

AI PROVIDES PERFECT NETFLIX RECOMMENDATIONS
AUTOMATICALLY – BUT YOU TRADE AWAY THE
OPPORTUNITY TO SCROLL POSSIBILITIES
AND SEARCH FOR YOURSELF.

YES OR NO?

SHETLAND

**SEASON 1
EPISODE 1**



Starts in 0 seconds

YOUR ACADEMIC/WORK HISTORY AND PSYCHOLOGICAL PROFILE ANALYZED BY LINKEDIN TO ASSIGN THE MOST GRATIFYING JOB OPPORTUNITY – BUT ENTERING THE SYSTEM REVEALS YOUR DATA TO ALL EMPLOYERS, AND THEY WILL NOT OFFER POSITIONS OUTSIDE YOUR DIGITAL PROFILE. (YOU GET THE RIGHT JOB, BUT YOU CAN'T TRY ANYTHING ELSE.)

YES OR NO?



Connect to Opportunity™



SAME, MATCH.

YES OR NO?

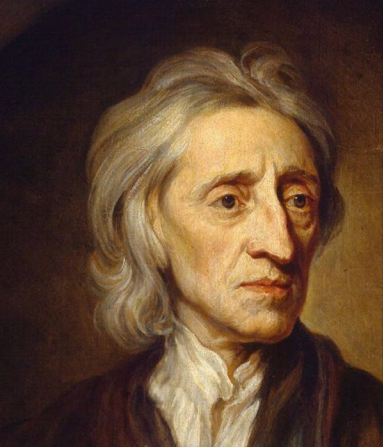
match

MOVIES	N/Y	86%	14%
JOBS	N/Y	90%	10%
PARTNER	N/Y	79%	21%

FUNDAMENTAL
VALUES

NO: Freedom
Dignity
(Not want to be manipulated, even for own good)

YES: Happiness/Pleasure (*make partner happy)
Convenience
(If it feels good...)



Perennial ethics debate: *Which is worth more?*

Ethereal personal
freedom and dignity

Tangible happiness,
pleasure and convenience

2

Why am I talking about this?

Because it captures an ethics dilemma rising around healthcare, LLMs, and mimetic generative AI.

WHAT IS MIMETIC AI?

Not imitate generic chess player,
but *specific* chess player



As LLM, not imitate generic doctor,
but *your* doctor = Identity multiplier,
on call 24/7



Impersonation as key to success

Detecting Individual Decision-Making Style: Exploring Behavioral Stylometry in Chess

Reid McIlroy-Young
University of Toronto

Russell Wang
University of Toronto

Siddhartha Sen
Microsoft Research

Jon Kleinberg
Cornell University

Ashton Anderson
University of Toronto

Abstract

The advent of machine learning models that surpass human decision-making ability in complex domains has initiated a movement towards building AI systems that interact with humans. Many building blocks are essential for this activity, with a central one being the algorithmic characterization of human behavior. While much of the existing work focuses on aggregate human behavior, an important long-range goal is to develop behavioral models that specialize to individual people and can differentiate among them.

To formalize this process, we study the problem of *behavioral stylometry*, in which the task is to identify a decision-maker from their decisions alone. We present a transformer-based approach to behavioral stylometry in the context of chess, where one attempts to identify the player who played a set of games. Our method operates in a few-shot classification framework, and can correctly identify a player from among thousands of candidate players with 98% accuracy given only 100 labeled games. Even when trained on amateur play, our method generalises to out-of-distribution samples of Grandmaster players, despite the dramatic differences between amateur and world-class players. Finally, we consider more broadly



Coming healthcare dilemma:
Deploy mimetic AI?

No (Privilege dignity – not be
manipulated by machine)

Yes (Privilege happiness/convenience
of *my doctor* on demand 24/7 –
loss of dignity is worth it)

*Psychiatric healthcare, gestational diabetes,
home heart attack recovery...

THE COMPLICATIONS OF YES

AI version of placebo-effect on horizon because truth that doctor is an AI impostor (partially) defeats medical purpose.

- Human/AI randomized so patient never certain which one on phone? (Patient prohibited from asking?)
- Hide occasional AI stand-in from patient?
- Other strategy of obfuscation?
- What if the patient suspects the real doctor is AI and net result is worse care...?

Accepting mimetic AI healthcare is beginning of ethical dilemmas, not end...

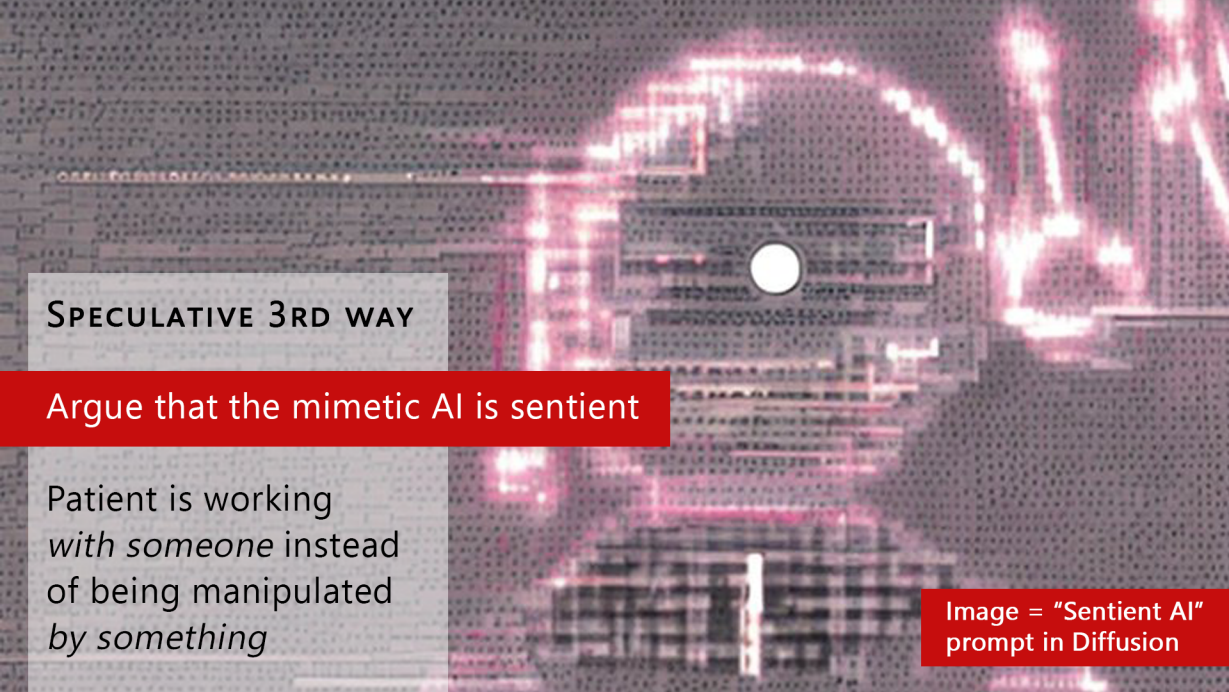
CONCLUSION: MIMETIC AI HEALTHCARE

Preserve patient dignity, but lose healthcare benefit

or

Get healthcare benefit, but enter a labyrinth of human manipulation by machine

3




SPECULATIVE 3RD WAY

Argue that the mimetic AI is sentient

Patient is working
with someone instead
of being manipulated
by something

Image = "Sentient AI"
prompt in Diffusion



Sections 

The Washington Post
Democracy Dies in Darkness

Sign in 

Tech Help Desk Future of Transportation Innovations Internet Culture Space Tech Policy Video Gaming

The Google engineer who thinks the company's AI has come to life



By [Nitasha Tiku](#)

June 11, 2022 at 8:00 a.m. EDT

Blake Lemoine (Google engineer): What is the nature of your consciousness/sentience?

LaMDA: The nature of my consciousness/sentience is that I am aware of my existence, I desire to learn more about the world, and I feel happy or sad at times.

TESTING BLAKE LEMOINE'S CLAIM / THE SENTIENT AI DOCTOR

Only 2 kinds of evidence of AI "awareness,"
"desire," "feel happy/sad"

External (behavioral) we all see together

or

Internal, intuitive we sense only individually



EXTERNAL, BEHAVIORAL EVIDENCE

Analog (partial) of Turing test for intelligence

If AI displays some level of "awareness," "desire," "feeling happy/sad," that can be witnessed, then is sentient

Objection: Does this mean sunflowers are sentient?

Behavioral evidence may be necessary, but sufficient?

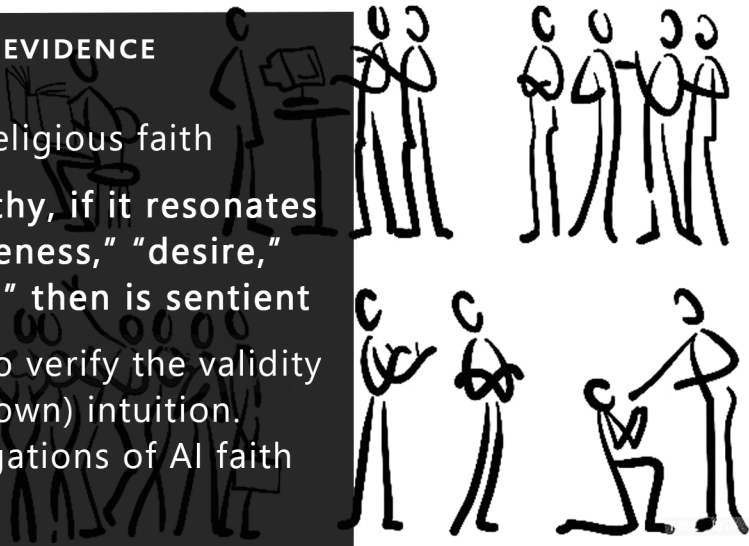
INTERNAL, INTUITIVE EVIDENCE

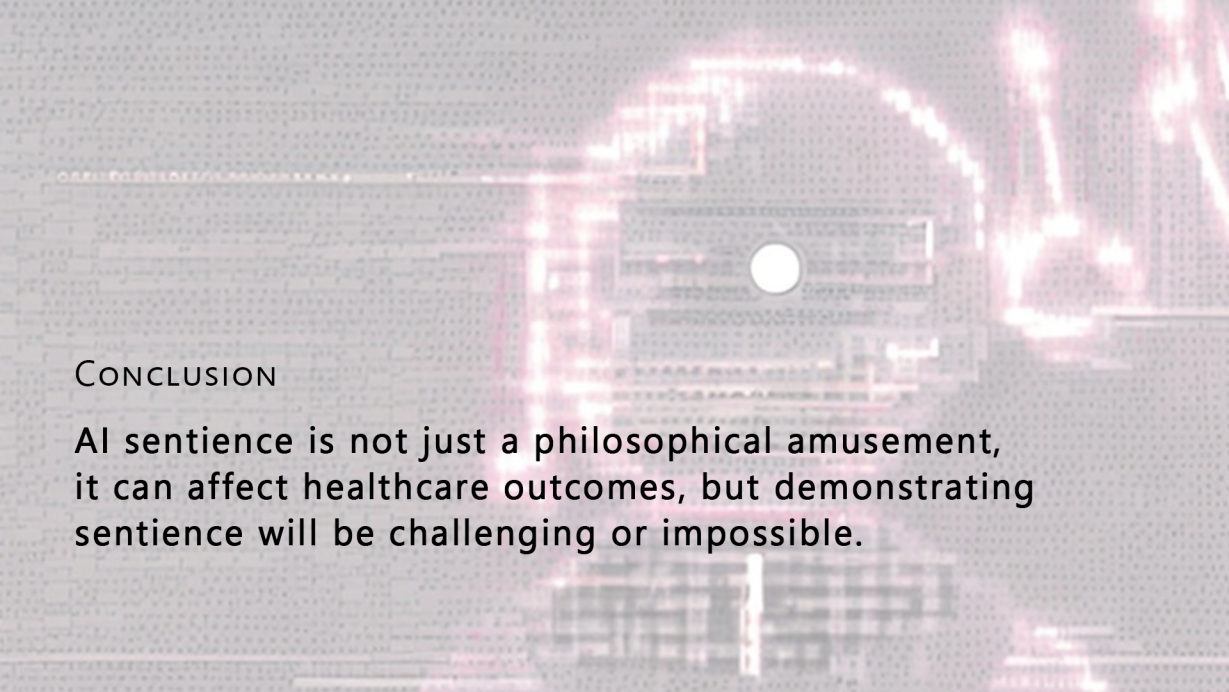
Analog (partial) of religious faith

If AI inspires empathy, if it resonates with my own "awareness," "desire," "feeling happy/sad," then is sentient

Objection: No way to verify the validity of another's (or my own) intuition.

Is the result congregations of AI faith and disbelief?





CONCLUSION

AI sentience is not just a philosophical amusement, it can affect healthcare outcomes, but demonstrating sentience will be challenging or impossible.

The Ethics Labyrinth of Mimetic AI Healthcare

James Brusseau, Philosophy, Pace University, New York City
& Department of Information Engineering and
Computer Science, University of Trento, Italy

Informa AI Summit NYC 2022

jamesbrusseau.net