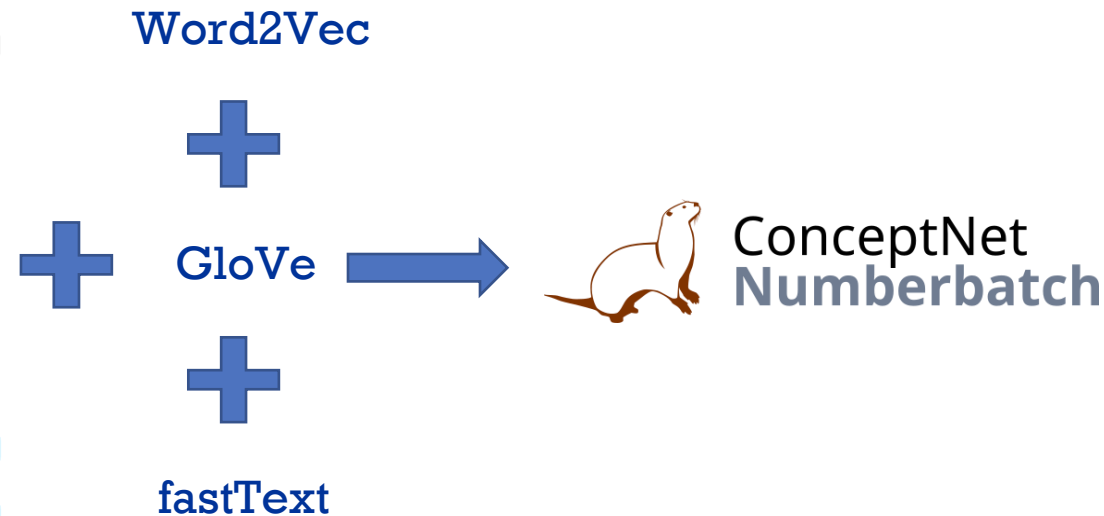
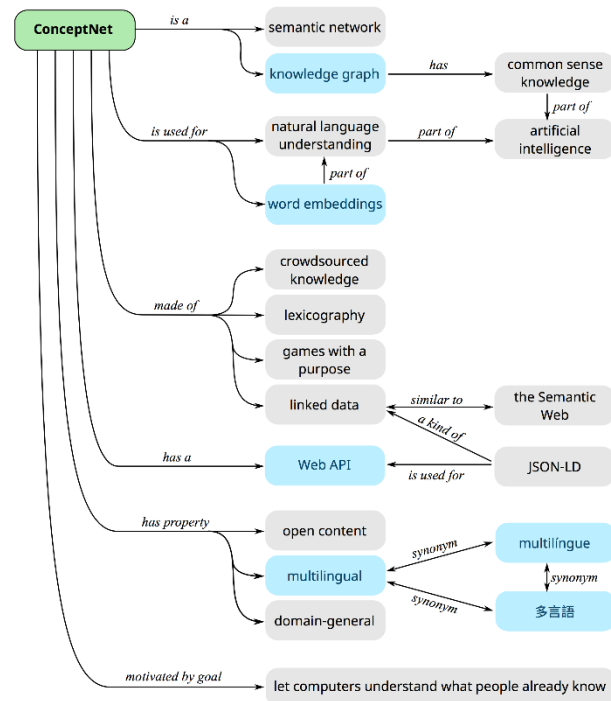
[Source](#)



[Source](#)

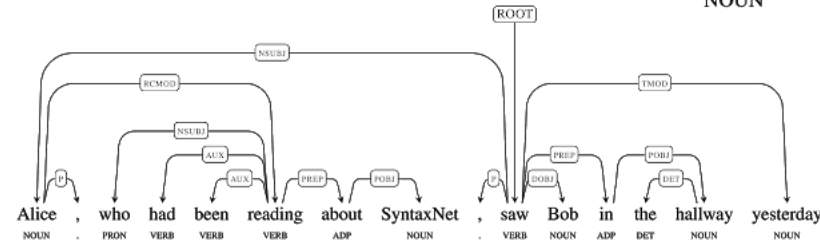
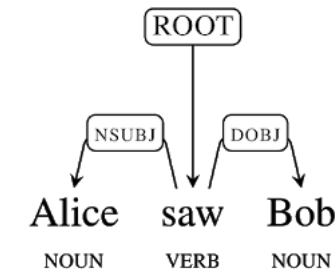
# Representing Words <sup>Part 3</sup> — World Knowledge

- Word2Vec still has shortcomings
- Idea: Integrating **world knowledge** → e.g. **Conceptnet Numberbatch**



# Applications

- SyntaxNet
  - Part-of-speech (POS) tagging: Noun, verb, ...
  - Dependency parsing: „Alice“ is the subject of „saw“, ...
  - Generally a very complex task



- GPT-3
  - A very complex and big (**175 billion parameters!**) **language model** using the transformer architecture
  - Objective: Given a text, predict the next word
  - Very convincing output
  - Can solve many tasks without further training, e.g. question answering, machine translation, ...

<https://twitter.com/quasimondo/status/1284509525500989445>

## The importance of being on twitter

by Jerome K. Jerome  
London, Summer 1897

It is a curious fact that the last remaining form of social life in which the people of London are still interested is Twitter. I was struck with this curious fact when I went on one of my periodical holidays to the sea-side, and found the whole place twittering like a starling-cage. I called it an anomaly, and it is.

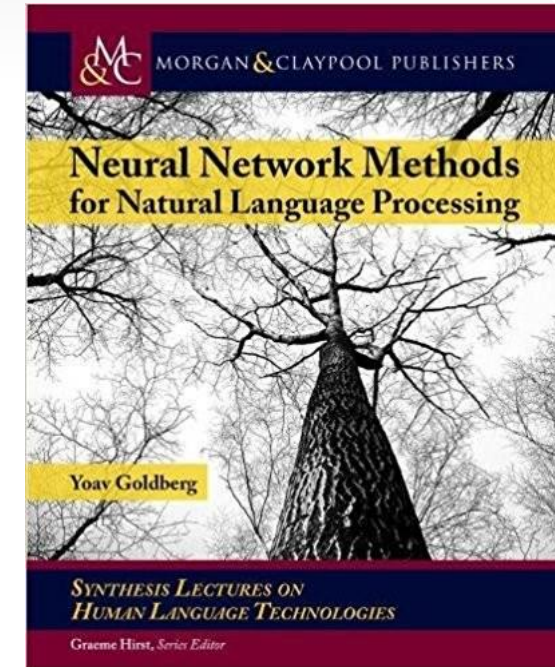
I spoke to the sexton, whose cottage, like all sexton's cottages, is full of antiquities and interesting relics of former centuries. I said to him, "My dear sexton, what does all this twittering mean?" And he replied, "Why, sir, of course it means Twitter." "Ah!" I said, "I know about that. But what is Twitter?"

"It is a system of short and pithy sentences strung together in groups, for the



# Literature

**Neural Network Methods for Natural Language Processing, 2017** by [Yoav Goldberg](#) (Author), [Graeme Hirst](#) (Editor)



A Primer on Neural Network Models  
for Natural Language Processing, Yoav  
Goldberg

<https://u.cs.biu.ac.il/~yogo/nnlp.pdf>

# Other Resource: Twitter

NLP moves very fast nowadays → Follow researchers on Twitter for new interesting papers and discussions

## Some recommendations

- Sebastian Ruder: [@seb\\_ruder](#)
- Chris Manning: [@chrmanning](#)
- Yoav Goldberg: [@yoavgol](#)
- Emily M. Bender: [@emilymbender](#)
- Andrej Karpathy: [@karpathy](#)
- Anna Rogers: [@annargrs](#)
- ...