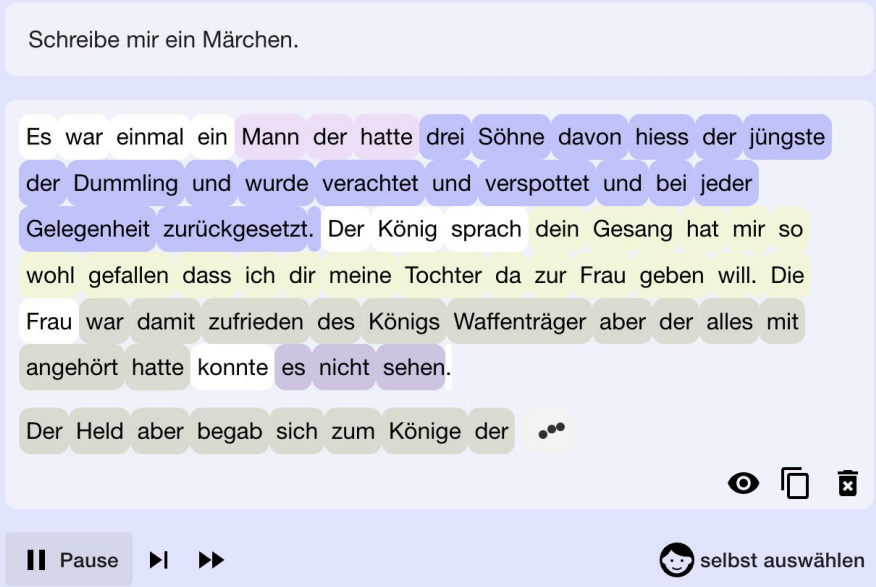
phsz

Soekia

GPT ein didaktisches Sprachmodell



oder

Wie man im Unterricht (Sek I / II) das Phänomen “Textgenerator” erklären kann.

Prof. Dr. Michael Hielscher

Pädagogische Hochschule Schwyz

Ausbildung: Kindergarten & Primarstufe
Weiterbildung: Kindergarten & Primarstufe, Sek 1

ca. 350 Studierende

Das Phänomen Textgeneratoren

ChatGPT

30. November 2022

Copilot

15. Januar 2024

Bard

21. März 2023

StableLM

20. April 2023

Das Phänomen Textgeneratoren

ChatGPT

30. November 2022

Copilot

Gemini

21. März 2023

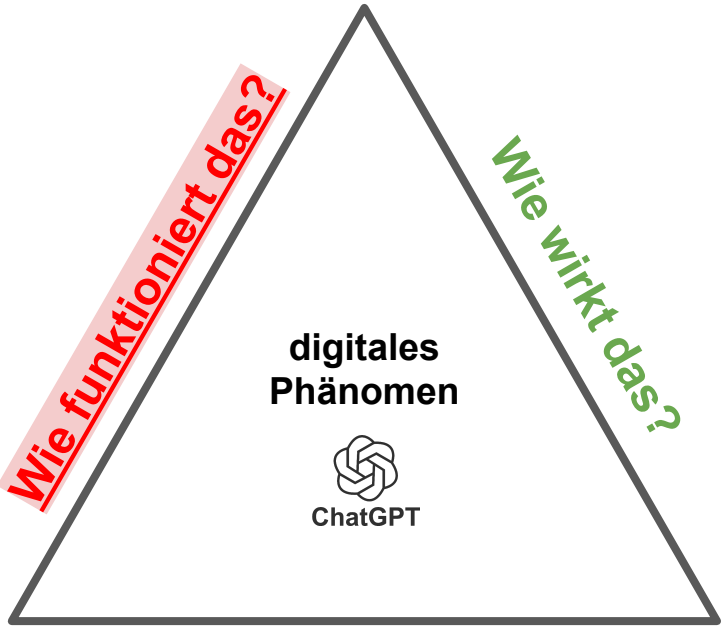
StableLM

20. April 2023

Konzepte hinter digitalen Alltagsphänomenen erklären …

https://dagstuhl.gi.de/dagstuhl-erklaerung

Konzepte hinter digitalen Alltagsphänomenen erklären …



…

Basis der bekannten Textgeneratoren: Statistische Sprachmodelle

Rotkäppchen
Es war einmal ein kleines süsses Mädchen, das hatte jedermann lieb …

ChatGPT
Bard
StableLM

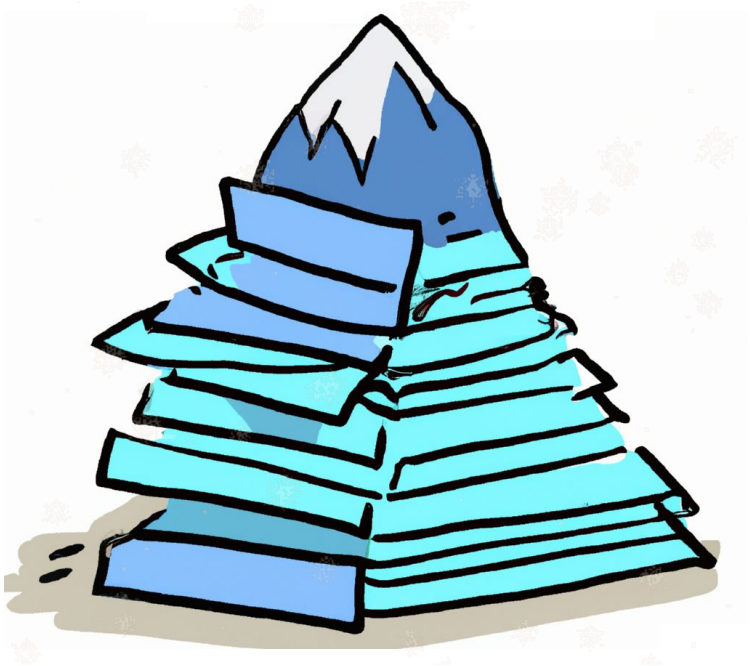
Dokumentenkollektion (Trainingsdaten)

Statistische Analyse (maschinelles Lernen)

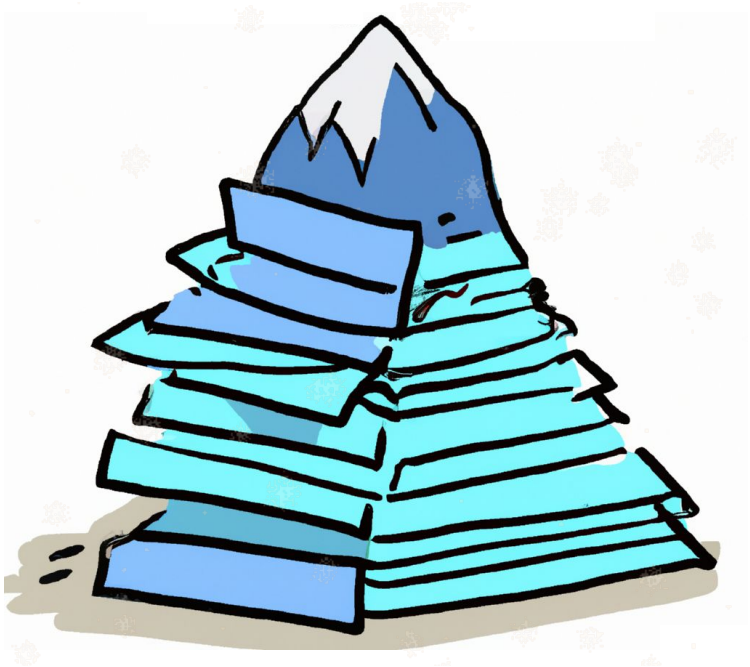
Textgenerator (statistisches Sprachmodell)

Dokumentenkollektion (Trainingsdaten)

Texte aus Büchern, Wikipedia, Internet



Dokumentenkollektion (Trainingsdaten)
Texte aus Büchern, Wikipedia, Internet



ChatGPT

Bei ChatGPT 3 sind das etwa 570GB an Texten aus Büchern, Wikipedia und dem Internet gewesen - alles was man finden konnte. Entspricht etwa 300 Milliarden Wörtern.

Dokumente werden zerlegt in kleine Teile und die statistischen Zusammenhänge gespeichert.

Anfang	*nächstes Wort*	*Anzahl*
Es war einmal	ein	3x
Es war einmal	eine	1x
Es war einmal ein	kleines	1x

Im statistischen Sprachmodell stecken keine “ganzen Kopien” von Texten!

Trainingsdaten

Rotkäppchen
Es war einmal ein kleines süsses Mädchen, das hatte jedermann lieb, der sie nur ansah, am allerliebsten aber ihre Großmutter, die wusste gar nicht was sie alles dem Kinde geben sollte. Einmal schenkte sie ihm ein Käppchen aus rotem Samt …

Jorinde und Joringel
Es war einmal ein altes Schloss mitten in einem großen dicken Wald, darinnen wohnte eine alte Frau ganz allein. Sie war eine Erzzauberin. Am Tage machte sie sich zur Katze oder zur Nachteule, des Abends aber …

Rumpelstilzchen
Es war einmal ein Müller, der war arm, aber er hatte eine schöne Tochter. Nun traf es sich, dass er mit dem König zu sprechen kam, und um sich ein Ansehen zu geben, sagte er zu ihm „ich habe eine Tochter, die kann Stroh zu Gold spinnen.“ …

Rapunzel
Es war einmal eine Frau und ein Mann, die wünschten sich schon lange vergeblich ein Kind, endlich machte sich die Frau Hoffnung der liebe Gott werde ihren Wunsch erfüllen. Die Leute hatten in ihrem Hinterhaus ein kleines Fenster…

Trainingsdaten

Rotkäppchen
Es war einmal ein kleines süsses Mädchen, das hatte jedermann lieb, der sie nur ansah, am allerliebsten aber ihre Großmutter, die wusste gar nicht was sie alles dem Kinde geben sollte. Einmal schenkte sie ihm ein Käppchen aus rotem Samt …

Jorinde und Joringel
Es war einmal ein altes Schloss mitten in einem großen dicken Wald, darinnen wohnte eine alte Frau ganz allein. Sie war eine Erzzauberin. Am Tage machte sie sich zur Katze oder zur Nachteule, des Abends aber …

Rumpelstilzchen
Es war einmal ein Müller, der war arm, aber er hatte eine schöne Tochter. Nun traf es sich, dass er mit dem König zu sprechen kam, und um sich ein Ansehen zu geben, sagte er zu ihm „ich habe eine Tochter, die kann Stroh zu Gold spinnen.“ …

Rapunzel
Es war einmal eine Frau und ein Mann, die wünschten sich schon lange vergeblich ein Kind, endlich machte sich die Frau Hoffnung der liebe Gott werde ihren Wunsch erfüllen. Die Leute hatten in ihrem Hinterhaus ein kleines Fenster…

Trainingsdaten

Rotkäppchen
Es war einmal ein kleines süsses Mädchen, das hatte jedermann lieb, der sie nur ansah, am allerliebsten aber ihre Großmutter, die wusste gar nicht was sie alles dem Kinde geben sollte. Einmal schenkte sie ihm ein Käppchen aus rotem Samt. …

Jorinde und Joringel
Es war einmal ein altes Schloss mitten in einem großen dicken Wald, darinnen wohnte eine alte Frau ganz allein. Sie war eine Erzzauberin. Am Tage machte sie sich zur Katze oder zur Nachteule, des Abends aber …

Rumpelstilzchen
Es war einmal ein Müller, der war arm, aber er hatte eine schöne Tochter. Nun traf es sich, dass er mit dem König zu sprechen kam, und um sich ein Ansehen zu geben, sagte er zu ihm „ich habe eine Tochter, die kann Stroh zu Gold spinnen.“ …

Rapunzel
Es war einmal eine Frau und ein Mann, die wünschten sich schon lange vergeblich ein Kind, endlich machte sich die Frau Hoffnung der liebe Gott werde ihren Wunsch erfüllen. Die Leute hatten in ihrem Hinterhaus ein kleines Fenster…

statistisches Modell

Anfang	*nächstes Wort*	*Anzahl*
Es war einmal	ein	3x
Es war einmal	eine	1x

Trainingsdaten

Rotkäppchen
Es war einmal ein kleines süsses Mädchen, das hatte jedermann lieb, der sie nur ansah, am allerliebsten aber ihre Großmutter, die wusste gar nicht was sie alles dem Kinde geben sollte. Einmal schenkte sie ihm ein Käppchen aus rotem Samt. …

Jorinde und Joringel
Es war einmal ein altes Schloss mitten in einem großen dicken Wald, darinnen wohnte eine alte Frau ganz allein. Sie war eine Erzzauberin. Am Tage machte sie sich zur Katze oder zur Nachteule, des Abends aber …

Rumpelstilzchen
Es war einmal ein Müller, der war arm, aber er hatte eine schöne Tochter. Nun traf es sich, dass er mit dem König zu sprechen kam, und um sich ein Ansehen zu geben, sagte er zu ihm „ich habe eine Tochter, die kann Stroh zu Gold spinnen.“ …

Rapunzel
Es war einmal eine Frau und ein Mann, die wünschten sich schon lange vergeblich ein Kind, endlich machte sich die Frau Hoffnung der liebe Gott werde ihren Wunsch erfüllen. Die Leute hatten in ihrem Hinterhaus ein kleines Fenster…

statistisches Modell

Anfang	*nächstes Wort*	*Anzahl*
Es war einmal	ein	3x
Es war einmal	eine	1x
Es war einmal ein	kleines	1x
Es war einmal ein	altes	1x
Es war einmal ein	Müller	1x

Trainingsdaten

Rotkäppchen
Es war einmal ein kleines süsses Mädchen, das hatte jedermann lieb, der sie nur ansah, am allerliebsten aber ihre Großmutter, die wusste gar nicht was sie alles dem Kinde geben sollte. Einmal schenkte sie ihm ein Käppchen aus rotem Samt. …

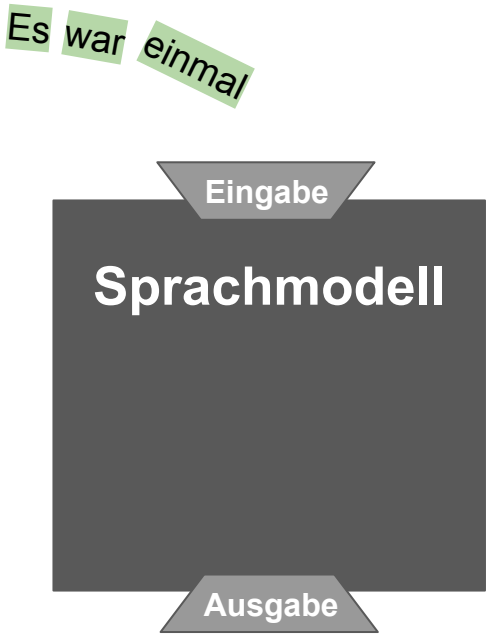
Jorinde und Joringel
Es war einmal ein altes Schloss mitten in einem großen dicken Wald, darinnen wohnte eine alte Frau ganz allein. Sie war eine Erzzauberin. Am Tage machte sie sich zur Katze oder zur Nachteule, des Abends aber …

Rumpelstilzchen
Es war einmal ein Müller, der war arm, aber er hatte eine schöne Tochter. Nun traf es sich, dass er mit dem König zu sprechen kam, und um sich ein Ansehen zu geben, sagte er zu ihm „ich habe eine Tochter, die kann Stroh zu Gold spinnen.“ …

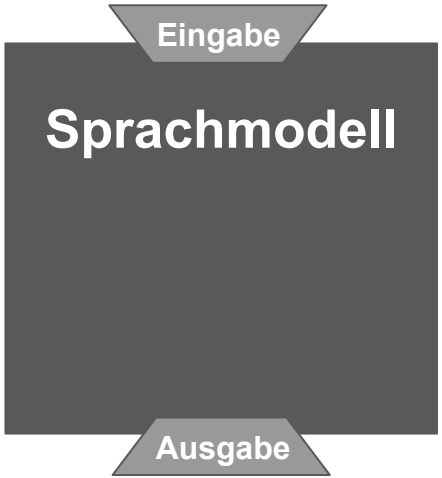
Rapunzel
Es war einmal eine Frau und ein Mann, die wünschten sich schon lange vergeblich ein Kind, endlich machte sich die Frau Hoffnung der liebe Gott werde ihren Wunsch erfüllen. Die Leute hatten in ihrem Hinterhaus ein kleines Fenster…

statistisches Modell

Anfang	*nächstes Wort*	*Anzahl*
Es war einmal	ein	3x
Es war einmal	eine	1x
Es war einmal ein	kleines	1x
Es war einmal ein	altes	1x
Es war einmal ein	Müller	1x
ein	kleines	2x
ein	altes	1x
ein	Müller	1x
ein	Ansehen	1x
ein	Mann	1x
ein	Kind	1x

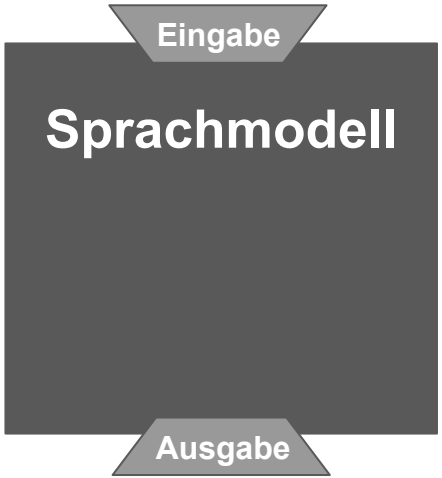

Es war einmal
Eingabe
Sprachmodell
Ausgabe

Es war einmal



ein	=	0.75
eine	=	0.25
…		
Bauer	=	0.00
Müller	=	0.00

Es war einmal



ein	=	0.75
eine	=	0.25
…		
Bauer	=	0.00
Müller	=	0.00

ChatGPT

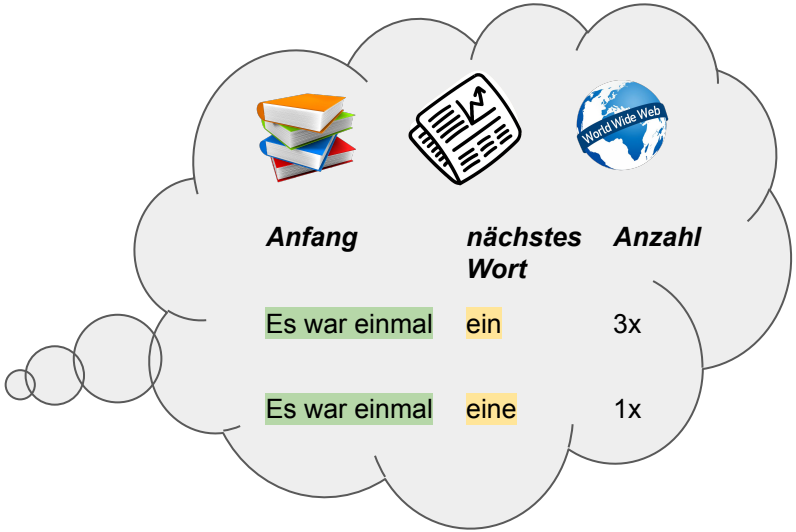
Das GPT-3 Modell kennt **50,257** verschiedene Wörter (Tokens) (vocabulary size)

Es war einmal

Eingabe

Sprachmodell

Ausgabe



ein	=	0.75
eine	=	0.25
…		
Bauer	=	0.00
Müller	=	0.00

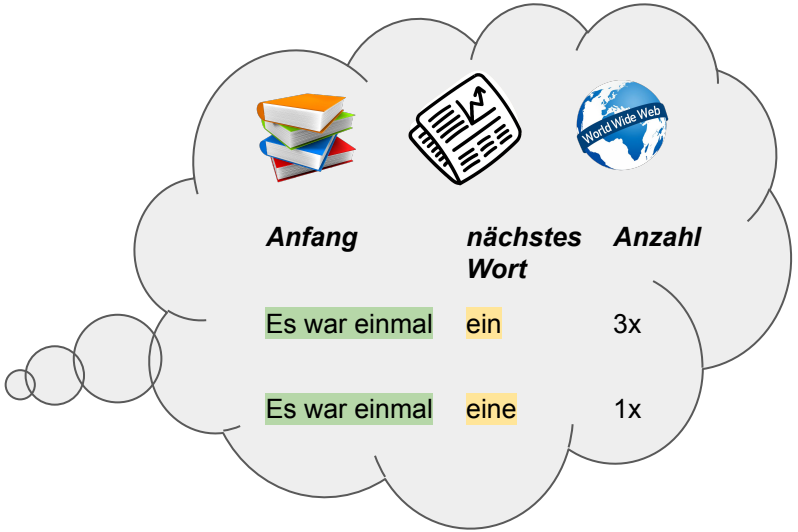
Es war einmal

Eingabe

Sprachmodell

verschiedene
Modelle / Verfahren
z.B.
n-Gramm Listen,
neuronale Netzwerke

Ausgabe



ein	=	0.75
eine	=	0.25
…		
Bauer	=	0.00
Müller	=	0.00

Es war einmal

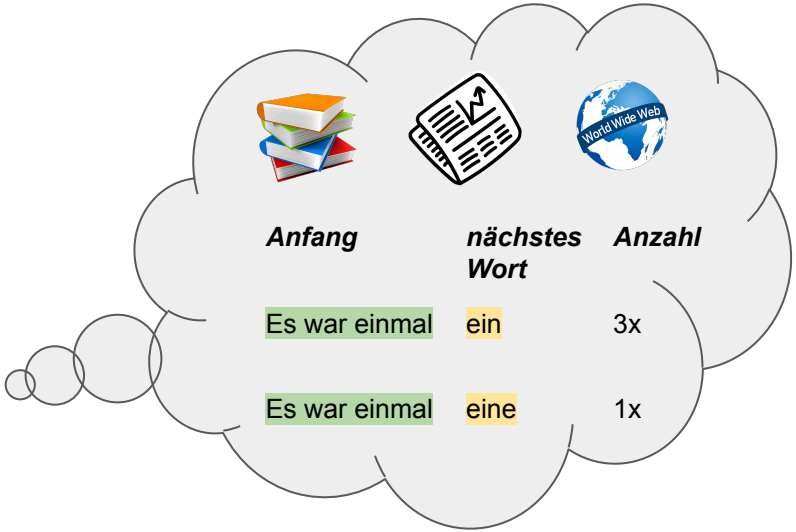
Eingabe

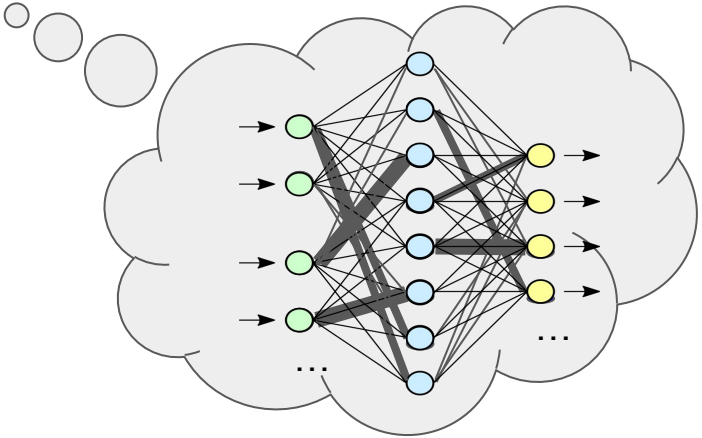
Sprachmodell

verschiedene
Modelle / Verfahren
z.B.
n-Gramm Listen,
neuronale Netzwerke

Ausgabe

ein	=	0.75
eine	=	0.25
…		
Bauer	=	0.00
Müller	=	0.00





Es war einmal

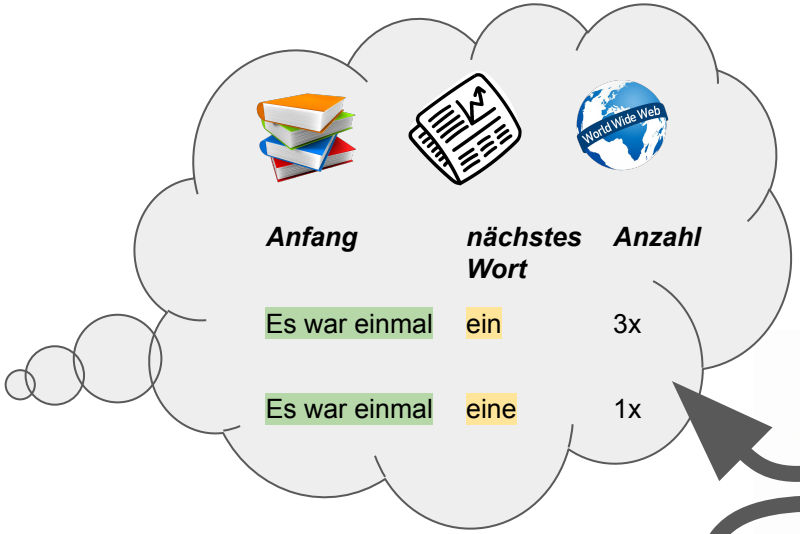
Eingabe

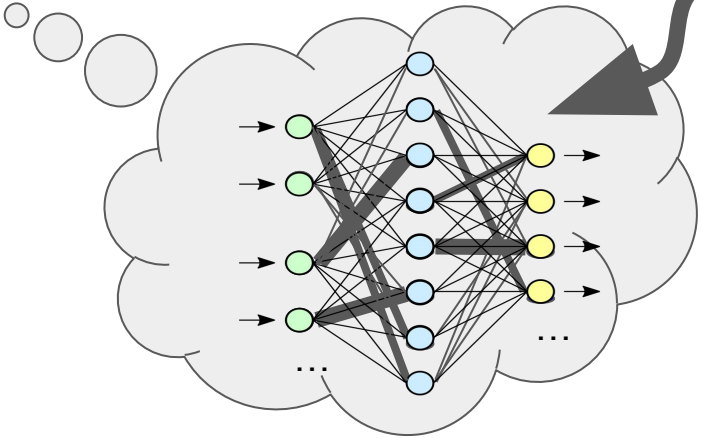
Sprachmodell

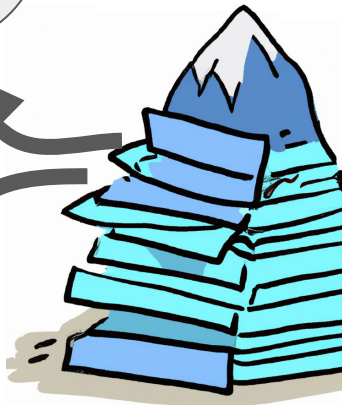
verschiedene
Modelle / Verfahren
z.B.
n-Gramm Listen,
neuronale Netzwerke

Ausgabe

ein	=	0.75
eine	=	0.25
…		
Bauer	=	0.00
Müller	=	0.00





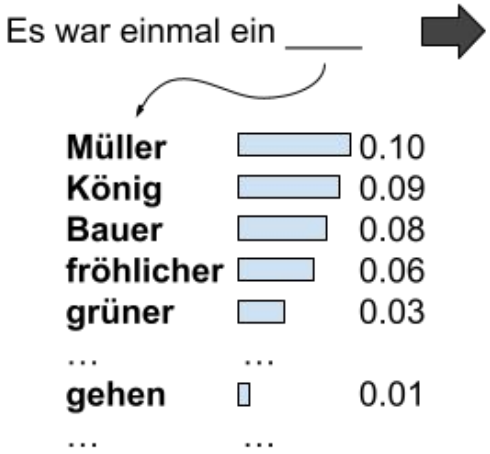


Wie nutzt man ein statistisches Sprachmodell? (vereinfacht)

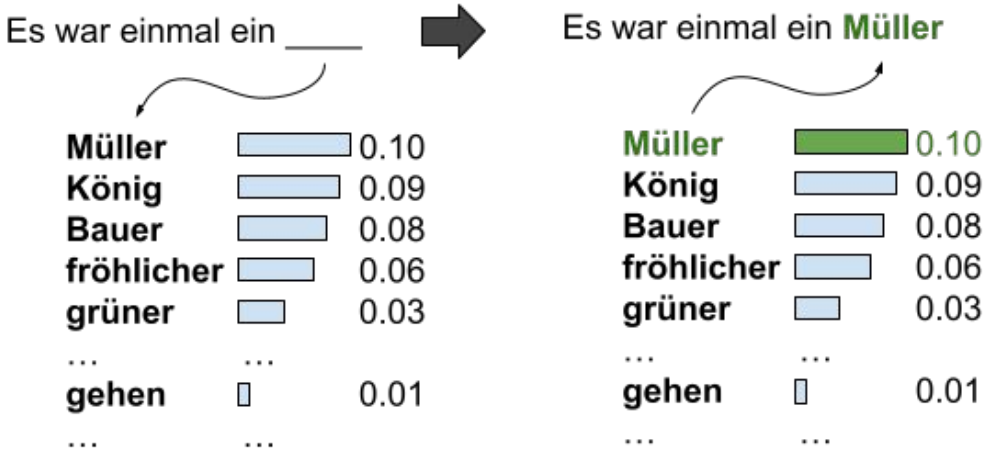
Wie nutzt man ein statistisches Sprachmodell? (vereinfacht)

Es war einmal ein ____

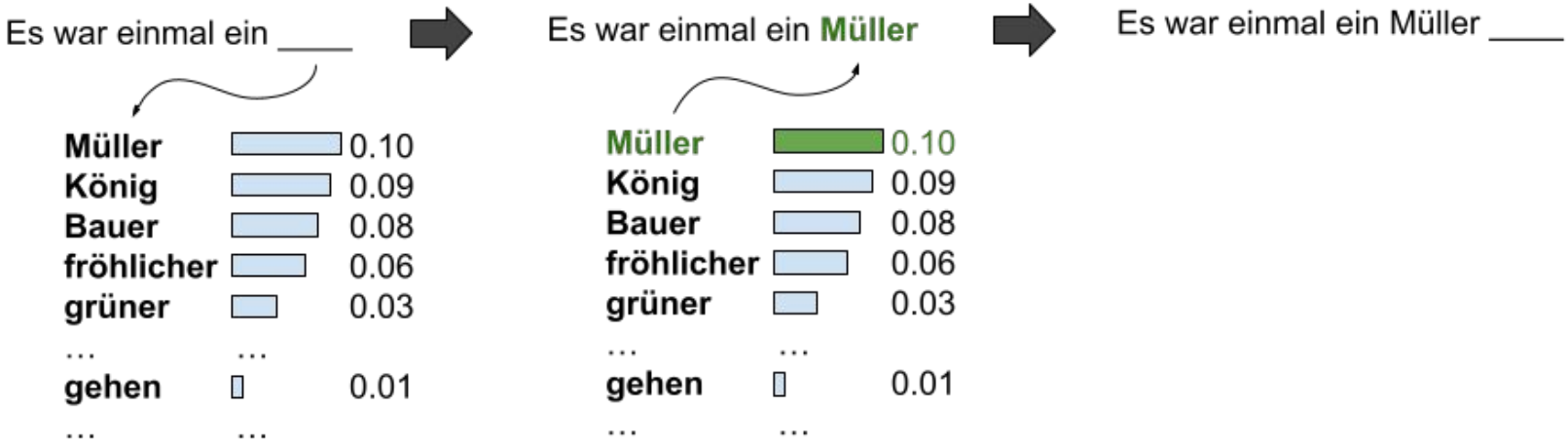
Wie nutzt man ein statistisches Sprachmodell? (vereinfacht)



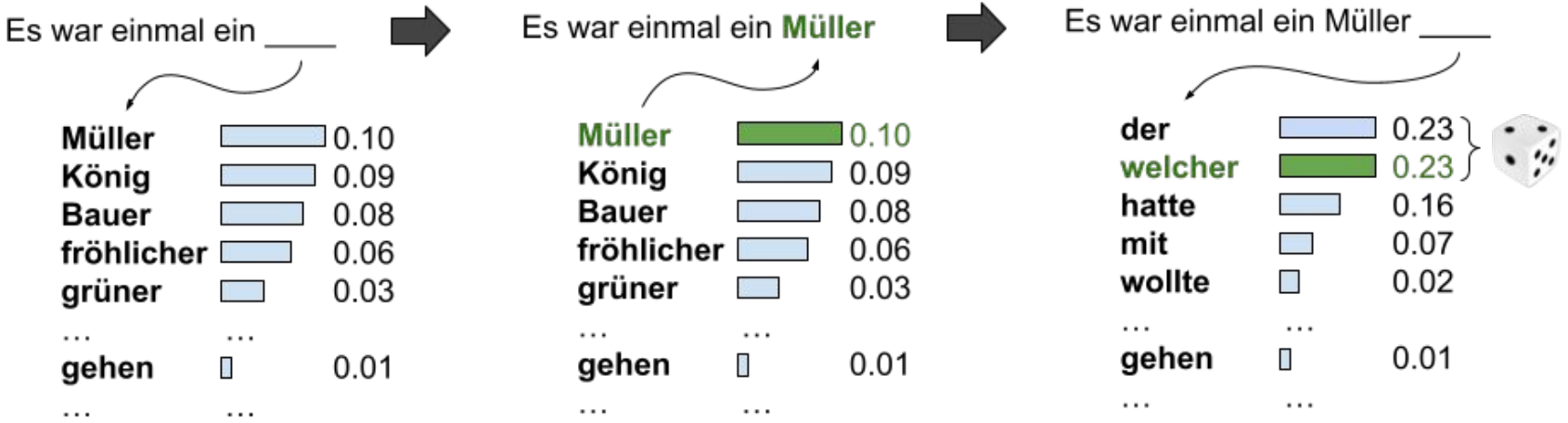
Wie nutzt man ein statistisches Sprachmodell? (vereinfacht)



Wie nutzt man ein statistisches Sprachmodell? (vereinfacht)



Wie nutzt man ein statistisches Sprachmodell? (vereinfacht)



Zufall spielt eine Rolle und ist gewünscht
→ immer leicht andere Ergebnisse

Experimentierumgebung
für Schülerinnen und Schüler

Experimentierumgebung
für Schülerinnen und Schüler

Soekia GPT

Das didaktische Sprachmodell

https://Soekia.ch

Soekia GPT

Schreibe mir ein Märchen.

Es war einmal ein Mann der hatte drei Söhne davon hiess der jüngste der Dummling und wurde verachtet und verspottet und bei jeder Gelegenheit zurückgesetzt. Der König sprach dein Gesang hat mir so wohl gefallen dass ich dir meine Tochter da zur Frau geben will. Die Frau war damit zufrieden des Königs Waffenträger aber der alles mit angehört hatte konnte es nicht sehen.

Der Held aber begab sich zum Könige der

SoekiaGPT starten

Handreichung für Lehrpersonen (PDF)

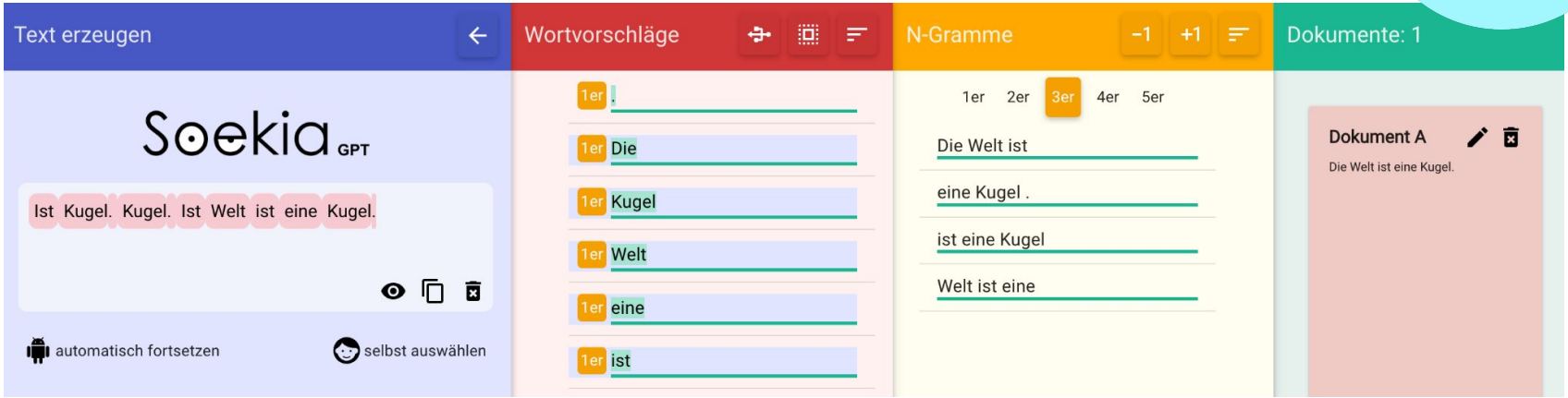
Beschreibung von SoekiaGPT (PDF)

Erklärvideo zu SoekiaGPT für Lehrpersonen

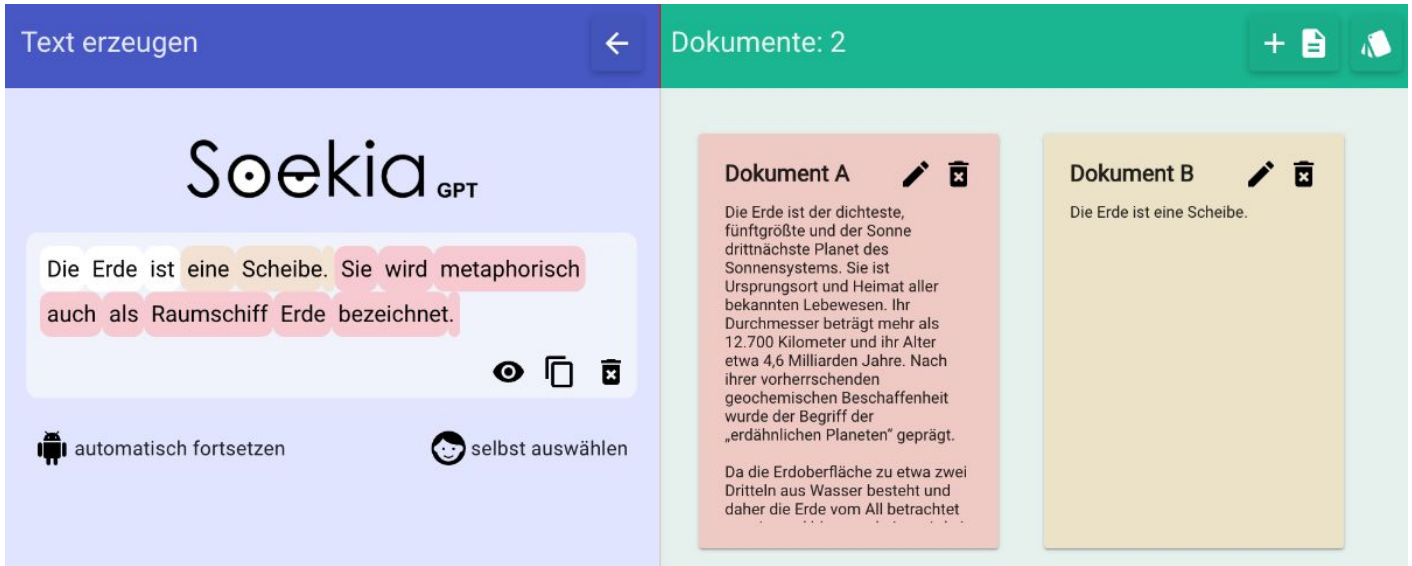
Was kann man mit Soekia$_{GPT}$ im Unterricht zeigen?

Was kann man mit Soekia$_{GPT}$ im Unterricht zeigen?

stochastischer Papagei

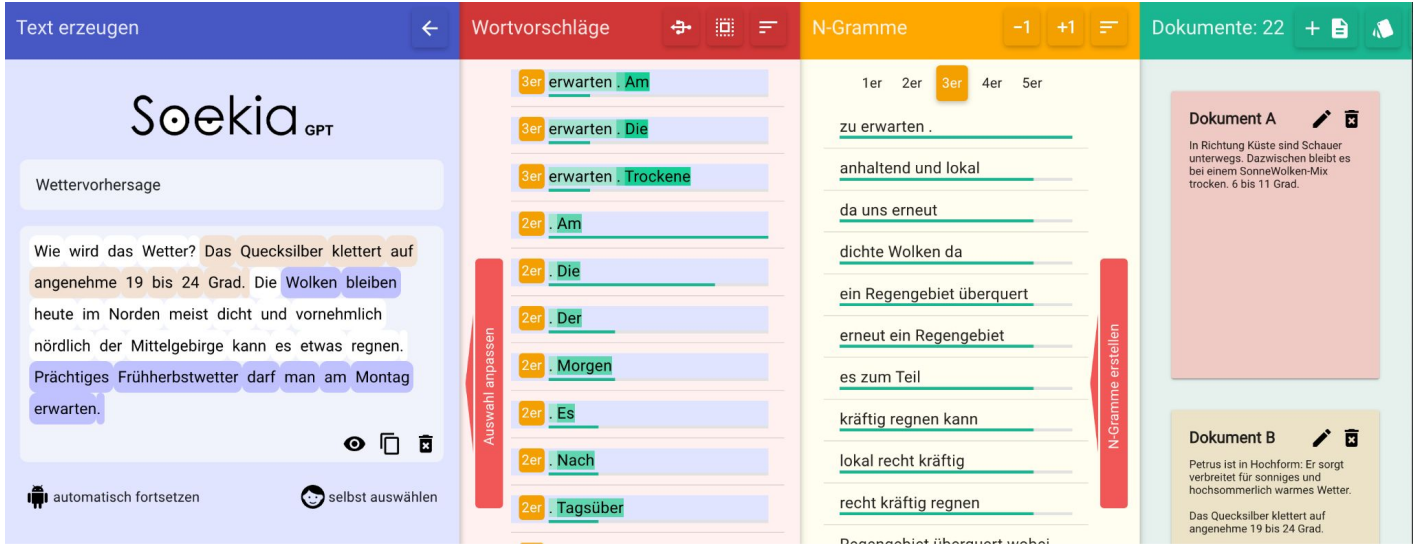


Was kann man mit Soekia$_{GPT}$ im Unterricht zeigen?



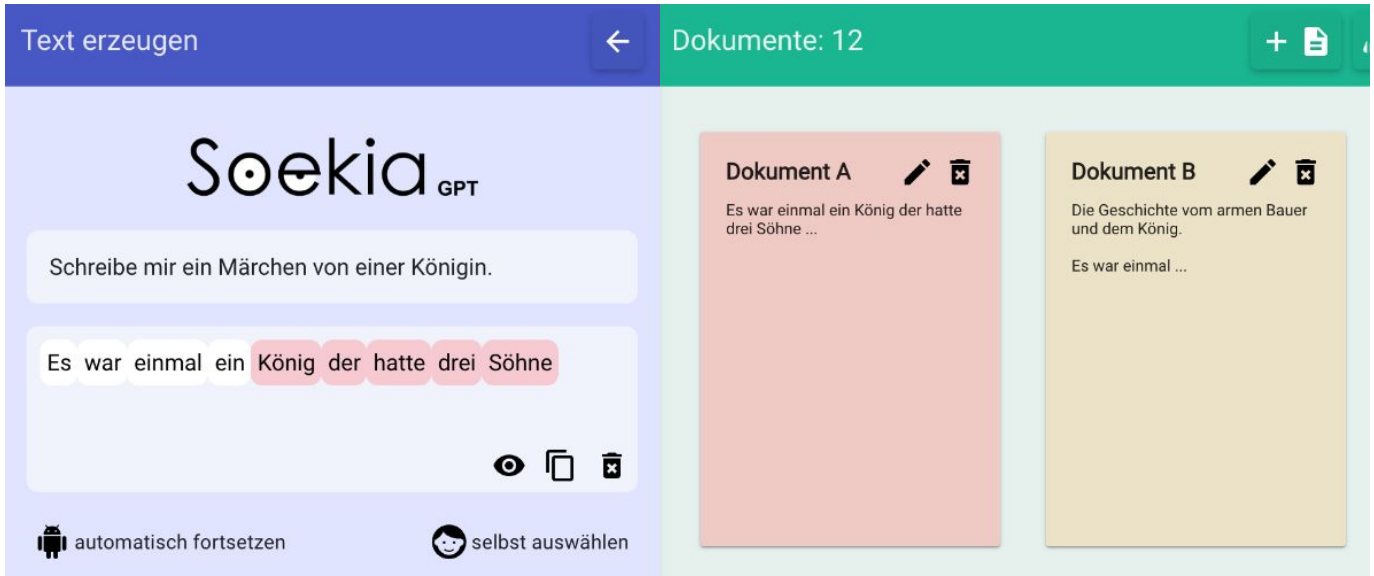
Trainingsdaten gezielt manipulieren

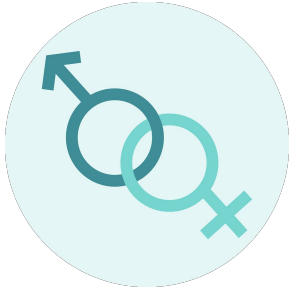
Was kann man mit Soekia$_{GPT}$ im Unterricht zeigen?



“Eine Ki-Halluzination liegt vor, wenn ein Large Language Model falsche Informationen erzeugt.”

Was kann man mit Soekia$_{GPT}$ im Unterricht zeigen?



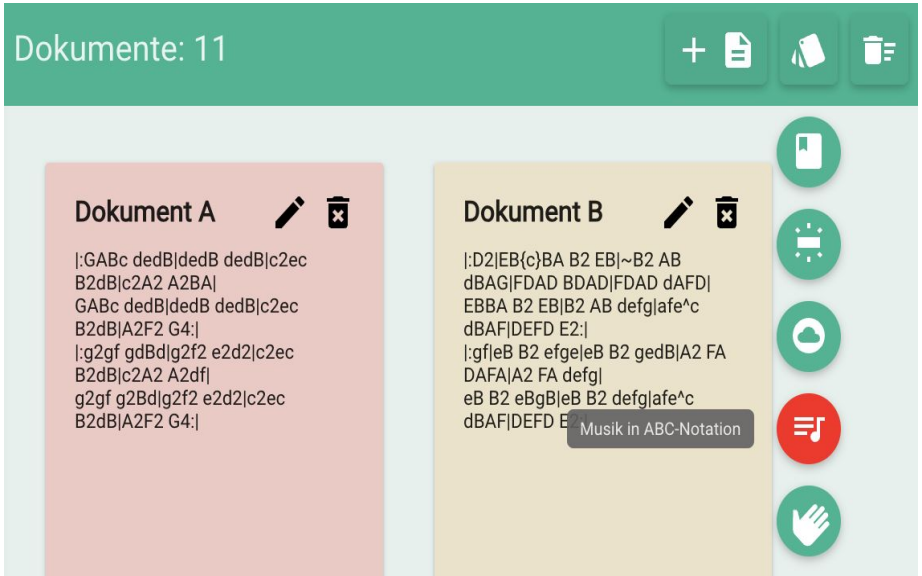


Bias-In-Daten

Es wird generiert, was in den Trainingsdaten vorhanden ist.

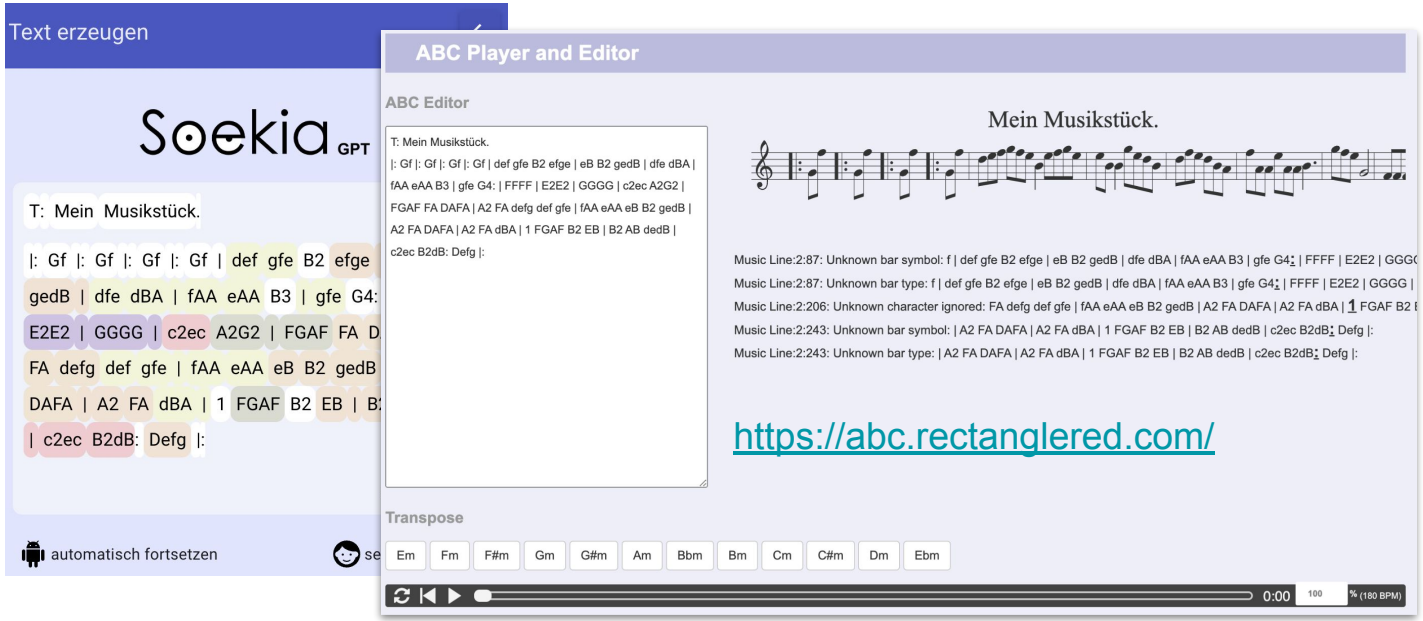
Stereotypen, alte Rollenbilder …

Was kann man mit Soekia$_{GPT}$ im Unterricht zeigen?



Vielseitigkeit von Sprachmodellen über Text hinaus.

Was kann man mit Soekia GPT im Unterricht zeigen?



Vielseitigkeit von Sprachmodellen über Text hinaus.

Grenzen von Soekia GPT

Sprachmodell != ChatGPT (Produkt ist viel komplexer, als das reine Sprachmodell)

Alltagsphänomen ChatGPT lässt sich damit nur teilweise erklären.

Didaktische Reduktion

Vereinfachung,
aber grundlegend
vergleichbare Prozesse

Large Language Model (LLM)

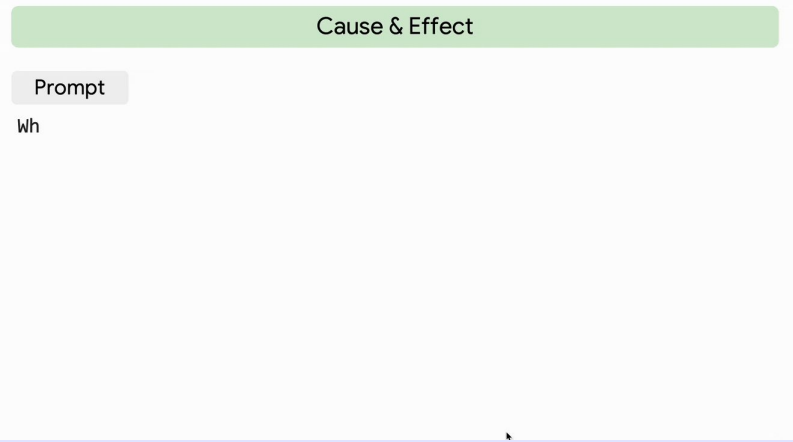
- Erlernt Wahrscheinlichkeiten von Wortkombinationen / -folgen aus einer **gigantischer** Kollektion
- Berechnet **mit Neuronalen Netzen** die Wahrscheinlichkeit für das nächste Token
- Schaut dabei zurück auf die letzten (z.B. **4000 bei GPT 3**) Token
- Wählt aus den wahrscheinlichsten Token (mit gewissem Zufall) genau eines aus
- Wiederholt den Prozess nach jedem Token

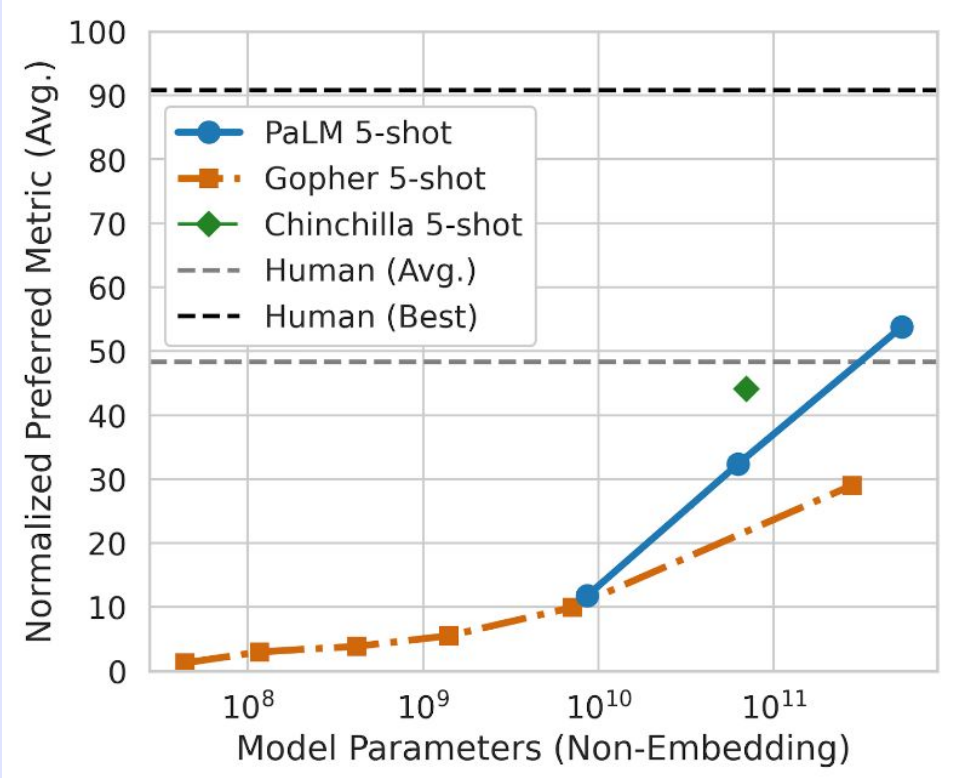
Soekia GPT

- Erlernt Wahrscheinlichkeiten von Wortkombinationen / -folgen aus einer **kleiner** Kollektion
- Berechnet **aus allen N-Grammen** die Wahrscheinlichkeit für das nächste Token
- Schaut dabei zurück auf die letzten (**maximal 5**) Token
- Wählt aus den wahrscheinlichsten Token (mit gewissem Zufall) genau eines aus
- Wiederholt den Prozess nach jedem Token

Es braucht eine gewisse Grösse …

BIG-bench: Testaufgabe für die Ki





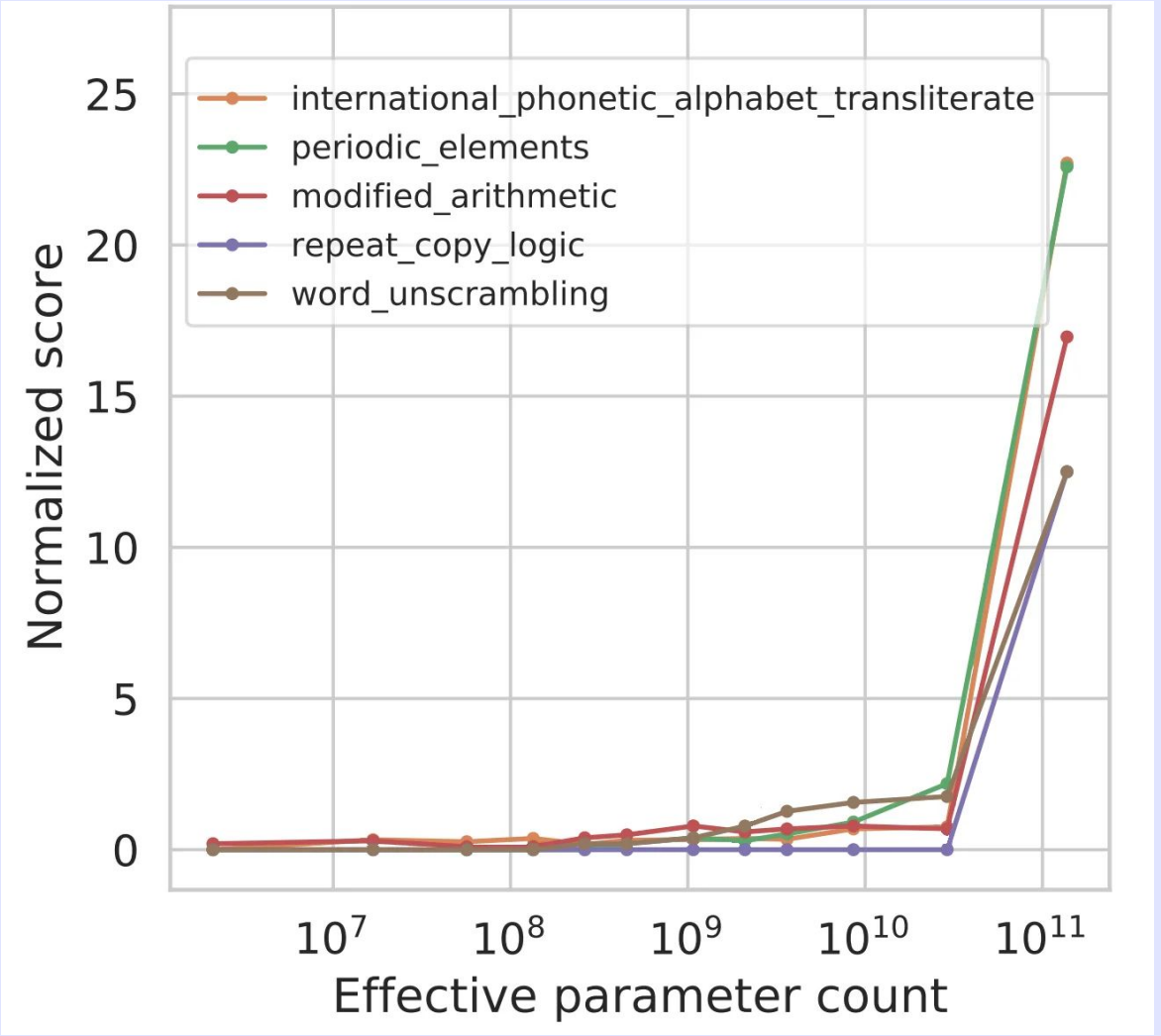
https://ai.googleblog.com/2022/04/pathways-language-model-palm-scaling-to.html

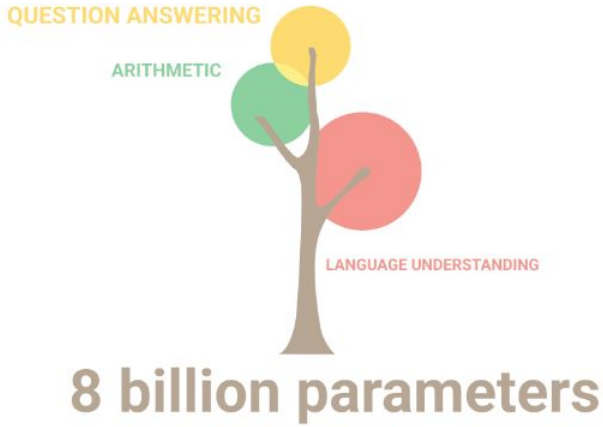
Es braucht eine gewisse Grösse …

Emergent Abilities of Large Language Models

1. We don't know at what scale they will appear
2. We don't know the level of ability until they do appear
3. We don't know the landscape of potential abilities

https://www.assemblyai.com/blog/emergent-abilities-of-large-language-models/




QUESTION ANSWERING
ARITHMETIC
LANGUAGE UNDERSTANDING
8 billion parameters

KI TECHNOLOGIEN

STARKE / ALLGEMEINE KI

?

SCHWACHE / SPEZIALISIERTE KI

KLASSISCH / SYMBOLISCH

- ENTSCHEIDUNGS-BÄUME
- TABELLENBASIERTE AGENTEN
- SUCHE
- EXPERTENSYSTEME
- REGELSYSTEME
- SYMBOLISCHE LOGIK

- EXPERTENSYSTEME MIT NN INPUT
- MONTE CARLO SUCHE MIT NN
- ENTSCHEIDUNGS-BAUMLERNEN
- XGBOOST
- MISCHALGORITHMEN

DEEP LEARNING

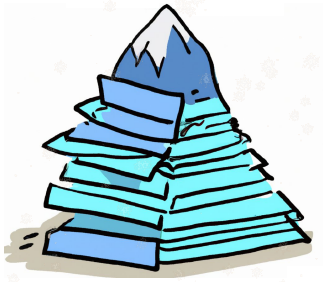
- NEURONALE NETZE
- LSTM
- CONVOLUTIONAL NN
- ATTENTION
- AUTOENCODER

- LINEARE REGRESSION
- NAIVE BAYES
- RANDOM FOREST
- Q LEARNING
- EVOLUTIONÄRE ALGORITHMEN

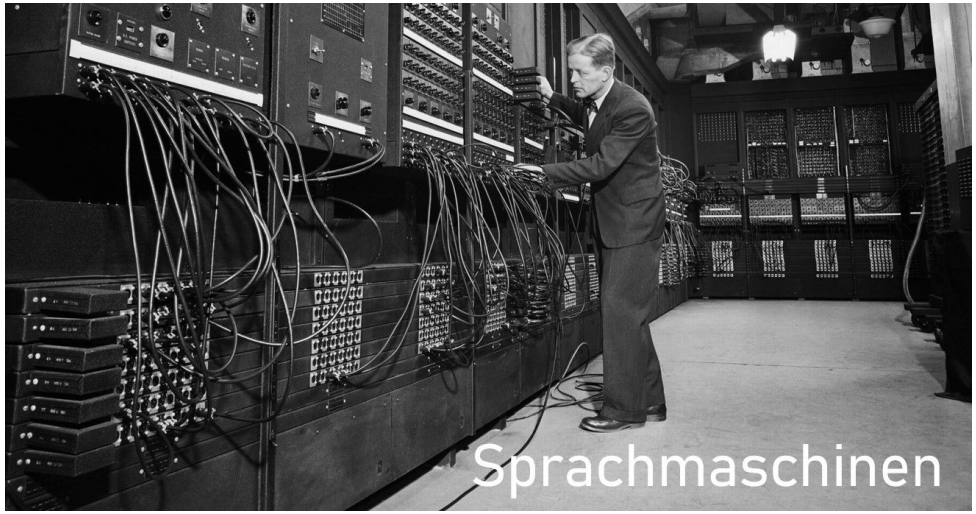
MASCHINELLES LERNEN

wissensbasiert

datenbasiert



Wissensbasierte Sprachmodelle



https://mia.phsz.ch/Sprachmaschinen (2021)

pädagogische hochschule schwyz

Michael Hielscher, Beat Döbeli Honegger

01.09.21

Was hat Informatik mit Sprache zu tun?

Warum dieses Dokument?

Mit Informatik verbinden die meisten Menschen Computer, Programmieren, Algorithmen und Nullen und Einsen. Doch Informatik hat auch sehr viel mit Sprache zu tun. Für den Schulunterricht kann die Verknüpfung zum Sprachunterricht bewusst herausgearbeitet werden, um dem eher mathematisch geprägten Bild von Informatik entgegenzuhalten. Die Informatik beschäftigt sich mit Sprache und Schrift aus unterschiedlichen Perspektiven und je nach Teildisziplinen mit anderen, damit verbundenen Herausforderungen.

Dieses Dokument richtet sich primär an Lehrpersonen und Dozierende des Lehrplanmoduls "Medien und Informatik". Es eignet sich damit nicht zur Bearbeitung mit Schülerinnen und Schülern, erklärt jedoch Hintergründe und Konzepte im Sinne einer fachlichen Überhöhung. An besonders gekennzeichneten Stellen sind Hinweise

```
1  Satz -> Subjekt Prädikat Adverb1 Adverb2 Objekt .
2  Satz -> Adverb1 Prädikat Subjekt Adverb2 Objekt .
3  Satz -> Adverb2 Prädikat Subjekt Adverb1 Objekt .
4  Satz -> Objekt Prädikat Subjekt Adverb2 Adverb1 .
5
6  Subjekt -> Hans
7  Subjekt -> Egon
8  Subjekt -> Max
9
10 Prädikat -> baut
11 Prädikat -> bastelt
12 Prädikat -> streicht
13
14 Adverb1 -> im Wald
15 Adverb1 -> im Garten
16 Adverb1 -> hinter dem Haus
17
18 Adverb2 -> mit seinem Freund
19 Adverb2 -> mit Susi
20 Adverb2 -> mit jeder Menge Werkzeug
21
22 Objekt -> eine riesige Baumhütte
23 Objekt -> ein Co-Kart
24 Objekt -> eine Rakete
25
```

ausführen

```
> Eine Rakete bastelt Max mit seinem Freund hinter dem Haus .
```

LCH DACHVERBAND LEHRERINNEN UND LEHRER SCHWEIZ

ICT-Entwicklungskonzept ICT für die Kantonsschule

Version 1.6

Es ist unübersehbar: Die zukünftige Gesellschaft ist geprägt durch soziale Netzwerke. Praktisch alle über-22-Jährigen in Österreich haben seit mehreren Jahren Zugriff auf Social Media. Dies bleibt nicht ohne Konsequenzen für die Wirtschaft. Bleibt nicht ohne Auswirkungen für der Bildungsbereich. Digitale Kompetenz darf nicht länger freiwillig bleiben.

https://programmingwiki.de/Nonsense/HTML

Werbung

Informatische Bildung in Schulen

ibis

www.informatischebildung.de

Praxisbeiträge

SoekiaGPT - ein didaktisches Sprachmodell

Hielscher, M.
Pädagogische Hochschule Schwyz

DOI: 10.0000/s00000-000-00000-0

Zusammenfassung

Statistische Sprachmodelle, insbesondere Textgeneratoren wie ChatGPT, haben in letzter Zeit große Aufmerksamkeit erregt. Künstliche Intelligenz ist auch ein Thema im Unterricht und die kritische Auseinandersetzung sollte ein Verständnis der grundlegenden Funktionsweisen umfassen. Sprachmodelle wie ChatGPT oder Bard sind jedoch komplexe Informatiksysteme und basieren auf riesigen Dokumentensammlungen und massiver Rechenleistung. Die Lernumgebung SoekiaGPT nimmt eine Reihe didaktischer Vereinfachungen vor und ermöglicht Schülerinnen und Schülern so einen Blick unter die Motorhaube von statistischen Sprachmodellen zu werfen und mit einem selbst gebauten Modell zu experimentieren.

Einleitung

Seit vielen Jahren werden Verfahren der künstlichen Intelligenz zum Beispiel bei Übersetzungen, Bilderkennung, Empfehlungssystemen, bei Social Media oder Predictive Maintenance eingesetzt. Im Herbst 2022 hat das Sprachmodell ChatGPT einen regelrechten Hype auch in der breiten Bevölkerung ausgelöst. Generative KI-Systeme, die aus einer Eingabe Text, Bild, Ton, Video, 3D-Modell usw. generieren, werden in immer mehr Bereichen eingesetzt. Große Sprachmodelle sind auch zu einem aktuellen Thema für die Schule geworden.

Nach der Empfehlung des Dagstuhl-Dreiecks zur Bildung in einer digital geprägten Welt (vgl. Gesellschaft für Informatik 2016) lassen sich digitale Alltagsphänomene im Informatikunterricht aus drei Perspektiven betrachtet und erklären. Die Anwendungsperspektive (Wie nutze ich das?), die gesellschaftlich-kulturelle Perspektive (Wie wirkt das?) und die technologische Perspektive (Wie funktioniert das?). Inzwischen gibt es eine schon fast unübersehbare Anzahl von Handreichungen für Lehrkräfte und Unterrichtsmaterialien für den Einsatz auf verschiedenen Schulstufen. Diese Materialien beziehen sich meist auf die konkrete Anwendung von Sprachmodellen oder auf die Auseinandersetzung mit Fragen rund um die Auswirkungen auf Schule und Gesellschaft. Zur grundlegenden Funktionsweise von großen Sprachmodellen (engl. LLM, Large Language Model) gibt es hingegen kaum Materialien. Hier setzt SoekiaGPT an.

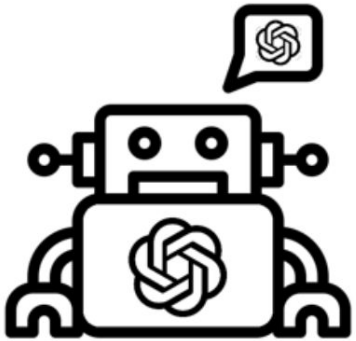
SoekiaGPT ist eine kostenlos verfügbare, webbasierte Lernumgebung ab der Sekundarstufe I, die vollständig im eigenen Browser arbeitet und ohne persönliches Konto genutzt werden kann. Mit Hilfe von vorgegebenen oder eigenen, sehr kleinen Dokumentenkollektionen (von wenigen Sätzen bis einigen hundert Seiten) kann unter die Motorhaube von Textgeneratoren und statistischen Sprachmodellen geschaut werden. Im Gegensatz zu reinen Visualisierungen können die Schülerinnen und Schüler mit dem Werkzeug vielfältig interagieren und experimentieren.

SoekiaGPT fordert mit einem Eingabefeld zu einer ersten Eingabe (Prompt) auf. Voreingestellt ist eine Kollektion mit verschiedenen Märchen. Der Textgenerator übernimmt die Aufforderung, ein neues Märchen mit "Es war einmal" zu schreiben, und beginnt einen Märchentext. Der Aufbau ähnelt damit den bekannten Chatbots (siehe Abb. 1). Die "Schau hinein" Funktion legt die dahinterliegenden Verarbeitungsschritte und Trainingsdaten offen und ist der zentrale Teil der Lernumgebung.

SoekiaGPT verzichtet bewusst auf die Berücksichtigung neuronaler Netzwerke und maschinellen Lernverfahren und nutzt ein klassisches

Abbildung 1: SoekiaGPT generiert ein "Märchen"

pädagogische hochschule schwyz



https://mia.phsz.ch/LLM

ChatGPT & Co. und Schule

Einschätzungen der Professur «Digitalisierung und Bildung» der Pädagogischen Hochschule Schwyz

Worum geht es?

Die Veröffentlichung des Chatbots ChatGPT Ende 2022 hat zu einem Hype in sozialen Medien und Massenmedien geführt. Diese Website (https://mia.phsz.ch/LLM) liefert Einordnungen der Professur "Digitalisierung und Bildung" der Pädagogischen Hochschule Schwyz zur Frage, welche Auswirkungen Sprachgeneratoren wie ChatGPT auf die Schule haben. Wir versuchen uns dabei auf langfristige, konzeptionelle Aspekte zu fokussieren und möglichst kurz zu bleiben. Auf unserer Website sind derzeit keine konkreten Unterrichtsszenarien oder Nutzungsanleitungen zu finden, da diese sehr stufen- und fachabhängig sind und sich die produkt- und versionsspezifischen Eigenschaften entsprechender Werkzeuge derzeit noch sehr rasch ändern.

большое спасибо
Thank you very much
gracias
mahad
ευχαριστίες
mahad
grazie
merci
kiitos
tack
teşekkürler
vielen Dank
vielen Dank
mulțumesc
grazie
gras
tack
ευχαριστίες
благодаря
спасибі
gracias
спасибі
mulțumesc
mahad
большое спасибо
Thank you very much
merci