

TRANSFORM 2023: AI for the Public Sector

# How AI is Revolutionizing the Future of Text

Paulina Grnarova  
CEO at DeepJudge  
[paulina@deepjudge.ai](mailto:paulina@deepjudge.ai)

# What is AI?

**Artificial**

Human-Made

**Intelligence**

??

# Early Days: AI Based on Rules

- Expert Systems
- Logic & Databases

Welcome to

```
EEEEEE LL      IIII  ZZZZZZ  AAAAA
EE      LL      II    ZZ      AA  AA
EEEEEE LL      II    ZZZ      AAAAAA
EE      LL      II    ZZ      AA  AA
EEEEEE LLLLLL  IIII  ZZZZZZ  AA  AA
```

Eliza is a mock Rogerian psychotherapist.  
The original program was described by Joseph Weizenbaum in 1966.  
This implementation by Norbert Landsteiner 2005.

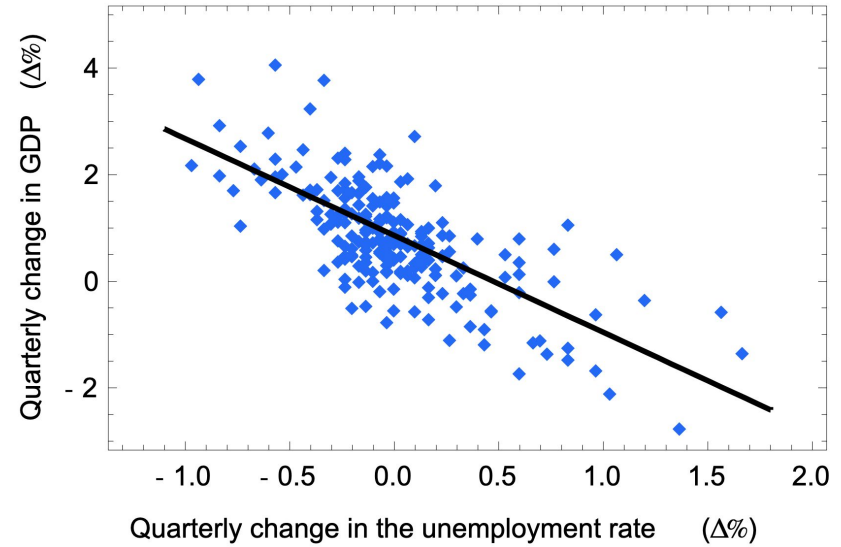
```
ELIZA: Is something troubling you ?
YOU:   Men are all alike.
ELIZA: What is the connection, do you suppose ?
YOU:   They're always bugging us about something or other.
ELIZA: Can you think of a specific example ?
YOU:   Well, my boyfriend made me come here.
ELIZA: Is it important to you that your boyfriend made you come here ?
YOU:   He says I'm depressed much of the time.
ELIZA: I am sorry to hear that you are depressed.
YOU:   It's true. I am unhappy.
ELIZA: Can you explain what made you unhappy ?
YOU:
```

# The Limits of Rules and Human Programming

What is on the roof?



# Learning from data: Machine Learning



# Deep Learning

Neural networks learn hierarchical representations

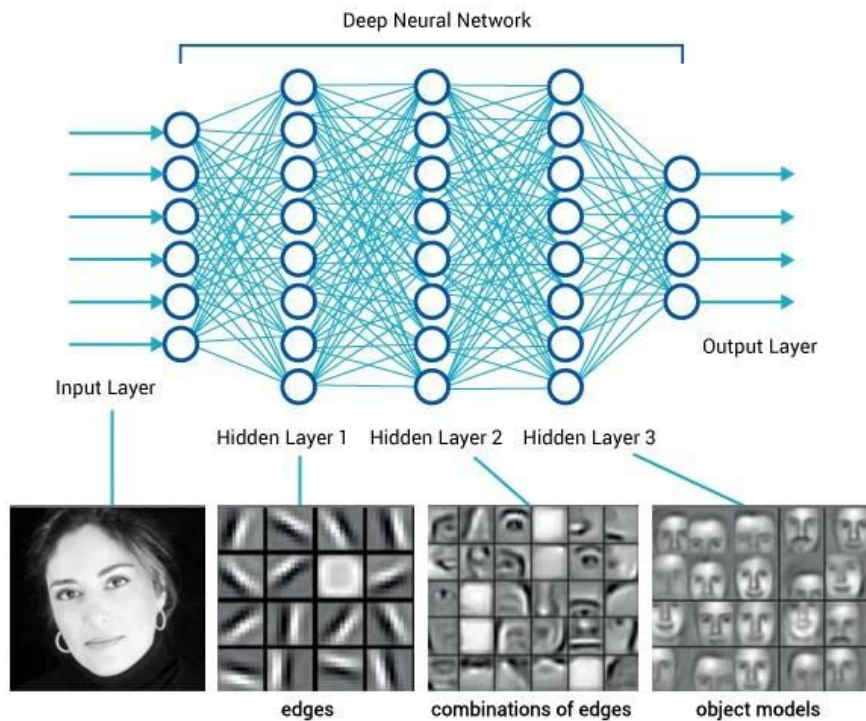


Image: Julien Vitay



# Natural Language Processing

# Language models

🔍 I saw a cat|

I saw a cat on the chair

I saw a cat running after a dog

I saw a cat in my dream

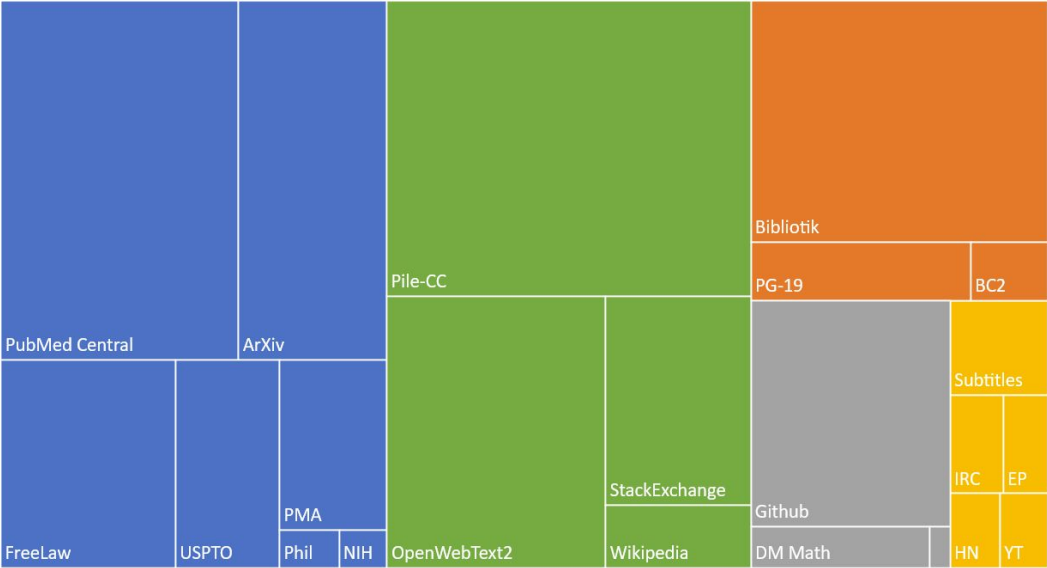
I saw a cat book



# Datasets

Composition of the Pile by Category

■ Academic ■ Internet ■ Prose ■ Dialogue ■ Misc



# Transformers

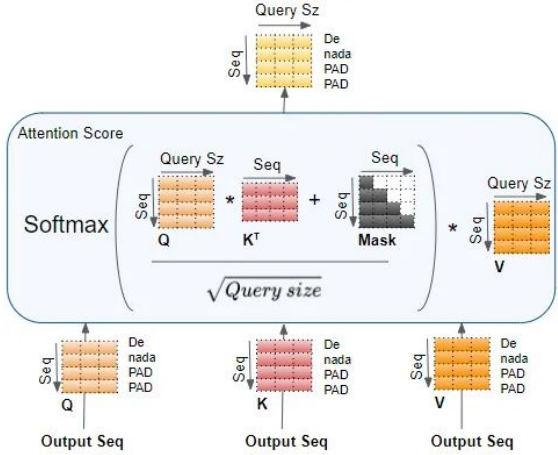


Image: Ketan Doshi

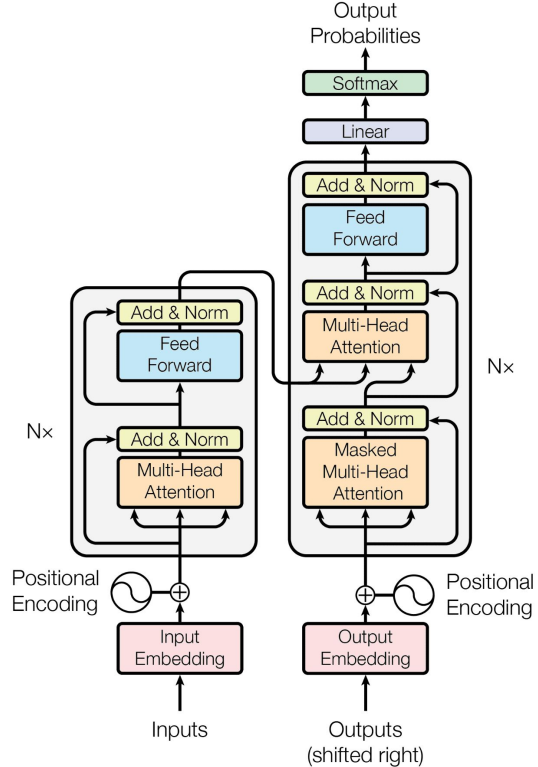


Image: Vasvani 2017

# Next Token Prediction

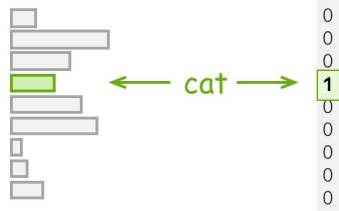
we want the model  
to predict this



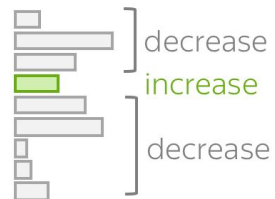
Training example: **I saw a cat** on a mat <eos>

Model prediction:  $p(* | \text{I saw a})$

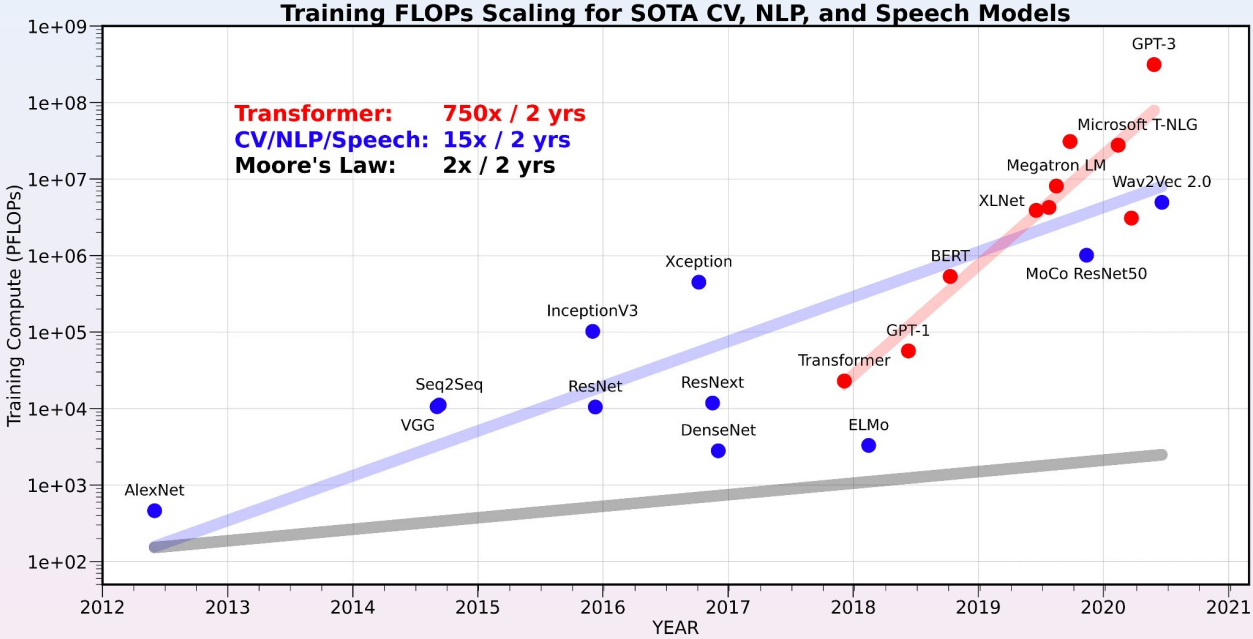
Target



Loss =  $-\log(p(\text{cat})) \rightarrow \min$



# Recent Developments: Scale





# ***AI for Text & Public Sector***

# Make Government Easier

Support citizens when interacting with the state

Increased transparency

Guidance through  
processes



Increased  
accessibility

Navigation of regulations

# Make Governance Easier

Support the public sector directly

## Reduce bureaucracy

Automation of repetitive tasks: reports, data entry, file management

## Speed up processes

Faster processing of applications and paperworks

## Increase communication

Via chatbots and AI assistants

# Oligopoly of Large US Firms

Increasing centralization

Intransparent data harvesting

Massive capital investments

Public (& academic) competition non-existent / much too slow



# Strengthen Switzerland

**New technologies enable  
Powerful applications**

**Switzerland is uniquely  
positioned**

- High concentration of expertise
- Exporter of high-end knowledge work
- Careful with AI regulation by EU

**The time to act is now!**

- Distribute to as many people as possible
- Do we need a SwissGPT

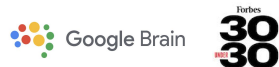
DeepJudge Founders

# World-class team with 25+ years AI experience



**Dr. Paulina Grnarova**  
CEO

PhD ETH in AI  
Researcher @ [Google Brain](#)  
Forbes 30 under 30  
Digital Shaper CH 2021, 2022



**Dr. Yannic Kilcher**  
CTO

PhD ETH in ML at scale  
[Google AI Language](#)  
World's largest ML YouTube  
channel



**Dr. Florian Schmidt**  
Head of Research

PhD ETH in NLP  
NLP advisor @ Typewise  
Published at [top NLP conferences](#)



**Dr. Kevin Roth**  
CDS / iCFO

PhD ETH in Data Science  
Theoretical Physicist  
Quant @ [Credit Suisse](#)  
Researcher @ [Google Brain](#)



Breakthrough technology

# How our AI is trained:

“High School”  
Pre training



×

Understanding of  
natural language

“Law School”  
Adaptive fine-tuning



×

Understanding of  
legal concepts

“Law Company”  
Active Learning



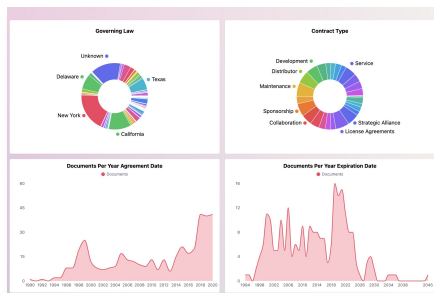
×

Understanding  
customer and task  
specific concepts

Unique recipe based on our own research and experience.  
These AI breakthroughs allow solving problems that were not possible a few years ago.

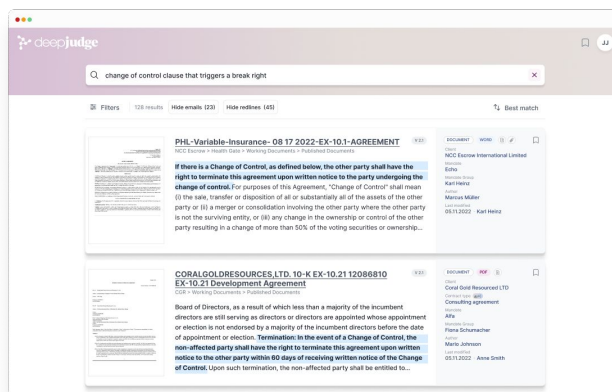
# Our AI assistant unlocks the content of any unstructured document

Make sense of legal unstructured data



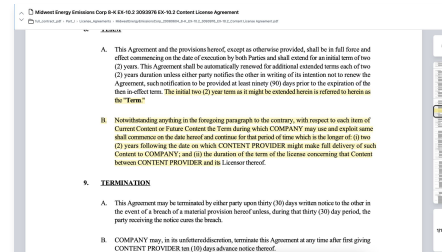
Explore

Surface relevant documents using natural language

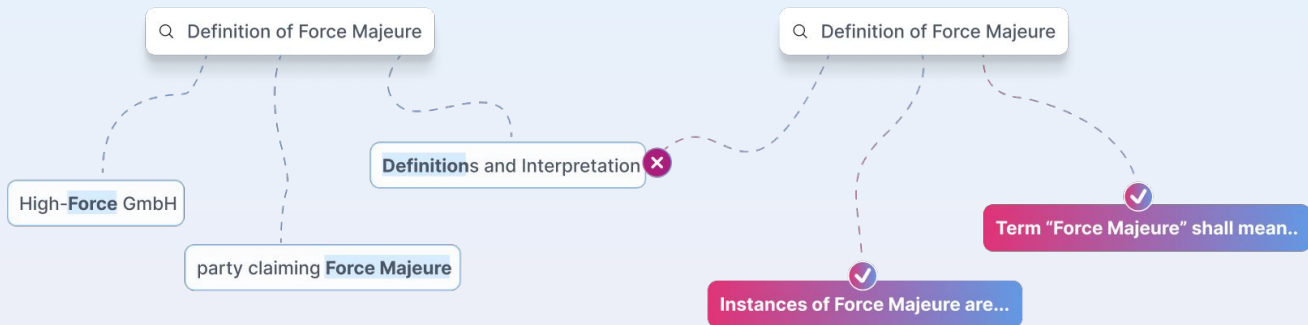


Search

Extract and process key information



Analyze



## Keyword search

the old way

- Too many irrelevant results
- Need to hit the correct words
- No understanding of concepts

## Semantic search

the new way

- Searches the concept instead of just the words
- Context-sensitive
- Finds what you're looking for



# Semantic Search Engine

Based on years of experience in academia  
and industry

- **Completely private**
- **Document understanding core**  
(using our own LLMs)
- **Scales to millions of documents**
- **In-House document collections**  
Previous matters, Contract collections, Regulations / Compliance,  
Knowledge management, Email, Data rooms

# Make Public Documents Accessible

Many documents of the  
public sector are open

Regulations, Laws,  
Proceedings & Court Rulings



... but finding them is  
hard

DeepJudge enables  
Semantic Search in  
millions of documents

Muss ich für einen Schaden im Rahmen eines Gefallens haften?

Filters 128 results

Best match



### 137 III 539

http://relevancy.bger.ch/php/clir/http/index.php?highlight\_docid=atf%3A%2F%2F137-...

Gefälligkeitshandlungen ohne Rechtsbindungswillen; Haftung des Gefälligen.  
**Abgrenzung zum Vertrag (E. 4.1); Kinderhüten unter Nachbarinnen für eine beschränkte Dauer als Gefälligkeitshandlung (E. 4.2 und 4.3).**  
**Der Gefällige haftet nach den Grundsätzen über die unerlaubte Handlung (E. 5.1); dabei gilt grundsätzlich der eingeschränkte Sorgfaltsmassstab der eigenüblichen Sorgfalt (diligentia quam in suis) (E. 5.2).**

DOCUMENT

Date  
20. Oktober 2011 -

Civil law



### 129 III 181

V 2.1

http://relevancy.bger.ch/php/clir/http/index.php?highlight\_docid=atf%3A%2F%2F129-...

**Gefälligkeitshandlungen ohne Rechtsbindungswillen.**  
**Analoge Anwendung von Art. 422 Abs. 1 OR betreffend Schadenersatzpflicht des Geschäftsherrn auf Gefälligkeitshandlungen ohne Geschäftsbindungswille (E. 3 und 4).**  
**Sachverhalt ab Seite 181**  
**BGE 129 III 181 S. 181**  
**A.- Am Nachmittag des 1. März 1997 begab sich A. zum Bauernhof von B., um dort ein Kalb zu besichtigen, das er eventuell übernehmen wollte. Im Verlaufe des**

DOCUMENT

Date  
21. Oktober 2002 -

Civil law



# Why Search is Necessary

## Problems with Generative LLMs

- Hallucinations
- Intransparency
- Inconsistency
- 3rd party APIs

## Search engine allows starting from your own documents

- Reference documents prevent hallucinations
- Data source is transparent
- Consistent style
- Private environment

**Volunteer Project**

**Open-Source code**

**Massive crowd-sourced dataset**

**Publicly available models**

**Currently best-performing open  
ChatGPT alternative**



# Open Assistant

**Conversational AI for everyone.**

We believe we can create a revolution.

In the same way that Stable Diffusion helped the world make art and images in new ways, we want to improve the world by providing amazing conversational AI.

[Try our assistant](#)

[Help us improve](#)

The public sector has a big potential to benefit from latest breakthroughs in AI for text



Get in touch via [paulina@deepjudge.ai](mailto:paulina@deepjudge.ai)



Spin-off | **ETH** zürich



BRIDGE



WISSENSCHAFT.  
BEWEGEN  
GEBERTHOP STIFTUNG