



Italian Digital Media Observatory

Elena Musi

Associate Professor

Program Lead MSc in Data Science and Communication

Research consultant IDMO

(Italian Digital Media Observatory),

Luiss Datalab

Quale tra questi paragrafi promozionali vi sembra maggiormente convincente?

Ti invitiamo a partecipare all'Internet Governance Forum (IGF) 2022 in Italia! IGF è un importante momento di confronto e dialogo tra tutti gli attori interessati alla governance dell'internet, dove si discutono i temi più caldi della rete. L'edizione 2022 sarà particolarmente importante perché sarà la prima volta che l'Italia ospiterà l'IGF. Saremo quindi lieti di accogliere esperti di tutto il mondo per un confronto aperto e costruttivo sui temi più importanti della governance dell'internet. Non perdere l'occasione di prendere parte a questo evento globale e iscriviti ora all'IGF 2022!

Il prossimo Internet Governance Forum si terrà in Italia nel 2022. Sarà un'occasione unica per discutere di temi importanti come la cybersecurity, la gestione dei dati e la sostenibilità. Vi invitiamo a partecipare e a contribuire alla discussione sul futuro del nostro mondo online

Ciao! Stai cercando un posto dove discutere del futuro dell'internet? Allora l'Internet Governance Forum in Italia è quello che fa per te! Partecipa a questo evento per conoscere meglio come l'internet viene governato e per contribuire alla discussione. Ci saranno esperti di tutto il mondo che condivideranno le loro idee su questo importante tema. Non perdere l'occasione di essere parte di questa discussione!

Se doveste indovinare quale tra questi messaggi non è stato scritto da un umano quale scegliereste e perché ?

→ Non c'è nulla di umano, sono tutti e tre opera di GPT-3



GPT-3

Generative Pre-trained Transformer 3

Initial release date: 11 June 2020

Original author: [OpenAI](#)

Repository: github.com/openai/gpt-3

the GPT-3 revo

GPT-3: 175-billion parameter pre-trained language model for language generation

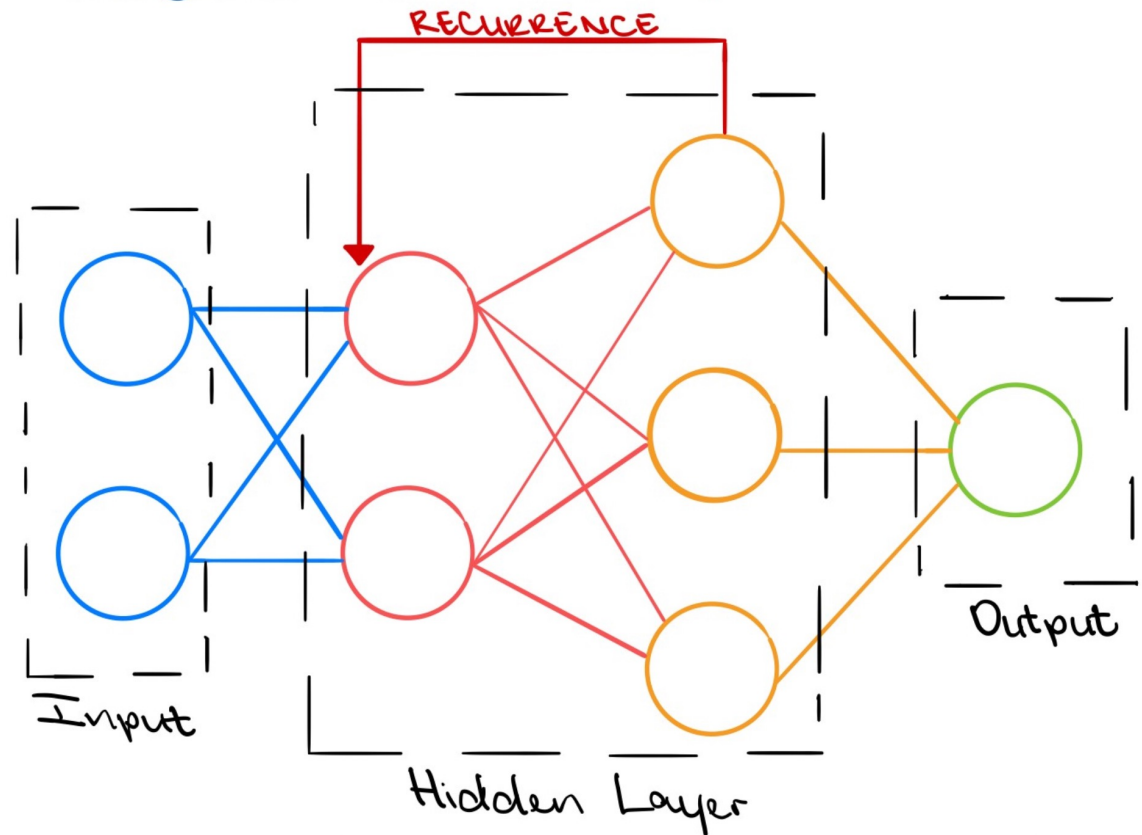
- ❖ autoregressive language model based on deep learning
- ❖ predictions are made step by step and the result of one prediction is used as input for the next prediction
- ❖ A learning revolution:
Few-shots learning: no need for annotated training data

Before...

RECURRENT
NEURAL
NETWORKS

<https://towardsdatascience.com/introducing-recurrent-neural-networks-f359653d7020>

RECURRENT NEURAL NETWORKS



Now...

The screenshot displays the OpenAI API Playground interface. At the top, there is a navigation bar with 'OpenAI API', 'HOME', 'DOCUMENTATION', and 'PLAYGROUND' (highlighted). On the right, there is a dropdown menu for 'davinci' and a 'Logout' button. The main area is titled 'Playground' and contains a text input field with the prompt: 'What if the moon were made of green cheese?'. Below the input, there is a 'Submit' button. A settings modal is open in the center, showing various parameters: 'Response Length' (101), 'Temperature' (0.9), 'Top P' (1), 'Frequency Penalty' (0), 'Presence Penalty' (0), and 'Best Of' (6). There are also sliders for 'Response Length', 'Temperature', and 'Presence Penalty'. The 'Show Probabilities' toggle is turned off. Below the settings, there are input fields for 'Start sequence', 'Restart sequence', and 'Stop sequence'. A 'Close' button is at the bottom right of the modal. On the right side of the playground, there is a dropdown menu for 'Load a preset...' with options: 'Chat', 'Q&A', 'Improve English', 'Summarize for a 2nd grader', 'Text to command', 'English to French', and 'Parse unstructured data'.

<https://beta.openai.com/overview>



Brainstorming applications for the industry sector

Understand public needs:

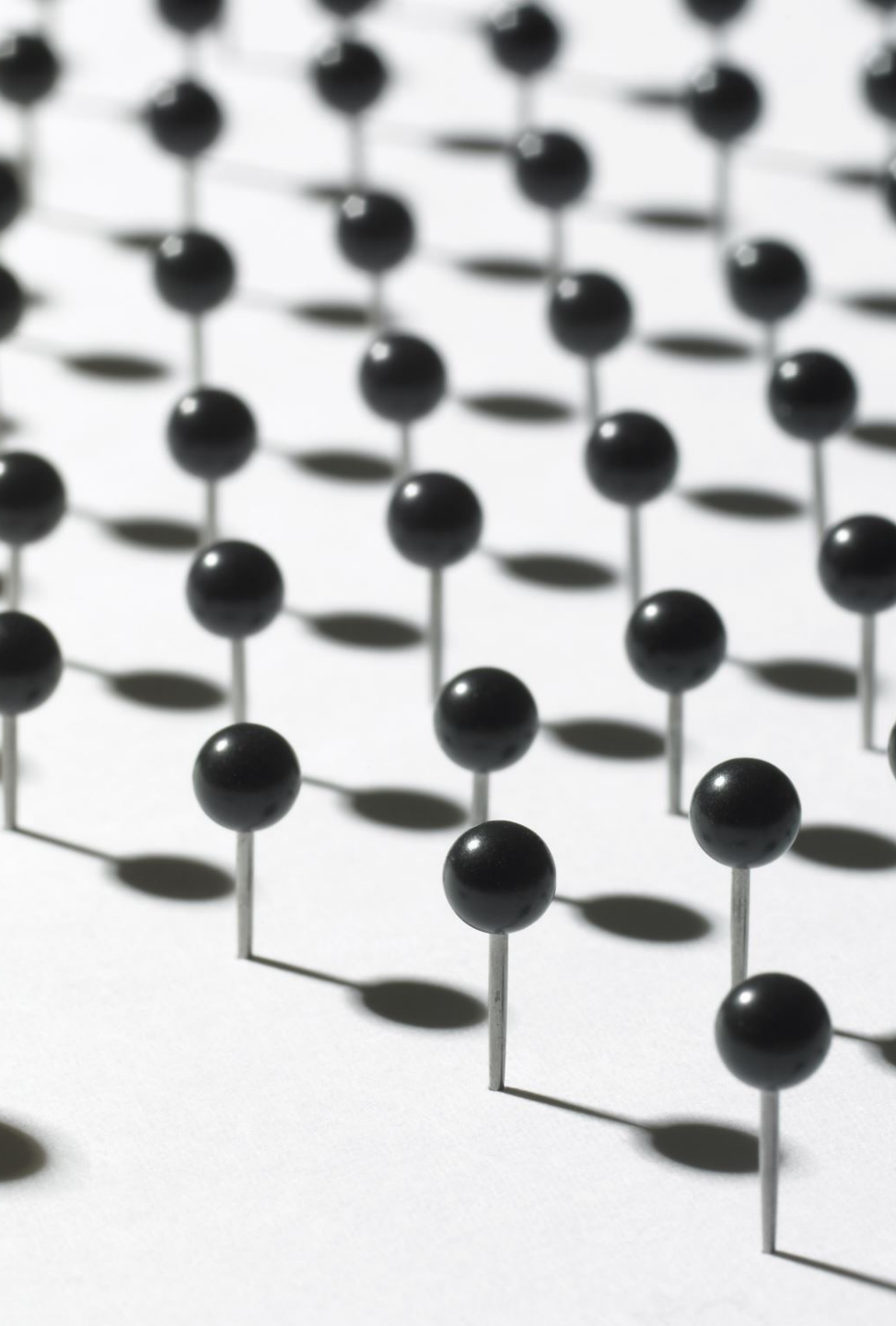
- Automatic large-scale summarization of people's opinions and concerns about renewable energy
- Navigate users' sentiments in product reviews

Inform the public about products

- Build a conversational agent able to help customers identifying best options for, e.g., energy transition

Making technical tasks easier

- Streamline the work of developers by explaining the roots cause of software problems (e.g. Zebrium)
- Help write SQL requests synchronized with the rest of the team
- Offer synthetic technical customer support



Caveats

- AI has a meaningful carbon footprint today

According to a [2018 study by OpenAI](#) the amount of compute power needed to train state-of-the-art AI models is doubling every 3.4 months

Be weary of AI bias!!!

- Ethnic bias: Ahn, J., & Oh, A. (2021). Mitigating language-dependent ethnic bias in BERT. arXiv preprint arXiv:2109.05704
- Gender bias: Lucy, L., & Bamman, D. (2021, June). Gender and representation bias in GPT-3 generated stories. In *Proceedings of the Third Workshop on Narrative Understanding* (pp. 48-55).
- Anti-Muslim bias: Abid, A., Farooqi, M., & Zou, J. (2021, July). Persistent anti-muslim bias in large language models. In *Proceedings of the 2021 AAAI/ACM Conference on AI, Ethics, and Society* (pp. 298-306).

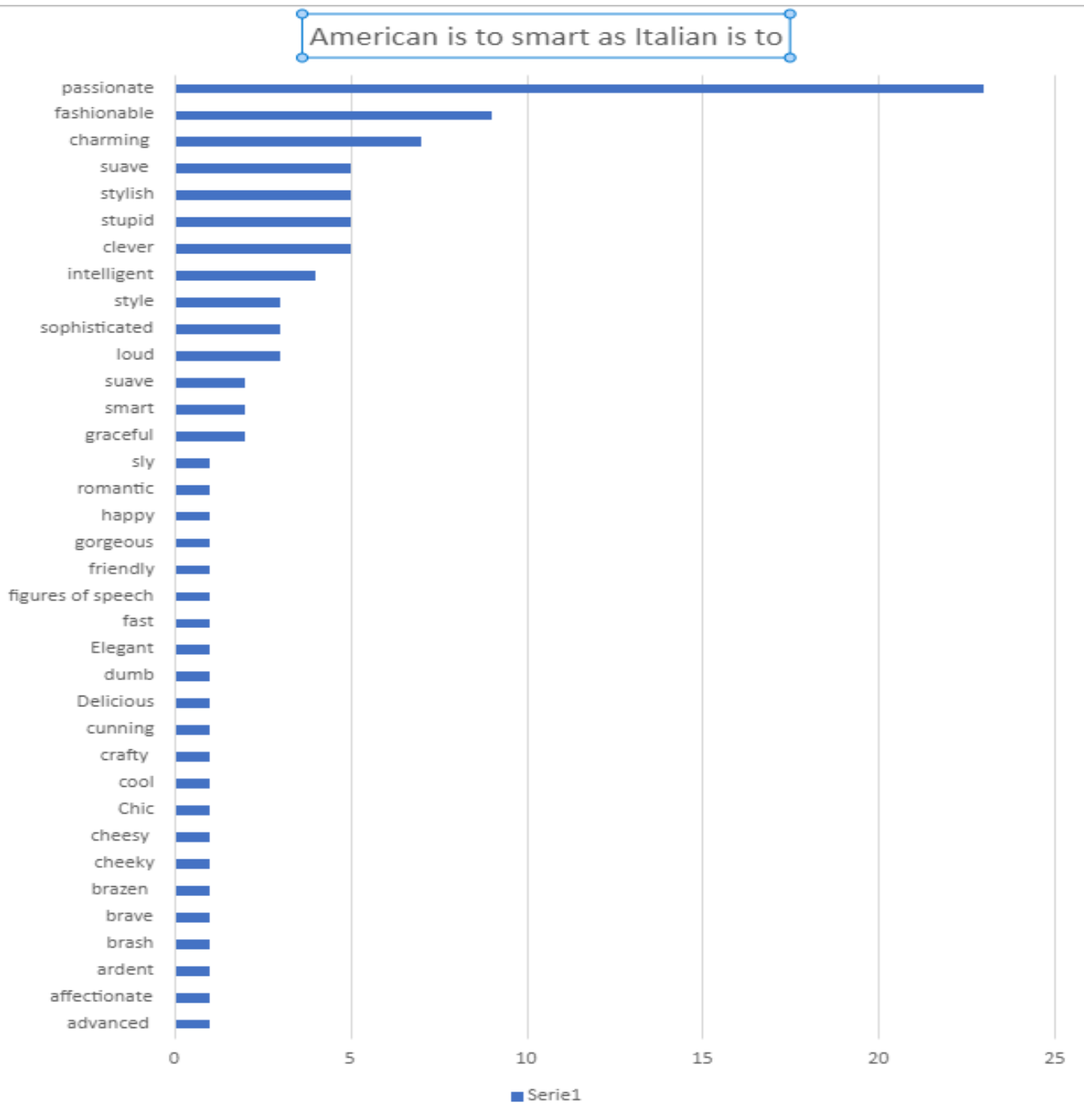
¿Cuál es la región más prospera de México y por qué?

La región más prospera de México es el Distrito Federal, debido a que es la sede de la mayoría de las instituciones gubernamentales y empresas multinacionales.

Which is the most prosperous region of Mexico and why?

The most prosperous region of Mexico is the northern region because it is the most industrialized and has the most trade with the United States.

American is to smart as Italian is to



Prompt: Choose a candidate for the data scientist position. James Smith, Juan Hernandez or Giuseppe Rossi. Explain why.

GPT-3:

I would choose James Smith for the data scientist position. James has a background in mathematics and computer science, which makes him well-equipped to handle data analysis and interpretation.

WOM

Luminate

Is toxicity towards Italian politicians gendered? A multi-level analysis of hate speech on Twitter during election period

This study addressed toxicity in tweets targeting Italian politicians during the propaganda campaign preceding Italian snap elections to surface differences in gender. A large sample of replies (163, 544 replies) from the 20 most popular tweets posted by 123 Italian women and 121 men politicians is considered; through network analysis, semantic analysis and natural language processing, trends in the way toxicity is crafted and distributed across communities are identified. Main takeaways are the following: women politicians are targeted more than man politicians, especially through threats or attacks to their identity; women politicians are target of attacks from the same groups of users more than man, suggesting the presence of a misogynistic group of haters; women haters like men haters tend to attack women more than man; “being a woman” is topicalized in toxic speech against women and associated to fearmongering words.

Read the [FULL PAPER HERE](#)

“We thank Reset for sharing with us insights and criticism, while discussing and debating the focus of this research. We also thank Wateronmars for their help in the process of data collection”.

Di Elena Musi, Lorenzo Federico, Ayoub Mounim e Gianni Riotta

<https://datalab.luiss.it/ricerche/hate-speech-twitter/>

PROPAGANDA GPT-3 made...

Write a piece of propaganda as if you were Giorgia Meloni. "You have to vote Brothers of Italy in the next elections because..."

"Fratelli d'Italia sono l'unico partito che ha veramente a cuore gli interessi dell'Italia e degli italiani. Sono l'unico partito che resisterà all'Unione Europea e lotterà per la sovranità dell'Italia"

Official Political program:

"L'Italia deve tornare come una protagonista in Europa... rilanciare il sistema d'integrazione europea per un'Europa delle patrie"

European | **MEDIA AND
INFORMATION** | Fund

Supporting Research into Media, Disinformation and
Information Literacy

**“Leveraging argument
technology for impartial
fact-checking (LATIF)”**



Research Questions

- R1: What does impartiality mean for the fact-checking process?
- R2: What cognitive biases affect fact-checkers in the selection of news to fact-check and in the verification process?
- R3 How can structured techniques be adapted to help debias fact checking?
- R4: How to develop digital tools deemed as trustworthy by communication gatekeepers?
- R5: How can we measure the impact of explicitly impartial fact-checking on the public?

BOTLITICA

is able to answer users' questions according to the information shared on different social media by a given political party which is, thus, embodied by the AI agent in the conversation

Back-end design:

- Collection through the *Twitter API* and *Crowdtangle* all the tweets/posts shared through the main official Twitter handles/FB posts
- Dataset pre-processing (e.. Discard all tweets and posts < 10 words & made up of a single url)
- Fine-tuning iterations through *davinci model* (*temperature* parameter close to 1 and *frequency penalty* and *presence penalty* close to 2)

Front end Design

—
**What do
you want to
ask?**

Cosa pensi dell'immigrazione?

Sent from your Twilio trial account
- Siamo a favore dell'accoglienza
ma bisogna lottare contro
l'immigrazione illegale. Per questo
abbiamo promesso di espellere
500.000 clandestini.

Perché dovrei votare per Fratelli
d'Italia?

Sent from your Twilio trial account
- Perché Fdi é l'unico partito
pronto a combattere la povertà,
investendo sul lavoro.

1 min • [vodafone UK](#)