

Masterclass in Social Research

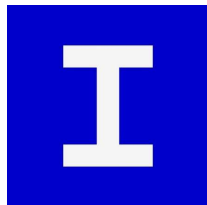
AI for Social Research: Qualitative

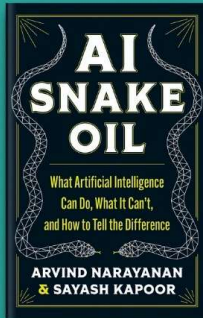






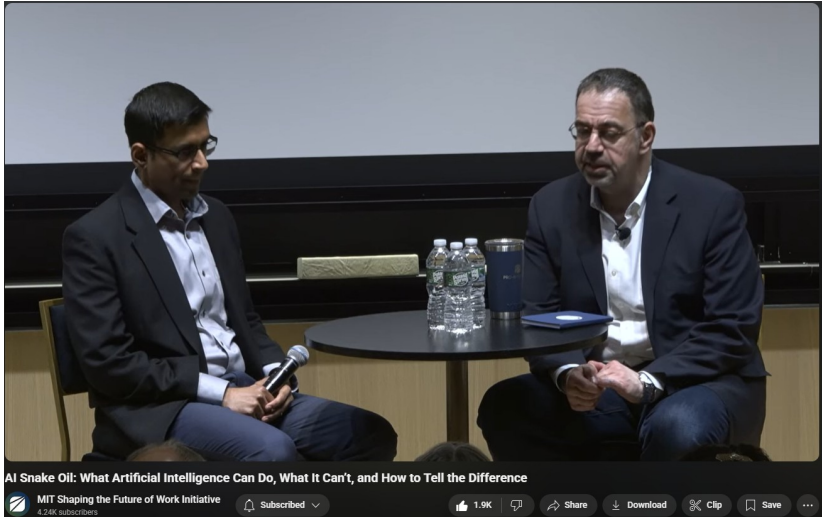
The
Alan Turing
Institute





<https://www.aisnakeoil.com/>





<https://www.youtube.com/watch?v=C3TqcUEFR58>



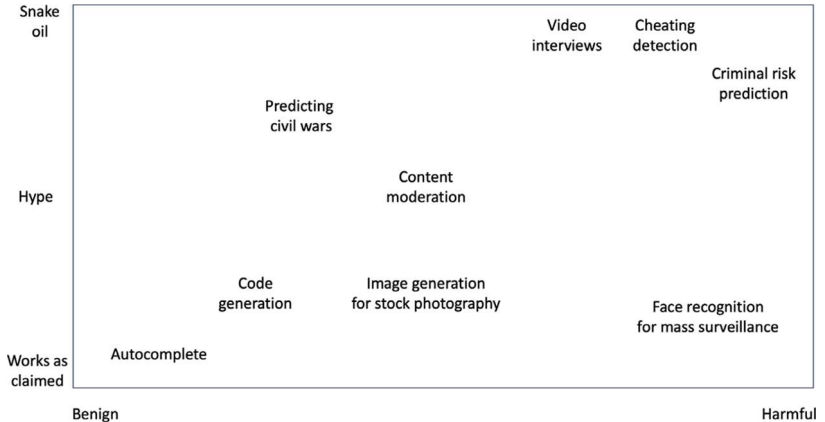


FIGURE 1.2. The landscape of AI snake oil, hype, and harms, showing a few illustrative applications.





Machine learning: Supervised, unsupervised and reinforcement

- ▶ Classification (or regression with regularization) – spam emails
- ▶ Pattern recognition (or clustering) – facial recognition
- ▶ Prediction (or forecasting) – stock market forecasting and GPT (generative predictive transformer)



Machine Learning

Supervised
Learning



Task Driven
(Classification/Regression)

Unsupervised
Learning



Data Driven
(Clustering)

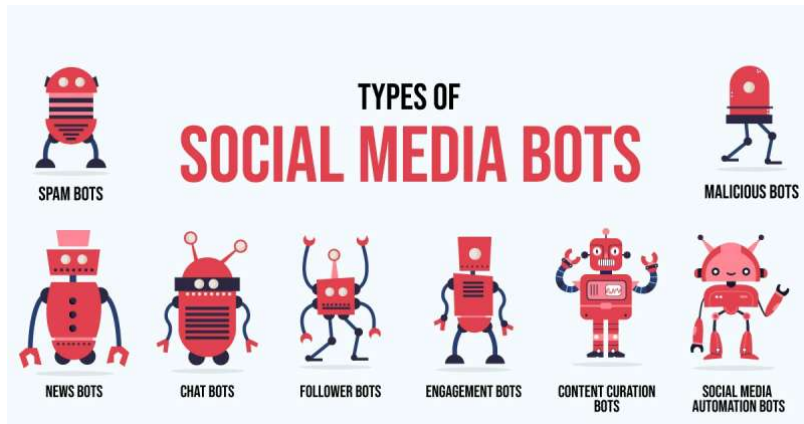
Reinforcement
Learning



Learning from
mistakes
(Playing Games)



Example: Bots on social media



<https://www.socialfollowers.uk/blogs/how-social-media-bots-work-and-how-to-detect-them/>



Example: Vocal pitch of women politicians

American Political Science Review (2019) 113, 4, 941–962

doi:[10.1017/S0003055419000467](https://doi.org/10.1017/S0003055419000467)

© American Political Science Association 2019

Pitch Perfect: Vocal Pitch and the Emotional Intensity of Congressional Speech

BRYCE J. DIETRICH *University of Iowa*

MATTHEW HAYES *Rice University*

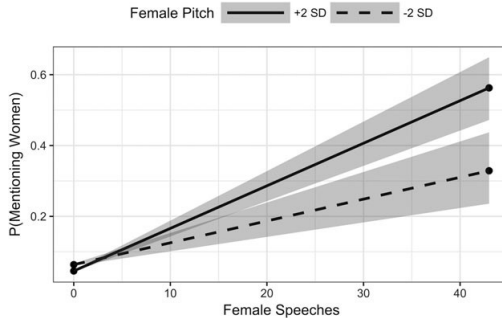
DIANA Z. O'BRIEN *Rice University*

Although audio archives are available for a number of political institutions, the data they provide receive scant attention from researchers. Yet, audio data offer important insights, including information about speakers' emotional states. Using one of the largest collections of natural audio ever compiled—74,158 Congressional floor speeches—we introduce a novel measure of legislators' emotional intensity: small changes in vocal pitch that are difficult for speakers to control. Applying our measure to MCs' floor speeches about women, we show that female MCs speak with greater emotional intensity when talking about women as compared with both their male colleagues and their speech on other topics. Our two supplementary analyses suggest that increased vocal pitch is consistent with legislators' broader issue commitments, and that emotionally intense speech may affect other lawmakers' behavior. More generally, by demonstrating the utility of audio-as-data approaches, our work highlights a new way of studying political speech.

<https://doi.org/10.1017/S0003055419000467>



FIGURE 1. The Quantity and Intensity of Women's Speech Affects the Quantity of Men's Speeches about Women



Note: Predicted male speaking behavior from Model 2 in Table 5 holding all other variables constant. Solid and dashed lines indicate Female Pitch was set to two standard deviations above (1.41) and below (-1.28) the mean, respectively. On the x-axis, Female Speeches is allowed to vary from its minimum (0) to maximum (43). The y-axis is the probability that the male speech included any of the Pearson and Dancey (2011b) women's dictionary terms. The gray ribbons represent 90% confidence intervals. The 95% confidence intervals overlap until the x-axis reaches approximately 25 speeches.



Example: Factional politics in post-1949 China

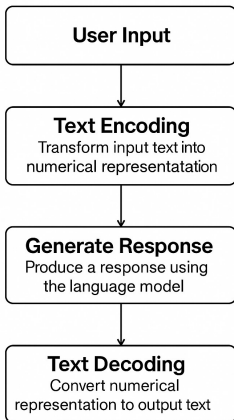


Generative AI, GPT and large language models (LLMs)

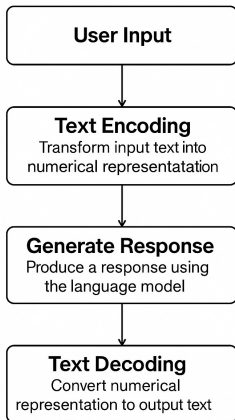
- ▶ Fine tuning and/or prompt engineering?
- ▶ General and/or domain-specific LLMs?
- ▶ Metrics for performance evaluation: Intelligent? accurate? efficient (computationally)? transparent? safe? ethical? democratic? open access?



How ChatGPT Works



How ChatGPT Works



- ▶ Is it intuitive to understand?
- ▶ Can you try to come up with your prompt to create a (similar) flowchart?
- ▶ What is the prompt you use? Share it with us.



Italy's regulator blocks Chinese AI app DeepSeek on data protection

By Reuters

February 4, 2025 5:02 PM GMT · Updated 3 months ago



The Deepseek logo is seen in this illustration taken on January 29, 2025. REUTERS/Dado Ruvic/Illustration/File Photo [Purchase Licensing Rights](#) 



Exercise: Running critical discourse analysis with AI

- ▶ **Step 1:** Ask AI to produce a brief definition of critical discourse analysis – critically evaluate the response provided by AI
 - What response do you get? Do you agree with it or not? Is it the same as what Margarita discussed last week?
 - Can you try to engage with AI to "improve" the definition?
- ▶ **Step 2:** Upload the "Return to Work" letter and ask AI to carry out a critical discourse analysis based on the definition – what do you get?
- ▶ **Questions to think about**
 - Are you comfortable with what we have just done? Why not? What might be the caveat (Morgan 2023)?
 - Does it make sense to analyze the "discourse" produced by AI or Human-AI interactions? How?



Exercise: Analyzing interview transcripts

- ▶ **Source:** "Urban conceptions of nature during the ongoing COVID-19 pandemic" (Matthew Wood 2021) – see <https://researchdata.bbk.ac.uk/id/eprint/161/>
- ▶ **Step 1:** Download the data and skim through the information on BiRD
- ▶ **Step 2:** Replicate the prompts used by Morgan (2023) to analyze the interview transcripts
- ▶ **Questions to think about**
 - Are you happy with Morgan's prompts? Can you get around the limitations or issues pointed out by Morgan?
 - How do you evaluate the quality of the analysis produced by AI? What might be some useful criteria?
 - How do you know AI is more or less nuanced when it comes to coding and **interpretation**?



Exercise: Conducting interviews using AI agent (Geiecke and Jaravel 2024)

Conversations at Scale: Robust AI-led Interviews with a Simple Open-Source Platform*

Friedrich Geiecke

Xavier Jaravel

LSE

LSE & CEPR

f.c.geiecke@lse.ac.uk

x.jaravel@lse.ac.uk

November 20, 2024

Abstract

The advent of large language models (LLMs) provides an opportunity to conduct qualitative interviews at a large scale, with thousands of respondents, creating a bridge between qualitative and quantitative methods. In this paper, we develop a simple, versatile open-source platform for researchers to run AI-led qualitative interviews. Our approach incorporates established best practices from the sociology literature, uses only a single LLM agent with low latency, and can be adapted to new interview topics almost instantaneously. We assess its robustness by drawing comparisons to human experts and using several respondents-based quality metrics. Its versatility is illustrated through four broad classes of applications: eliciting key factors in decision making, political views, views of the external world, and subjective mental states. High performance ratings are obtained in all of these domains. The platform is easy to use and deploy: we provide detailed explanations and code for researchers to swiftly set up and test their own AI-led interviews. In addition, we develop, validate, and share a simple LLM-based pipeline for textual analysis and coding of large volumes of interview transcripts.

Keywords: qualitative interviews; large language models; surveys.



Concluding remarks

- ▶ Combining the use of AI with other software packages for qualitative research (e.g., ATLAS.ti and MAXQDA)?
- ▶ "Strategic ignorance" is key – the priority is to use the tool and treat LLM as a "black" box?
- ▶ End of "lone wolf" social researchers? Check out the **Turing Way**
- ▶ "Only because we can does not mean we should" (from Jurassic Park)
– do you agree?
- ▶ Substitute versus empowerment?



THERE'S MORE TO COLLABORATION

THAN YOU MIGHT THINK!



Scribbleria

<https://book.the-turing-way.org/collaboration/collaboration>

