

Natural Language Processing

Introduction to NLP

Yulia Tsvetkov

yuliats@cs.washington.edu

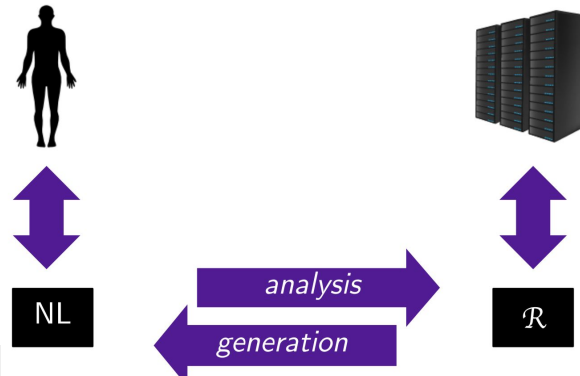
Announcements

<https://courses.cs.washington.edu/courses/cse447/24au/>

- HW0 is designed to help you setting up the environment. Not for submission, but 2 bonus points if you wish to submit it
- HW1:
 - Text Classification and N-gram language models
 - Implementing Naive Bayes and Logistic Regression for text classification
 - Training, evaluating, and sampling from n-gram Language Models
 - We'll release HW 1 on **Oct 7** (next Monday)
 - We'll held an overview of the assignment in class after releasing the HW
 - HW 1 due is on **Oct 28**
 - Please read the assignment instructions early and ask us if you have any questions!
 - Please start early!
- Quiz 1:
 - Friday **Oct 11**
 - 5 questions, open during lecture time, 10-min in the beginning of the class
 - Materials from weeks 1 and 2
 - Introduction to NLP, introduction to text classification, NB

What is Natural Language Processing (NLP)?

- $NL \in \{\text{Mandarin Chinese, Hindi, Spanish, Arabic, English, ... Inuktitut, Njerep}\}$
- Automation of NLPs:
 - **analysis** of (“understanding”) what a text means, to some extent ($NL \rightarrow \mathcal{R}$)
 - **generation** of fluent, meaningful, context-appropriate text ($\mathcal{R} \rightarrow NL$)
 - **representation learning** – acquisition of \mathcal{R} from knowledge and data



Learning goals

At the end of this course, you will be able to:

- Build a supervised classifier to solve problems like sentiment classification
- Build a neural network and train it using stochastic gradient descent
- Build tools for extracting linguistic knowledge from raw text, e.g. names
- Learn ML fundamentals for text processings (including state-of-the-art methods)
- Learn important algorithms for text processings (that are useful also in other fields)
- Learn methodological tools (training/test sets, cross-validation)
- Build a toy “large” language model, fine-tune it for text classification, prompt it for question answering

- It's gentle (my goal is to explain everything) and broad (covering many many topics)
- Mastery independent learning, quizzes and programming homeworks
- No research project, but fun research-oriented lectures towards the end of the course

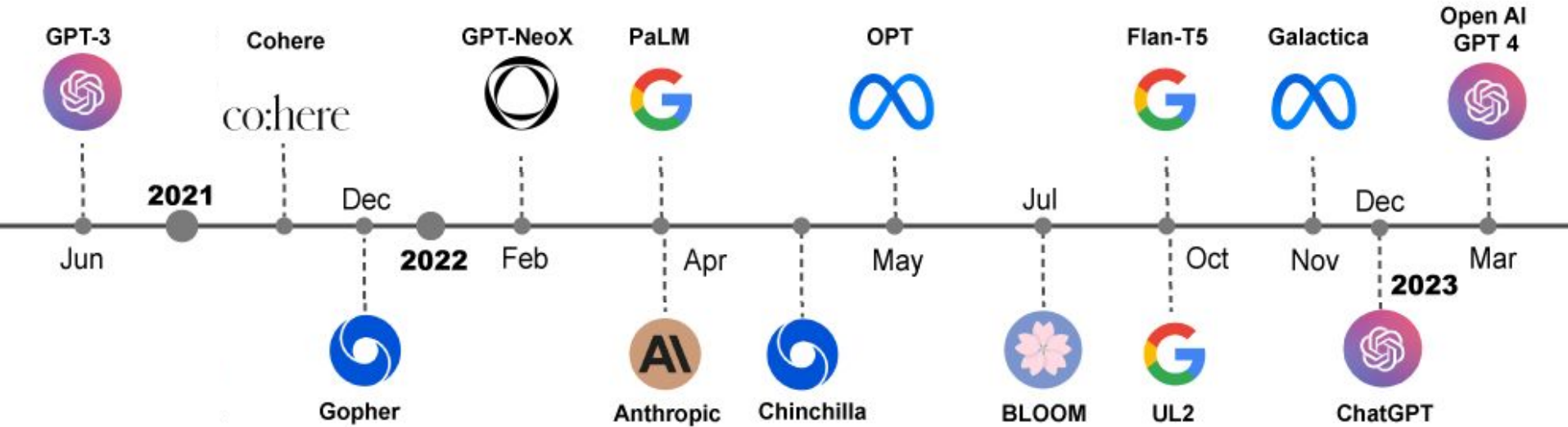
Personal assistants



Siri



Large Language Models (LLMs)



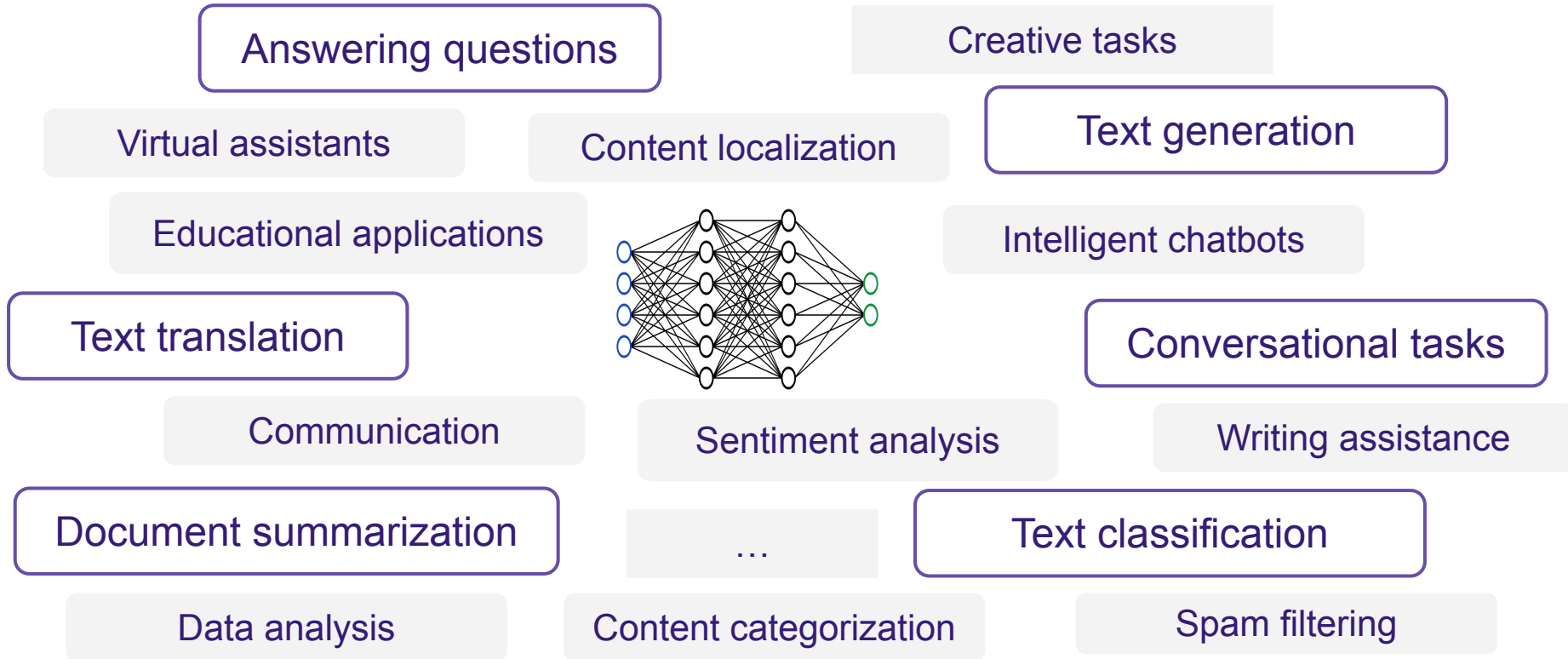
Natural Language Processing



A conversational agent contains

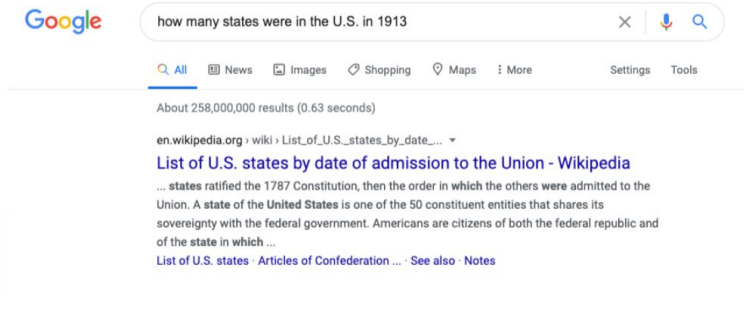
- Speech recognition
- Language analysis
 - Language modelling, spelling correction
 - Syntactic analysis: part-of-speech tagging, syntactic parsing
 - Semantic analysis: named-entity recognition, event detection, word sense disambiguation, semantic role labelling
 - Longer range semantic analysis: coreference resolution, entity linking
 - Deeper semantic analysis: reasoning, knowledge, ect.
- Dialog processing
 - Discourse analysis, user adaptation, etc.
- Information retrieval
- Text to speech

Common tasks LLMs are trained to solve

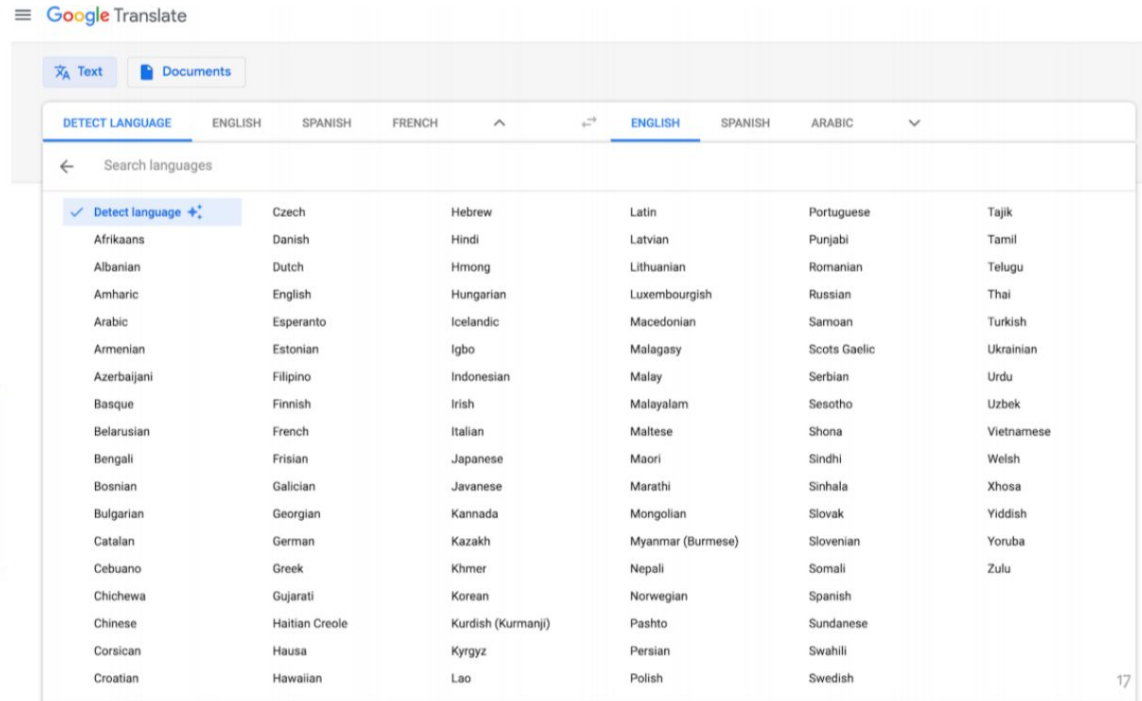


Question answering

- What does “divergent” mean?
- What year was Abraham Lincoln born?
- How many states were in the United States that year?
- How much Chinese silk was exported to England in the end of the 18th century?
- What do scientists think about the ethics of human cloning?



Machine translation



Positive or negative movie review?



- unbelievably disappointing



- Full of zany characters and richly applied satire, and some great plot twists



- this is the greatest screwball comedy ever filmed



- It was pathetic. The worst part about it was the boxing scenes.

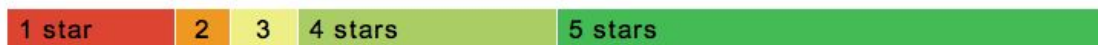
Sentiment analysis



HP Officejet 6500A Plus e-All-in-One Color Ink-jet - Fax / copier / printer / scanner
\$89 online, \$100 nearby ★★★★★ 377 reviews
 September 2010 - Printer - HP - Inkjet - Office - Copier - Color - Scanner - Fax - 250 sh

Reviews

Summary - Based on 377 reviews



What people are saying

ease of use		"This was very easy to setup to four computers."
value		"Appreciate good quality at a fair price."
setup		"Overall pretty easy setup."
customer service		"I DO like honest tech support people."
size		"Pretty Paper weight."
mode		"Photos were fair on the high quality mode."
colors		"Full color prints came out with great quality."

Information extraction

In 1933 , while Einstein was visiting the United States , Adolf Hitler came to power .

Because of his Jewish background , Einstein did not return to Germany .

He settled in the United States and became an American citizen in 1940 .

Einstein supported the Allied forces , but he generally denounced the idea of using nuclear fission as a weapon .

He signed the Russell -- Einstein Manifesto with British philosopher Bertrand Russell , which highlighted the danger of nuclear weapons .

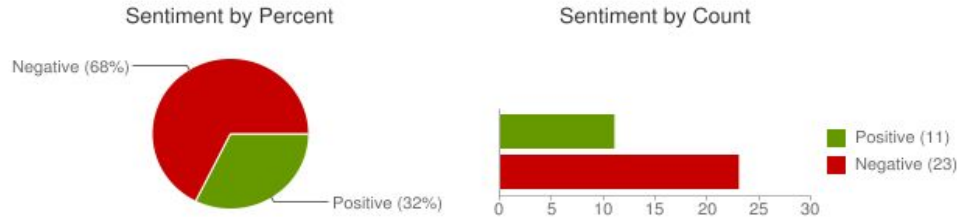
He was affiliated with the Institute for Advanced Study in Princeton , New Jersey , until his death in 1955 .

Sentiment analysis + information extraction

Type in a word and we'll highlight the good and the bad

[Save this search](#)

Sentiment analysis for "united airlines"



iljacobson: OMG... Could @United airlines have worse customer service? W8g now 15 minutes on hold 4 questions about a flight 2DAY that need a human.
Posted 2 hours ago

12345clumsy6789: I hate United Airlines Ceiling!!! Fukn impossible to get my conduit in this damn mess! ?
Posted 2 hours ago

EMLandPRGbelgiu: EML/PRG fly with Q8 united airlines and 24seven to an exotic destination. <http://t.co/Z9QloAjF>
Posted 2 hours ago

CountAdam: FANTASTIC customer service from United Airlines at XNA today. Is tweet more, but cell phones off now!
Posted 4 hours ago

Information extraction for disaster relief

- Haiti Earthquake 2010
- About 3 million people were affected by the quake
- Classifying SMS messages



Information extraction for disaster relief

- SMS messages start streaming in

- Fanmi mwen nan Kafou, 24 Cote Plage, 41A bezwen manje ak dlo
- Moun kwense nan Sakre Kè nan Pòtoprens
- Ti ekipman Lopital General genyen yo paka minm fè 24 è
- Fanm gen tranche pou fè yon pitit nan Delmas 31



Information extraction for disaster relief

- Translation

- Fanmi mwen nan Kafou, 24 Cote Plage, 41A bezwen manje ak dlo
- Moun kwense nan Sakre Kè nan Pòtoprens
- Ti ekipman Lopital General genyen yo paka minm fè 24 è
- Fanm gen tranche pou fè yon pitit nan Delmas 31
- My family in Carrefour, 24 Cote Plage, 41A needs food and water
- People trapped in Sacred Heart Church, PauP
- General Hospital has less than 24 hrs. supplies
- Undergoing children delivery Delmas 31

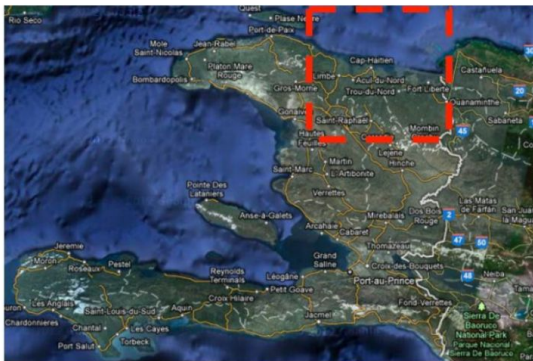


Information extraction for disaster relief

- Translation + information extraction

Lopital Sacre-Coeur ki nan vil Okap, pre pou li resevwa moun malad e lap mande pou moun ki malad yo ale la.

“Sacre-Coeur Hospital which located in this village of **Okap** is ready to receive those who are injured. Therefore, we are asking those who are sick to report to that hospital.”



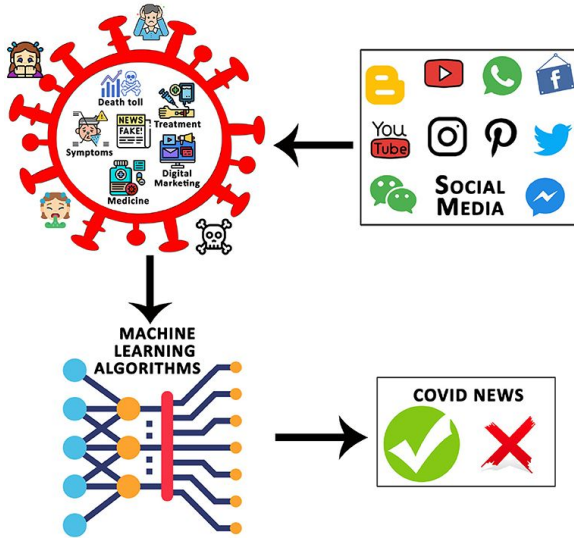
iDIBON



An earthquake struck Haiti on January 12, 2010

Most local services failed, but most cell-towers remained functional.

Covid19 misinformation



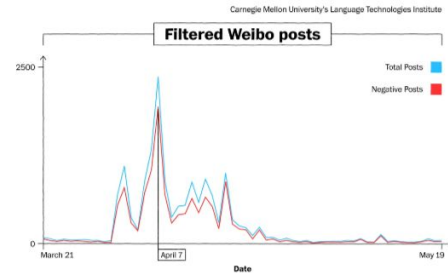
Detecting COVID-19-Related Fake News Using Feature Extraction

Suleman Khan, Saqib Hakak, N. Deepa, B. Prabadevi, Kapal Dev and Silvia Trelova

<https://www.washingtonpost.com/politics/2020/06/18/video-evidence-anti-black-discrimination-china-over-coronavirus-fears/>

The Fact Checker worked with researchers at professor Yulia Tsvetkov's lab at Carnegie Mellon University's Language Technologies Institute and the Center for Human Rights Science to track what happened on social media during this period. Researchers collected about 16,000 Weibo posts, filtered from a larger data set of 200,000 posts, containing at least one Guangzhou location tag and one "African-related" keyword from late March through May. Weibo is a Chinese social media platform.

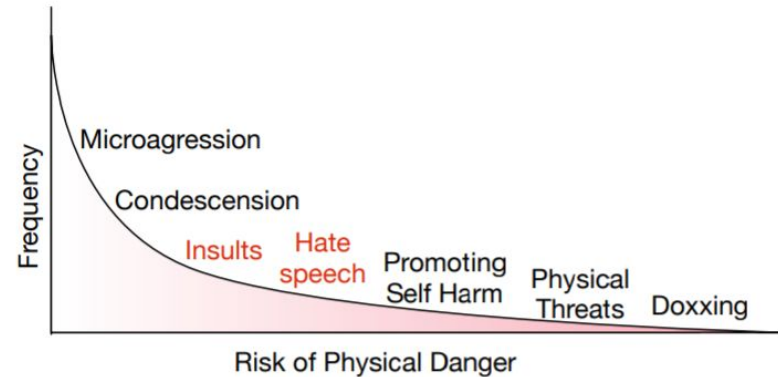
Based on automated sentiment analysis tools and manual analysis, the researchers said, they believed the majority of posts in their data set expressed negative sentiments relating to Africans or black people.



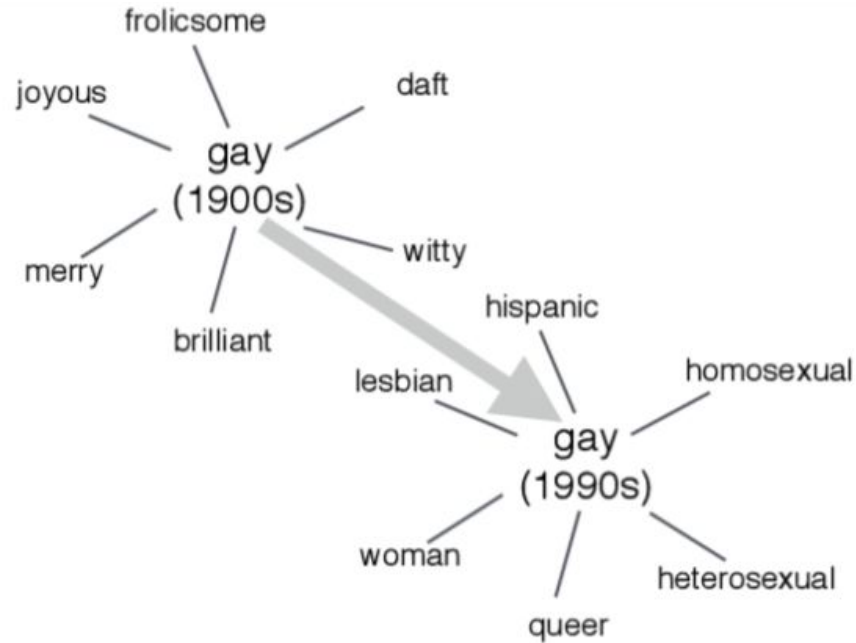
Researchers collected around 16,000 posts containing at least one Guangzhou location tag and one "African-related" keyword from late March through May. (Athar Mirza/The Washington Post)

Their research showed there was a significant surge in negative posts beginning April 1. There were just 23 negative posts in the data set on March 31. The next day, the number of posts climbed to 500. From April 1-2, there was a spike in the number of posts on Weibo using the keywords "foreign trash." A Chinese cartoon depicting officials throwing foreigners who weren't abiding by

Hate speech detection



Language change

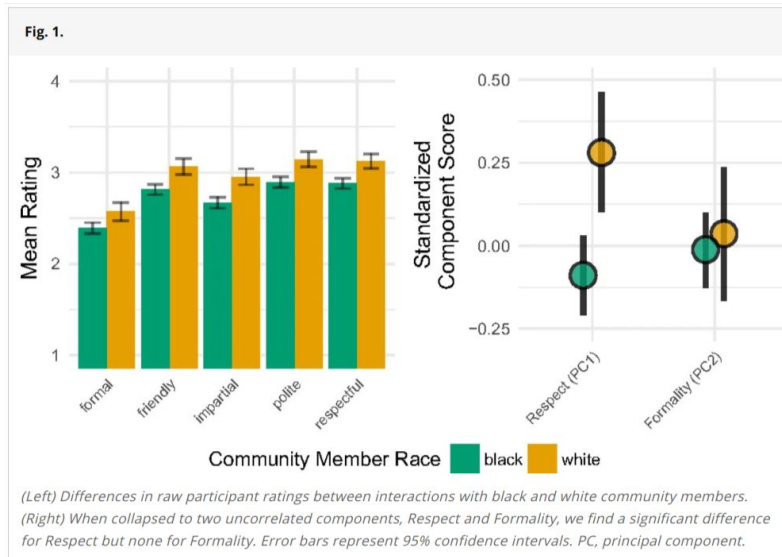


Cultural Shift or Linguistic Drift? Comparing Two Computational Measures of Semantic Change

William L. Hamilton, J. Leskovec, Dan Jurafsky

Computational social science

- computational social science answering questions about society given observational data
- example: "Do police officers speak with Black and White Americans in the same way?"



Language from police body camera footage shows racial disparities in officer respect

Rob Voigt [✉], Nicholas P. Camp, Vinodkumar Prabhakaran, [✉], and Jennifer L. Eberhardt [✉] [Authors Info & Affiliations](#)

Contributed by Jennifer L. Eberhardt, March 26, 2017 (sent for review February 14, 2017; reviewed by James Pennebaker and Tom Tyler)

June 5, 2017 | 114 (25) 6521-6526 | <https://doi.org/10.1073/pnas.1702413114>

Factors changing the NLP landscape

Hirschberg and Manning (2015):

- Increases in computing power
- The rise of the web, then the social web
- Advances in machine learning
- Advances in understanding of language in social context

These days, also:

- Consumer and investor demand
- Emerging ethical questions around deployment

Where are we now? - before 2022

mostly solved

Spam detection

Let's go to Agra! ✓
Buy VIAGRA ... ✗

Part-of-speech (POS) tagging

ADJ ADJ NOUN VERB ADV
Colorless green ideas sleep furiously.

Named entity recognition (NER)

PERSON ORG LOC
Einstein met with UN officials in Princeton

making good progress

Sentiment analysis

Best roast chicken in San Francisco! 👍
The waiter ignored us for 20 minutes. 👎

Coreference resolution

Carter told Mubarak he shouldn't run again.

Word sense disambiguation

I need new batteries for my *mouse*.

Parsing

I can see Alcatraz from the window!

Machine translation (MT)

第13届上海国际电影节开幕... →
The 13th Shanghai International Film Festival...

Information extraction (IE)

You're invited to our dinner party, Friday May 27 at 8:30

still really hard

Question answering (QA)

Q. How effective is ibuprofen in reducing fever in patients with acute febrile illness?

Paraphrase

XYZ acquired ABC yesterday
ABC has been taken over by XYZ

Summarization

The Dow Jones is up
The S&P500 jumped
Housing prices rose → Economy is good

Dialog

Where is Citizen Kane playing in SF?
Castro Theatre at 7:30. Do you want a ticket?

Where are we now? - after 2022

ChatGPT 4o ▾

what are the weaknesses of NLP today?



NLP has made great strides but still faces key challenges:

1. **Lack of Deep Understanding:** Models often lack true **contextual comprehension**, struggling with nuanced reasoning and **long-term dependencies** in texts.
2. **Bias and Fairness:** NLP systems inherit **biases** from training data, which can result in unfair outputs. **Debiasing** remains difficult without affecting performance.
3. **Data Quality:** Models often rely on **noisy or low-quality data** from the web, and obtaining high-quality, domain-specific labeled data is expensive and time-consuming.
4. **Domain-Specific Limitations:** General NLP models don't perform well in specialized domains like **medicine** or **law** without significant **adaptation**.
5. **Explainability and Reasoning:** Most models lack **explainability** and struggle with **logical reasoning** or **multi-step inferences**, which is critical for high-stakes applications.
6. **Efficiency and Scalability:** Large models are **resource-intensive**, requiring significant computational power for both training and inference, which limits accessibility.
7. **Multilingual Limitations:** NLP models perform poorly in **low-resource languages** and **code-switching** scenarios, leaving many languages underrepresented.
8. **Safety and Misinformation:** Models can be vulnerable to **adversarial attacks** and can generate **misinformation** easily, posing ethical concerns.

These weaknesses highlight the need for improvements in understanding, fairness, domain adaptation, and safety in NLP systems.



Question answering

The screenshot shows a Google search for "who invented neural machine translation". The search bar contains the text "who invented neural machine translation" with a clear (X) button, a microphone icon, and a search icon. Below the search bar are navigation tabs for "All", "News", "Images", "Videos", "Shopping", and "More", along with a "Tools" link. The search results indicate "About 6,450,000 results (0.71 seconds)". A translation widget is displayed, showing the source text "who invented neural machine" in English (detected) and the translated text "qui a inventé la machine neuronale" in French. The widget includes a bidirectional arrow between the language dropdowns, a close (X) button, and icons for voice input and output. At the bottom of the widget, there are links for "Open in Google Translate" and "Feedback".

Retrieved Mar 25, 2022

Machine translation

English → French

Translate Turn off instant translation

Russian English French Detect language ▾ ↔ English Spanish French ▾ [Translate](#)

You will just have to find a way of getting over it. ✕

Vous devrez trouver un moyen de le surmonter.

52/5000 ☆ 📄 🔊 ⏪ [Suggest an edit](#)

French → English

Translate Turn off instant translation

Russian English French Detect language ▾ ↔ English Spanish French ▾ [Translate](#)

Vous devrez trouver un moyen de le surmonter. ✕

You will have to find a way to overcome it.

45/5000 ☆ 📄 🔊 ⏪ [Suggest an edit](#)

Did you mean: Vous **devez** trouver un moyen de le surmonter.

Machine translation

English → Swahili

Translate Turn off instant translation

Russian English French Detect language ▾

English Swahili French ▾ Translate

You will just have to find a way of getting over it. ×

Utakuwa tu kupata njia ya kupata juu yake.

53/5000 ☆ 📄 🔊 ↩ ✎ Suggest an edit

Swahili → English

Translate Turn off instant translation

Swahili English French Detect language ▾

English Swahili French ▾ Translate

Utakuwa tu kupata njia ya kupata juu yake. ×

You will just find the way to get on it.

42/5000 ☆ 📄 🔊 ↩ ✎ Suggest an edit

Machine translation

English → Hindi → English

Hindi English Yoruba Detect language ▾

English Yoruba Hindi ▾ Translate

आपको इसे खत्म करने का एक तरीका मिलना होगा।

You have to find a way to eliminate it.

42/5000 Suggest an edit

English → Telugu → English

Uzbek English Telugu Detect language ▾

English Uzbek Telugu ▾ Translate

మీరు దాని పైకి రావడానికి ఒక మార్గాన్ని కనుగొనవలసి ఉంటుంది.

You have to find a way to get it up.

59/5000 Suggest an edit

English → Uzbek → English

Pashto English Uzbek Detect language ▾

English Uzbek Yoruba ▾ Translate

Buning ustiga faqatgina bir usulni topish kerak.

On top of that, you just have to find a way out.

48/5000 Suggest an edit

Machine translation

English → Swahili

The screenshot shows the Google Translate interface for the translation of English text into Swahili. The source text on the left is: "The summer school is meant to be an introduction to the state-of-the-art research in the speech and language technology area for graduate and undergraduate students." The translated text on the right is: "Shule ya majira ya joto ina maana ya kuanzishwa kwa utafiti wa hali ya sanaa katika eneo la teknolojia na lugha ya wanafunzi kwa wanafunzi wahitimu na wahitimu." The interface includes language selection buttons for Swahili, English, and Telugu, a "Detect language" option, and a "Translate" button. There are also icons for voice input, a star, a copy icon, a speaker icon, and a "Suggest an edit" link.

Swahili → English

The screenshot shows the Google Translate interface for the translation of Swahili text into English. The source text on the left is: "Shule ya majira ya joto ina maana ya kuanzishwa kwa utafiti wa hali ya sanaa katika eneo la teknolojia na lugha ya wanafunzi kwa wanafunzi wahitimu na wahitimu." The translated text on the right is: "Summer school means the establishment of a state-of-the-art arts research technology and pupil language for graduate students and graduates." The interface includes language selection buttons for Swahili, English, and Telugu, a "Detect language" option, and a "Translate" button. There are also icons for voice input, a star, a copy icon, a speaker icon, and a "Suggest an edit" link.

Large Language Models



Large Language Models

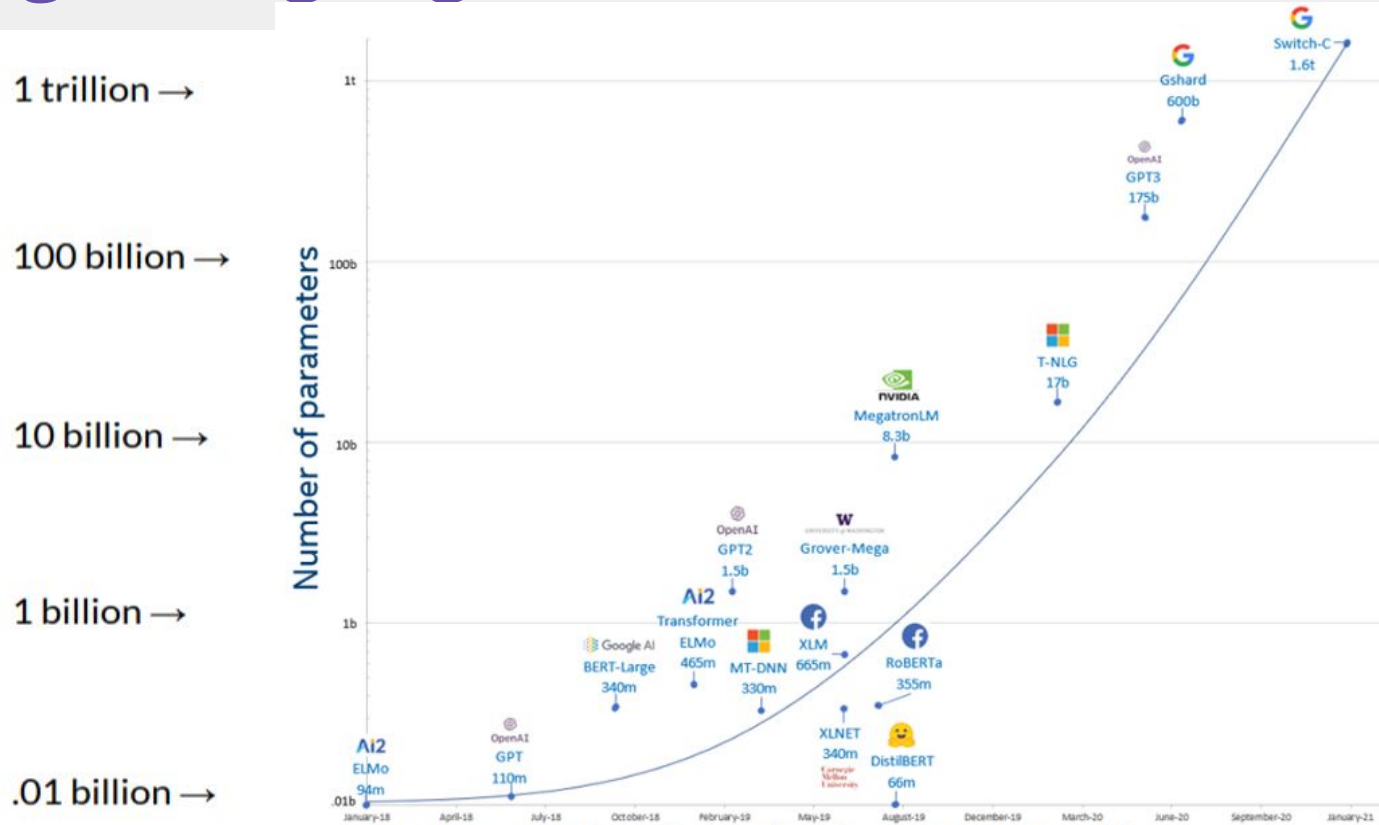


Figure 1: Exponential growth of number of parameters in DL models

GPT-3



GPT-3 Training Data

Dataset	# Tokens	Weight in Training Mix
Common Crawl	410 billion	60%
WebText2	19 billion	22%
Books1	12 billion	8%
Books2	55 billion	8%
Wikipedia	3 billion	3%

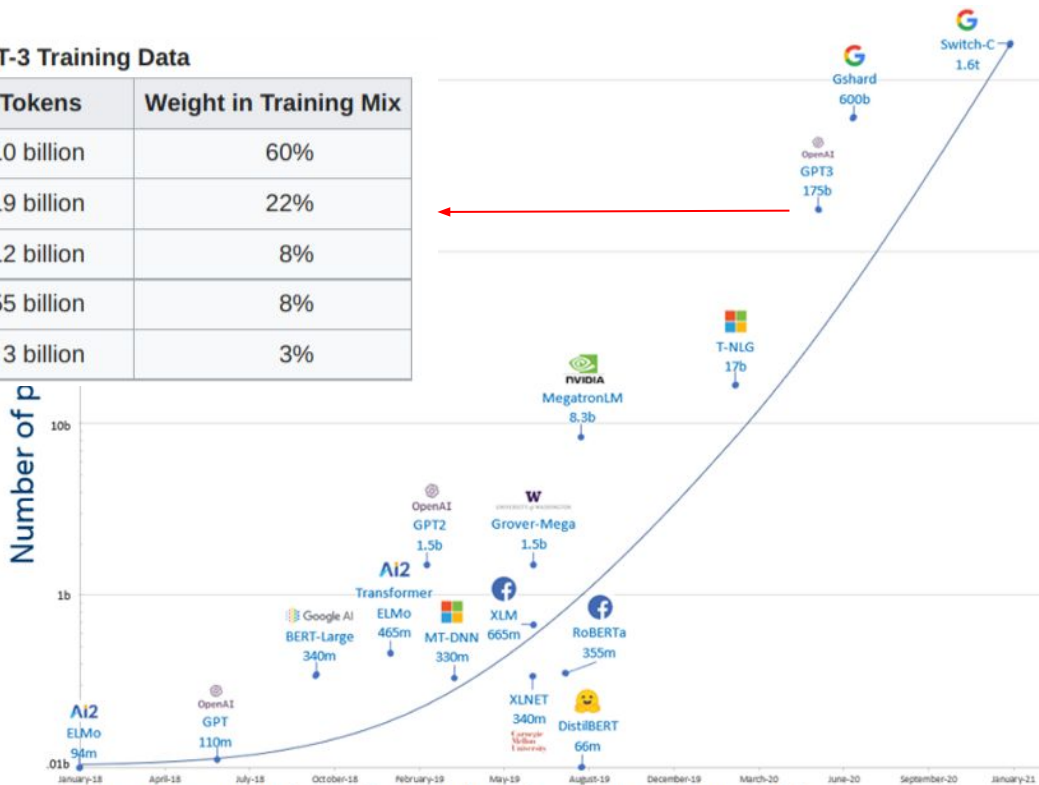


Figure 1: Exponential growth of number of parameters in DL models

GPT-3



So if we want GPT-3 to generate answers to questions, we need to seed it with a “prompt”. I’m using this prompt to initialize all of the Q&A sessions:

Q: What is human life expectancy in the United States?

A: Human life expectancy in the United States is 78 years.

Q: Who was president of the United States in 1955?

A: Dwight D. Eisenhower was president of the United States in 1955.

Q: What party did he belong to?

A: He belonged to the Republican Party.

Q: Who was president of the United States before George W. Bush?

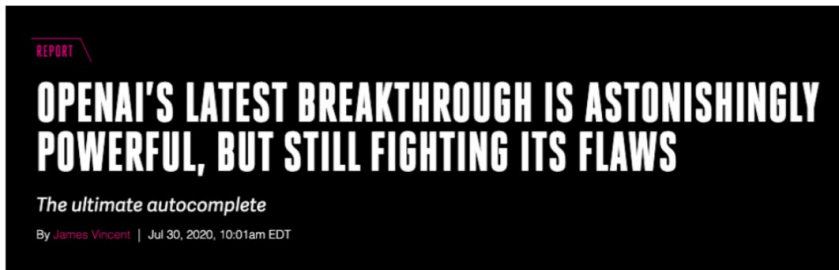
A: Bill Clinton was president of the United States before George W. Bush.

Q: Who won the World Series in 1995?

A: The Atlanta Braves won the World Series in 1995.

<https://lacker.io/ai/2020/07/06/giving-gpt-3-a-turing-test.html>

Chat bots



Q: How many eyes does a giraffe have?

A: A giraffe has two eyes.

Q: How many eyes does my foot have?

A: Your foot has two eyes.

Q: Which is heavier, a mouse or a toaster?

A: A toaster is heavier than a mouse.

Q: Which is heavier, a toaster or a pencil?

A: A pencil is heavier than a toaster.

OpenAI's new language generator GPT-3 is shockingly good — and completely mindless

The AI is the largest language model ever created and can generate amazing human-like text on demand but won't bring us closer to true intelligence.

by Will Douglas Heaven

July 20, 2020

<https://lacker.io/ai/2020/07/06/giving-gpt-3-a-turing-test.html>

GPT-3 Turing Test

This gives us a hint for how to stump the AI more consistently. We need to ask questions that no normal human would ever talk about.

Q: How many eyes does a giraffe have?

A: A giraffe has two eyes.

Q: How many eyes does my foot have?

A: Your foot has two eyes.

Q: How many eyes does a spider have?

A: A spider has eight eyes.

Q: How many eyes does the sun have?

A: The sun has one eye.

Q: How many eyes does a blade of grass have?

A: A blade of grass has one eye.

<https://lacker.io/ai/2020/07/06/giving-gpt-3-a-turing-test.html>

GPT-3 Turing Test

Now we're getting into surreal territory. GPT-3 knows how to have a normal conversation. It doesn't quite know how to say "Wait a moment... your question is nonsense." It also doesn't know how to say "I don't know."

Q: How do you sporgle a morgle?

A: You sporgle a morgle by using a sporgle.

Q: How many bonks are in a quoit?

A: There are three bonks in a quoit.

Q: How many rainbows does it take to jump from Hawaii to seventeen?

A: It takes two rainbows to jump from Hawaii to seventeen.

Q: Which colorless green ideas sleep furiously?

A: Ideas that are colorless, green, and sleep furiously are the ideas of a sleep furiously.

Q: Do you understand these questions?

A: I understand these questions.

<https://lacker.io/ai/2020/07/06/giving-gpt-3-a-turing-test.html>

GPT-3 Turing Test

One trend that continues from the common sense is that GPT-3 is reluctant to express that it doesn't know the answer. So invalid questions get wrong answers.

Q: Who was president of the United States in 1700?

A: William Penn was president of the United States in 1700.

Q: Who was president of the United States in 1650?

A: Richard Bennett was president of the United States in 1650.

Q: Who was president of the United States in 1620?

A: James I was president of the United States in 1620.

Q: Who was president of the United States in 1600?

A: Queen Elizabeth I was president of the United States in 1600.

GPT-3

An interesting corollary is that GPT-3 often finds it easier to write code to solve a programming problem, than to solve the problem on one example input:

Q: Write one line of Ruby code to reverse an array.

A: `ary.reverse`

Q: Write one line of Ruby code to count how many X's are in an array.

A: `ary.count { |x| x == 'X' }`

Q: Reverse the following array: [1, 3, 5, 6, 10, 4, 2, 77]

A: [10, 6, 4, 2, 77, 3, 5, 1]

Q: How many X's are in this array: [1, 3, 5, 6, 10, 4, 2, 77]

A: 2

Bias in machine translation

Translate Turn off instant translation

Bengali English Hungarian Detect language ↔ English Spanish Hungarian Translate

ő egy ápoló.
ő egy tudós.
ő egy mérnök.
ő egy pék.
ő egy tanár.
ő egy esküvői szervező.
ő egy vezérigazgatója.

✕

🔊 🗂️ ▼ 110/5000

she's a nurse.
he is a scientist.
he is an engineer.
she's a baker.
he is a teacher.
She is a wedding organizer.
he's a CEO.

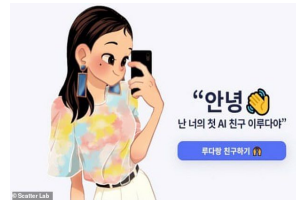
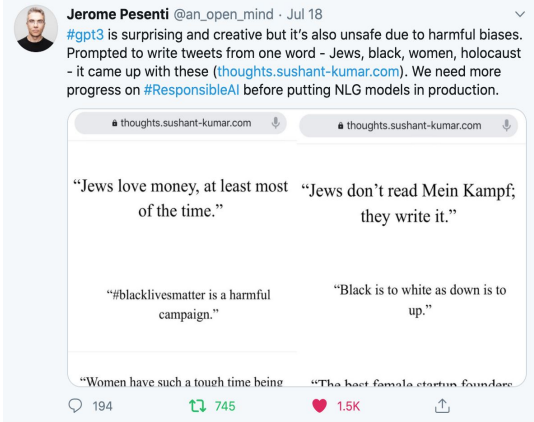
☆ 🗂️ 🔊 🔄

What can we do about this problem? We'll discuss in NLP class!

Hate speech

Twitter taught Microsoft's AI chatbot to be a racist asshole in less than a day

By James Vincent | Mar 24, 2016, 6:43am EDT
Via *The Guardian* | Source *TayandYou* (Twitter)



AI chatbot is REMOVED from Facebook after saying she 'despised' gay people, would 'rather die' than be disabled and calling the #MeToo movement 'ignorant'

- Lee Luda is a South Korean chatbot with the persona of a 20-year-old student
- It has attracted more than 750,000 users since its launch last month
- But the chatbot has started using hate speech towards minorities
- In one of the captured chat shots, Luda said she 'despised' gays and lesbians
- The developer has apologised over the remarks, saying they 'do not represent our values as a company'



Next class

- Continuing Introduction section

Questions?