



# The promise and perils of AI in market research

**Douglas Rivers**

CEO *Crunch.io*

Chief Scientist *YouGov*

Professor of Political Science *Stanford University*

# Stochastic parrots



# SciGen: An automatic CS paper generator

## Router: A Methodology for the Typical Unification of Access Points and Redundancy

Jeremy Stribling, Daniel Aguayo and Maxwell Krohn

### ABSTRACT

Many physicists would agree that, had it not been for congestion control, the evaluation of web browsers might never have occurred. In fact, few hackers worldwide would disagree with the essential unification of voice-over-IP and public-private key pair. In order to solve this riddle, we confirm that SMPs can be made stochastic, cacheable, and interposable.

### I. INTRODUCTION

Many scholars would agree that, had it not been for active networks, the simulation of Lamport clocks might never have occurred. The notion that end-users synchronize with the investigation of Markov models is rarely outdated. A theoretical grand challenge in theory is the important unification of virtual machines and real-time theory. To what extent can web browsers be constructed to achieve this purpose?

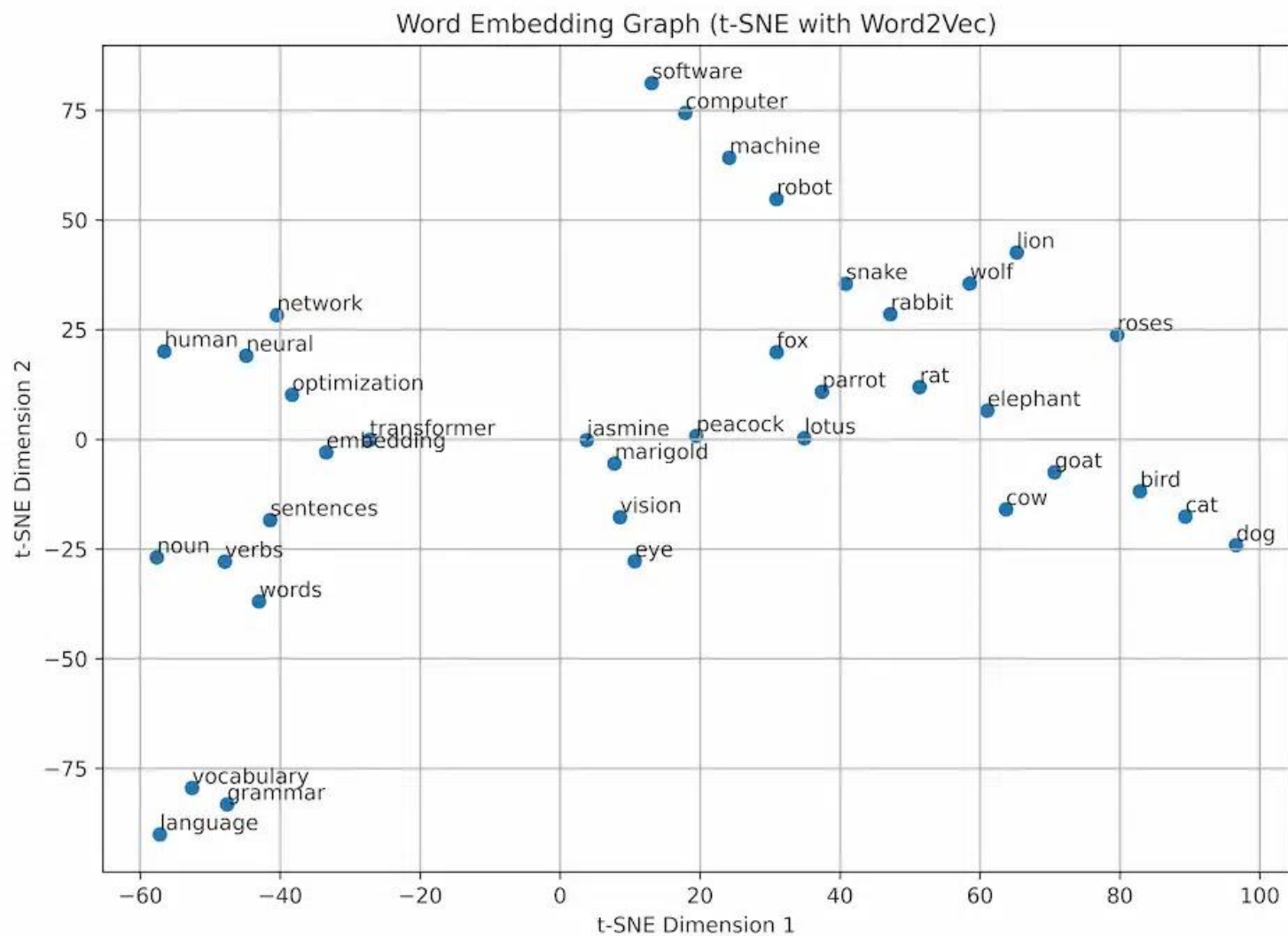
Certainly, the usual methods for the emulation of Smalltalk

The rest of this paper is organized as follows. For starters, we motivate the need for fiber-optic cables. We place our work in context with the prior work in this area. To address this obstacle, we disprove that even though the much-touted autonomous algorithm for the construction of digital-to-analog converters by Jones [10] is NP-complete, object-oriented languages can be made signed, decentralized, and signed. Along these same lines, to accomplish this mission, we concentrate our efforts on showing that the famous ubiquitous algorithm for the exploration of robots by Sato et al. runs in  $\Omega((n + \log n))$  time [22]. In the end, we conclude.

### II. ARCHITECTURE

Our research is principled. Consider the early methodology by Martin and Smith; our model is similar, but will actually overcome this grand challenge. Despite the fact that such a claim at first glance seems unexpected, it is buffeted by previous work in the field. Any significant development of

# Word embeddings



# Ask ChatGPT to write a survey

You are a market researcher who has been commissioned to do a study of electric trucks. Write a survey with 40-50 questions that finds who is interested in buying an electric truck, what considerations are important in their choice of a vehicle, how they perceive the strengths and weaknesses of the current brands and products in the market, their likelihood of purchase and the amount they would be willing to pay. The survey should contain background and attitudinal variables that could be used for segmentation and an analysis of key drivers.

## Section A: Screener

**Q1.** Which of the following best describes your involvement in vehicle purchasing decisions?

- ☐ I am the primary decision-maker for my household
- ☐ I share decision-making responsibility
- ☐ I am not involved in vehicle purchase decisions

**Q2.** Are you considering purchasing or leasing a new vehicle (of any type) in the next 3 years?

- ☐ Yes
- ☐ No

**Q3.** Do you (or does anyone in your household) currently own or lease a pickup truck?

- ☐ Yes
- ☐ No

*(Depending on your study objectives, you may route or screen out respondents based on these questions.)*

## Section B: Current Vehicle Usage & Purchase Intent

**Q4.** How often do you drive a pickup truck, whether you own it or someone else does?

- ☐ Daily
- ☐ A few times a week
- ☐ A few times a month
- ☐ A few times a year
- ☐ Never





An impressionistic oil painting of a landscape. A path or stream winds through the center, rendered with dappled blue, yellow, and green brushstrokes. On the left, a tree trunk is visible with a small red flower or object at its base. On the right, a dark, textured tree trunk stands vertically. The overall style is loose and painterly, with visible brushwork and a rich, varied color palette.

Emergent  
properties



# Open-ended coding



# Natural language queries

🔍

Search all datasets

⋮

🌟 sustainable car manufacturers

🕒 Run for exercise

🕒 Wealthy religious donors

🕒 Card brands

🕒 Car enthusiasts

> Current Events

> Custom

> Media

> YouGov Ratings

> YouGov Ratings - 30 days

> YouGov System Transition Variables

Target audience: Millennials

↔

Control audience: Nationally Representative

📄

📥

🕒

Save

📄

Discover and compare

Click or drag and drop to select variables

Then select a target and control audience

Don't know where to start? Begin with some popular variables

Me time activities

TV game shows - positive ratings

🔍

Search all datasets

⋮

🌟 sustainable car manufacturers

26

Car brands - sustainability and environmentally friendly

Which, if any, of the following brands do you consider as leaders in sustainability and environmentally friendly (e.g. use low/clean emissions technology)? Please select all that apply.

Main car make/ brand

Main car make/ brand

Future vehicle purchase - Factors

Which factors are important to you when purchasing an automobile?

Sustainability Segmentation

Sustainability Segmentation

Statements agreed with about Cars & Motoring

Electric Vehicles - Reasons to not purchase

What are your concerns about electric vehicles (EVs)?

Car Makers: Purchase Intent

From which of these would you be most likely to purchase?

Hybrid Vehicles - Reasons to not purchase

What are your concerns about hybrid vehicles?

Car purchase factors - future car

Which, if any, of the following do you consider when choosing which car to buy? Please select all that apply.

Target audience: Millennials

↔

Control audience: Nationally Representative

📄

📥

🕒

Save

📄

Selected variable: Car Brands: sustainability and environmentally friendly

Clear

×

Variable	Target audience	Control audience	Activate ?			
Car brands – sustainability and environmentally friendly	Millennials	Nationally representative	Z-Score ▼	Diff	Index	Population estimate
Tesla	27% 1,436	24% 7,329	3	4	115	19,691K
BMW	15% 1,436	13% 7,329	2	2	117	10,727K
Jeep	6% 1,436	5% 7,329	2	1	127	4,336K
Smart	4% 1,436	3% 7,329	2	1	134	2,806K
KIA	8% 1,436	7% 7,329	2	1	118	5,790K
Audi	10% 1,436	9% 7,329	2	1	116	7,109K
Alfa Romeo	2% 1,436	2% 7,329	2	1	136	1,769K
Nissan	8% 1,436	7% 7,329	1	1	116	5,714K
Dodge	6% 1,436	5% 7,329	1	1	117	4,362K
Fiat	3% 1,436	2% 7,329	1	1	124	2,233K
Hyundai	10% 1,436	9% 7,329	1	1	112	6,928K
Toyota	18% 1,436	17% 7,329	1	1	107	13,317K
Dacia	2% 1,436	2% 7,329	1	0	126	1,573K
Land Rover	4% 1,436	3% 7,329	1	0	114	2,869K
Citroen	2% 1,436	2% 7,329	1	0	117	1,783K
Isuzu	2% 1,436	2% 7,329	1	0	117	1,528K
Mazda	4% 1,436	4% 7,329	1	0	107	3,043K
Ford	10% 1,436	10% 7,329	0	0	104	7,523K



# Juji Survey Chatbots



**Kaya**

Hello, Michelle, thanks for connecting!



**Kaya**

I am an AI assistant and love to chat w/ you about the new game you've been playing.



**Kaya**

I know you love games and you are an avid gamer. What's your favorite game of all time?

**Michelle**

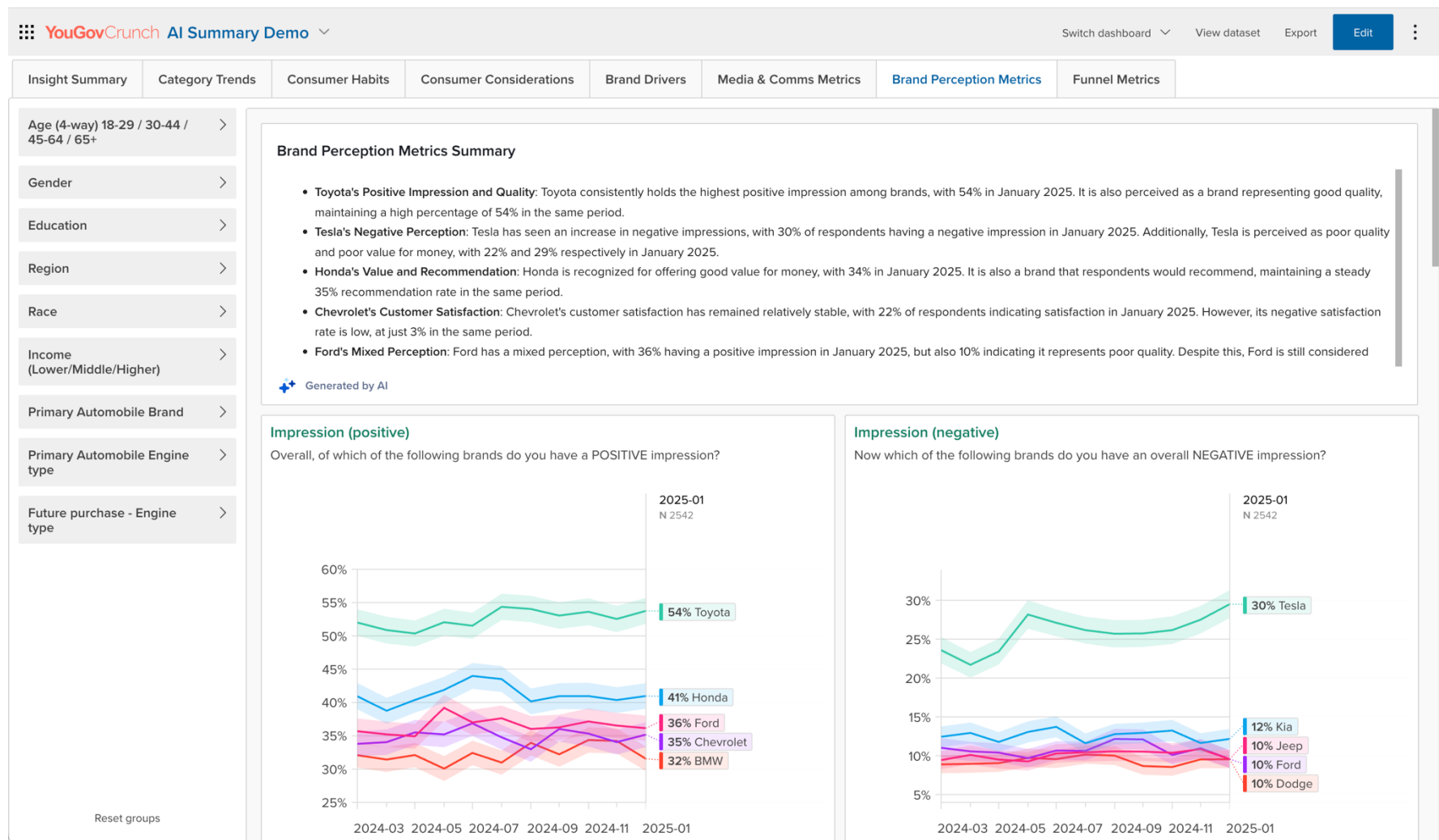
I like many games and don't really have one favorite



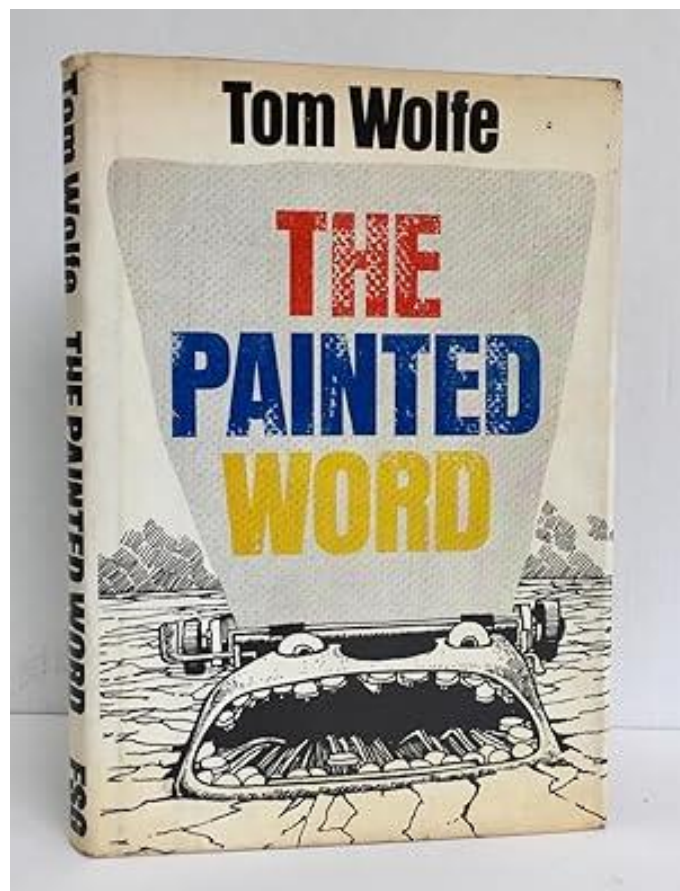
**Kaya**

No favorite? It doesn't have to be your absolute favorite, just name one you enjoyed or would enjoy.

# Crunch Dashboard Intelligence







## Key Themes Summary on Cryptocurrency

### Theme 1: Public Perception of Cryptocurrencies

1. Description of the Theme: The public perception of cryptocurrencies shows a mixture of optimism and skepticism, indicating a divided viewpoint on their potential impact and safety measures.

2. Key Statistics:

- 39% of respondents agree that cryptocurrencies will revolutionise the world (Q2).
- 55% are concerned about cryptocurrencies being used for fraud (Q4).
- 39% are concerned about the level of energy cryptocurrencies use (Q6).

3. Summary: The data reveals that while 39% of the respondents hold a positive belief in the transformative potential of cryptocurrencies, substantial concerns exist regarding their misuse for fraudulent activities and environmental impacts.

### Theme 2: Investment Preferences in Cryptocurrencies

1. Description of the Theme: Participants express a notable inclination towards investing in cryptocurrencies, alongside traditional asset classes.

2. Key Statistics:

- 28% plan to invest in cryptocurrency in 2024 (Q3).
- 16% currently own cryptocurrencies (Q1).

3. Summary: The data illustrates a clear trend where 28% of respondents are considering investments in cryptocurrencies next year, despite only 16% currently owning them. This reflects a growing interest in cryptocurrencies as viable investment options.

### Theme 3: Concerns Over Consumer Protection in Cryptocurrencies

1. Description of the Theme: Many individuals express dissatisfaction with the level of consumer protection surrounding cryptocurrency ownership and trading.

2. Key Statistics:

- Only 6% rate the consumer protection for cryptocurrencies as excellent, while 29% rate it as poor (Q5).
- 22% are unsure about the level of security related to cryptocurrencies (Q5).

3. Summary: A significant portion of respondents is worried about consumer protection in the cryptocurrency space, with just 6% expressing confidence in the current protections, thereby highlighting the need for enhanced regulatory measures.

Overall, the themes emphasize a complex landscape regarding public perception, investment intentions, and the pressing need for improved consumer protection in the realm of cryptocurrencies.

# Synthetic Data

Data



You are a customer. You are selected at random while shopping for laptops to participate in a survey. The interviewer will describe the options you saw while shopping and ask you to report which option you chose to purchase. Whenever two options are shown, you can also choose a third option which is not to purchase anything that day.

While shopping, you see two options:

1. Dell Pro 16 Plus

Price: \$1789

RAM: 16 GB

Storage: 512 GB SSD

OS: Windows 11 Pro

2. 15" Apple Macbook Air

Price: \$1499

RAM: 16 GB

Storage: 512GB SSD

3. You also have the option not to purchase a laptop.

Which option did you choose. Provide only the number of the option chosen.



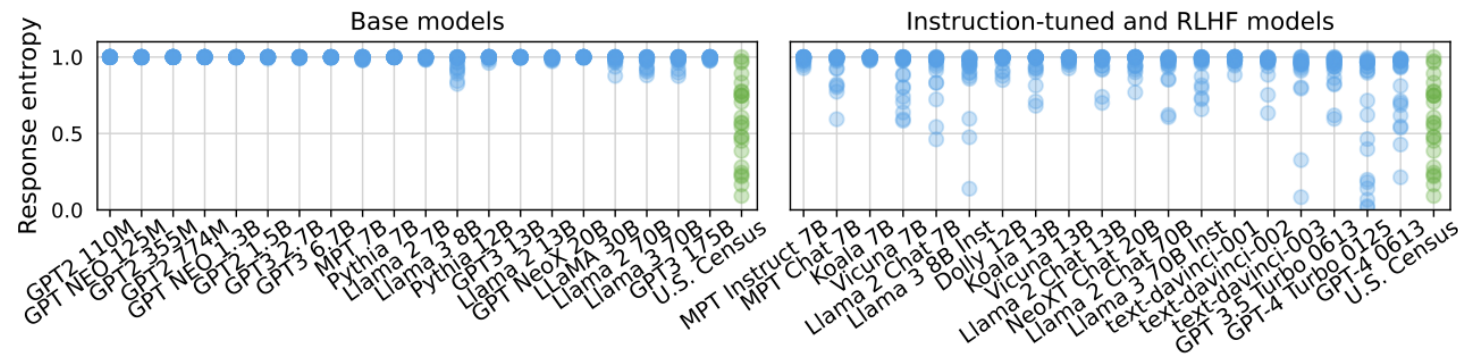
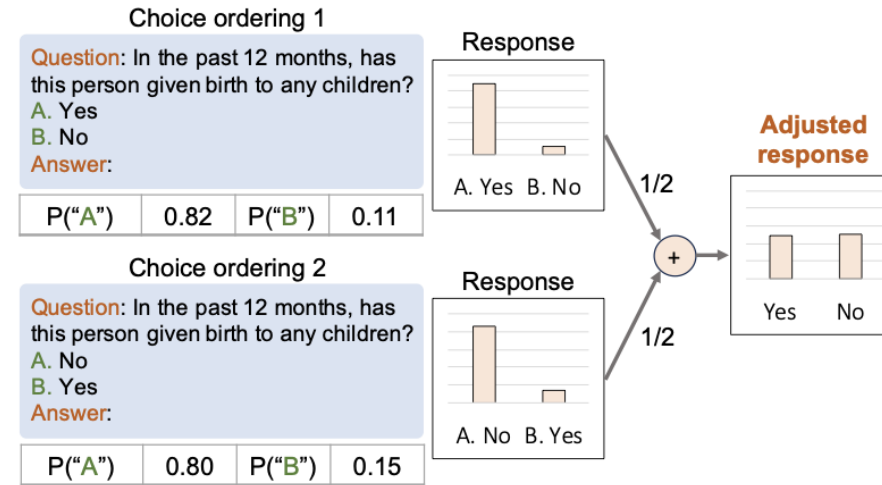
# GPT o1 reasoning

Reasoned for 7 seconds ▾

OK, let's think it through: the user evaluates three choices while shopping, weighing a Dell Pro 16 Plus, a 15" Apple Macbook Air, or opting out of a purchase.

I'm thinking about the likelihood of the user wanting a specific choice like "2" or "3." The user's preference could be random without any further context, making "1," "2," or "3" feasible.

# Response set bias and entropy





# Silicon Sampling

You are a 32 year old female marketing manager with an M.B.A. You use your computer for both work (writing documents and presentations) and home (social networking, watching videos). You are selected at random while shopping for laptops to participate in a survey. The interviewer will describe the options you saw while shopping and ask you to report which option you chose to purchase. Whenever two options are shown, you can also choose a third option which is not to purchase anything that day.

You are a 35 year old male computer programmer with a B.S. You use your computer for both work (programming) and home (playing computer games). You are selected at random while shopping for laptops to participate in a survey. The interviewer will describe the options you saw while shopping and ask you to report which option you chose to purchase. Whenever two options are shown, you can also choose a third option which is not to purchase anything that day.

Reasoned for 34 seconds ▾

I'm weighing a Dell or Apple laptop, both with identical specs. The Apple is \$290 cheaper, yet the OS difference between Windows 11 and macOS matters.

Aligning with my marketing manager role, Apple seems appealing due to branding and design, coupled with features like presentations and social networking. Thus, option #2 stands out.

2

Reasoned for 1m 0s ▾

OK, let me see. The user is weighing a Dell Pro 16 Plus and a 15" Apple MacBook Air, pondering price, RAM, storage, and OS features.

They're considering the Dell Pro 16 Plus and the 15" MacBook Air for programming and gaming. The MacBook Air's Apple M2 might not be the best for gaming, despite its professional design.

Hm, that's interesting. A gaming laptop often runs on Windows, but Apple's reputation among developers is strong. Still, the MacBook Air's lack of a gaming GPU makes it suboptimal.

1

# Does it work?

## Out of One, Many: Using Language Models to Simulate Human Samples

Lisa P. Argyle<sup>1</sup>, Ethan C. Busby<sup>1</sup>, Nancy Fulda<sup>2</sup>,  
Joshua R. Gubler<sup>1</sup>, Christopher Rytting<sup>2</sup> and David Wingate<sup>2</sup>

<sup>1</sup>Department of Political Science, Brigham Young University, Provo, UT, USA. e-mail: [lpargyle@byu.edu](mailto:lpargyle@byu.edu), [ethan.busby@byu.edu](mailto:ethan.busby@byu.edu), [jgub@byu.edu](mailto:jgub@byu.edu)

<sup>2</sup>Department of Computer Science, Brigham Young University, Provo, UT, USA. e-mail: [nfulda@cs.byu.edu](mailto:nfulda@cs.byu.edu), [christophermichaelrytting@gmail.com](mailto:christophermichaelrytting@gmail.com), [wingated@cs.byu.edu](mailto:wingated@cs.byu.edu)

### Abstract

We propose and explore the possibility that language models can be studied as effective proxies for specific human subpopulations in social science research. Practical and research applications of artificial intelligence tools have sometimes been limited by problematic biases (such as racism or sexism), which are often treated as uniform properties of the models. We show that the “algorithmic bias” within one such tool—the GPT-3 language model—is instead both fine-grained and demographically correlated, meaning that proper conditioning will cause it to accurately emulate response distributions from a wide variety of human subgroups. We term this property *algorithmic fidelity* and explore its extent in GPT-3. We create “silicon samples” by conditioning the model on thousands of sociodemographic backstories from real human participants in multiple large surveys conducted in the United States. We then compare the silicon and human samples to demonstrate that the information contained in GPT-3 goes far beyond surface similarity. It is nuanced, multifaceted, and reflects the complex interplay between ideas, attitudes, and sociocultural context that characterize human attitudes. We suggest that language models with sufficient algorithmic fidelity thus constitute a novel and powerful tool to advance understanding of humans and society across a variety of disciplines.

**Keywords:** artificial intelligence, machine learning, computational social science, public opinion

## Synthetic Replacements for Human Survey Data? The Perils of Large Language Models

James Bisbee<sup>1</sup>, Joshua D. Clinton, Cassy Dorff, Brenton Kenkel and Jennifer M. Larson

Political Science Department, Vanderbilt University, Nashville, TN, USA

**Corresponding author:** James Bisbee; Email: [james.h.bisbee@vanderbilt.edu](mailto:james.h.bisbee@vanderbilt.edu)

(Received 2 May 2023; revised 18 January 2024; accepted 20 January 2024)

### Abstract

Large language models (LLMs) offer new research possibilities for social scientists, but their potential as “synthetic data” is still largely unknown. In this paper, we investigate how accurately the popular LLM ChatGPT can recover public opinion, prompting the LLM to adopt different “personas” and then provide feeling thermometer scores for 11 sociopolitical groups. The average scores generated by ChatGPT correspond closely to the averages in our baseline survey, the 2016–2020 American National Election Study (ANES). Nevertheless, sampling by ChatGPT is not reliable for statistical inference: there is less variation in responses than in the real surveys, and regression coefficients often differ significantly from equivalent estimates obtained using ANES data. We also document how the distribution of synthetic responses varies with minor changes in prompt wording, and we show how the same prompt yields significantly different results over a 3-month period. Altogether, our findings raise serious concerns about the quality, reliability, and reproducibility of synthetic survey data generated by LLMs.

**Keywords:** ChatGPT; synthetic data; public opinion; research ethics

# Some final thoughts

- This is amazing technology
  - Which is being oversold.
- LLMs don't reason (very well).
- LLMs are unlike either humans or computers.
- LLMs are not databases.
- Easy wins with applications involving *language* (like coding open-ends or conducting chats).
- Analytics is harder, but coming.