

# How **Artificial Intelligence** May One Day Threaten the **Political Capacity** of **Human Intelligence**

Benjamin Gregg

University of Texas at Austin

Digital Humanism Initiative

Technische Universität Wien

29 March 2022

# Thesis

# AI: Capacity to Make Daily Life Easier

- AI: reduce some of complexities that make aspects of modern life difficult
  - and vulnerable to failure
- Yet political communities endanger themselves if they outsource, to AI, the tasks and challenges of politics

# HI: Capacity for Mutual Attribution of Responsibility

- Unlike AI,

**HI** [= *human intelligence*]

- has a *peculiarly human* capacity for mutual attribution of responsibility
- Necessary for liberal democratic politics

# New Perspective on an Old Danger

- democratic politics requires an attitude of mutual responsibility-taking
- responsibility-taking is a capacity of human intelligence
  - but *not* a capacity of AI

# Part 1. HI: A Social Phenomenon

# The Brain and Its Environments

- What happens in the mind depends on aspects of the brain's environments
- Its most immediate and permanent environment: the body of the brain-bearer

# Mutual Influence

- Neurophysiological changes occur in brain as a person engages with her environments
- Brain's relationship with its environments is one of mutual influence
- That relationship of mutual influence sometimes has political dimensions



# HI Involves Social Relationships

- Certain features of HI derive from social relationships
- Individuals construct social environments on basis of affective and social responses that involve HI

# Features of HI that Mark Social Relationships

- each participant's conception of self
- multiple participants' recognition of each other's selfhood
- participants' investment of meaning in their own social and affective responses
- and in social and affective responses of others

# Embeddedness of Human Mind

- We are embedded in the physical world
- We are no less “embedded” in the mental world of others
- What we are currently doing and thinking is molded by whomever we are interacting with

# Part 2. AI Cannot Be Analogized to Human Intelligence

# HI: Evolved; AI: Human Artefact

- HI: product of the deep history of undirected, natural evolution
  - Develops from mix of biology, natural environments, cultural environments
- AI: emerged within a very brief time
  - Within the continuum of humans developing technologies
  - From a process highly directed, highly reflected

# Analogies & Disanalogies: HI & AI

- Both involve manipulation of information toward “insight” that enhances our ability to negotiate reality
- Still, many reasons *not* to view AI by analogy to HI

# 1<sup>st</sup> Reason Not to Analogize

- HI: capacity to analyze, accept or reject or modify internalized norms
- HI: self-conscious and self-reflexive
- HI: can reject its cultural programming
- AI: unlikely or unable to spontaneously override its own programming

# Moral Capacity of HI

- HI: can respond to rules from a normative or moral perspective or political standpoint
- Rule-breaking is sometimes politically warranted



## 2<sup>nd</sup> Reason Not to Analogize

- HI is embodied
- Range of different body states results from play of chemical & electrical signals in our brain-body
- Drives, appetites, motivations, predispositions, emotions, moods, phobias
- Interact with the things we see and hear and feel

## 3<sup>rd</sup> Reason Not to Analogize

- Humans have emotions
- A phenomenon of our biology that plays significant role in our behavior
- Including in political sphere

# Part 3. Consciousness is Co- Constitutive of Its Environments

# Consciousness and Interaction

- To have intelligence is to interact
  - with one's self, with others, with environments both natural and social
- Interaction involves consciousness
- Consciousness is something we achieve rather than something that happens inside us
- Consciousness involves the person's social context

# HI and Interaction

- HI does not begin and end with the brain
- We are what we do, where we are, how we interact with our environments
- We interact through tools, through language, through collective practices

# Intersubjective Invention of Politics

- In interacting with our social and political environments, we constitute ourselves as political beings with an intelligence capable of doing politics
- HI is not independent of its social and cultural contexts

# AI: Not Intersubjective, Not Political

- AI does not constitute itself
  - in its relationship with the data it processes
  - in its relation to human beings
- AI is not capable of intersubjectivity

# Self as Interaction with Other Selves

- Self-consciousness: self-identity
- Experiences have a “mine-ness” that makes them distinctively our own



# Personhood Through Intersubjectivity

- The human-self grasps itself vis-à-vis other selves
- But one instance of AI does not “grasp” itself
  - vis-à-vis another instance of AI
  - or vis-à-vis humans

# Part 4. Consciousness Involves Exchanges with Its Environments

# Consciousness in Exchange with its Environments

- Consciousness is co-constitutive of its environments
- Consciousness engages in *forms of exchange* with its environments

# Symbolic Exchange

- Both HI and AI engage in symbolic exchange
- But only HI has a capacity to engage in the *political* manipulation of symbols
- *Political*: normative contestation of competing moral values, commitments, and visions on basis of discursive arguments

# Symbolic Exchange

- Only human intelligence is capable of political act of moral debate, of debating competing value-commitments
- Consider values that influence the organization and perpetuation of political community
  - from parliamentary democracy to an authoritarian welfare-state  
from a unitary
  - from a theocratic republic to a monarchy

# Moral Exchange

- The neural networks of human intelligence are not symbolic machines
- The digital networks of AI are nothing but symbolic machines
- Unlike human intelligence, AI has no *moral* capacity

# Exchange as Distribution

- For both HI and AI, emergent patterns give the entire system capacities not available to components in isolation
- HI: units are individual persons
- AI: units are bits of information
- HI: emergent pattern is human integration
- AI: emergent pattern is integration of information

# Political Legitimacy through Exchange

In democratic political community, some power is distributed among citizens

This can legitimize power

In digital cultures,  
dynamic distribution has no legitimizing function



# Political Legitimacy v. Administration

To replace dynamic distribution in political community

with that of digital cultures

would displace political legitimacy

(achieved through human interaction)

with AI-directed administration of humans

# Part 5. HI: Political Capacity for Mutual Attribution of Responsibility

# Mental Model of Our Environments

- The human brain produces a mental model of our physical and social environments
- It checks this model against experience
- Our interaction with others is based on our respective mental models

# Empathy as Politically Relevant

- Empathy occurs when the brain activity of two or more humans closely mirrors one another's
- We humans empathize with each other by creating similar cognitive states

# By Creating Similar Cognitive States

- We intersubjectively co-constitute a shared concern with others
- A possible element in rational debates about political alternatives

# Capacity for Social Cooperation

- Development of a moral consciousness, such as empathy, is a social phenomenon
- To see intentionally acting beings in other human beings makes social cooperation possible

# Capacity for Mutual Attribution of Responsibility

In democratic community,  
citizens must be able  
to attribute responsibility  
to others  
for their convictions and actions

# Political Significance of 3<sup>rd</sup>-Person-Perspective

- Development of capacity for politics in democracy: based on complementary entanglement of participants' perspectives
- By means of intersubjective communication, humans construct a “third person perspective” by which participants can
  - judge themselves and others
  - verify agreement or disagreement with each other
  - empathize with each other



# Political Significance of Empathy

- Meaning of what happens in digital settings comes from non-digital settings of political community
- Example: balancing individual privacy (heightened by on-line anonymity) with security of private individuals on-line as well as public security

# Empathy to Address AI Challenges

- Empathy for others: one goal of such balancing
- Empathy (a phenomenon of HI) sometimes may address some challenges raised by some forms of AI

# Empathy as Political Consciousness

- One cannot be forced to be empathetic
- One can bear empathy toward others only voluntarily
- Perhaps our brains make us experience ourselves as free agents because we get an advantage from that experience
- That experience facilitates our capacity for individual and communal self-determination

# AI Cannot Substitute for Empathy

- AI cannot participate in the intersubjective mutual attribution of responsibility
  - that is necessary for politics in a liberal democratic community
- But communities often motivated to delegate, to AI, aspects of political participation and deliberation

# Part 6. Algorithms Cannot Meet the Challenges of Political Life

# AI Offers Superior Data Analysis

- AI facilitated social integration along non-political dimensions of public life in modern societies
- Can solve complex problems at limits of human decision-making capacities
- Boosts capacities for data manipulation toward greater insight and prediction

# In Healthcare, Fair Employment, Criminal Justice: AI Contributes

- By by treating responsibility along dimension of accountability for public policy on algorithmic approaches, AI may be able to
  - balance accountability, efficiency, fairness
  - support greater evidence-based decision-making
  - make statistical predictions & recommendations more robust and accurate than those made by humans

# But AI Cannot Guarantee Such Outcomes

Delegation of political and economic tasks  
and decisions to algorithms

only enhances their capacity

to include or exclude particular groups of people, and  
information, in diverse settings



# AI May Perpetuate Discrimination

- AI may create new forms of injustice by reproducing prejudices of prior participants
- or by replicating persistent social biases
- Only humans can evaluate the outcomes of algorithmic processes to identify possible harms created

# AI Cannot Take Responsibility

- Mutual attribution of responsibility by members of a political community: one basis for demanding accountability
  - By requiring decision-makers to present themselves before those persons whose interests they either represent or otherwise affect
- Responsibility-taking is a human activity
  - not a capacity of AI

# Conclusion

# Danger of AI

- AI does not pose a political danger inherently or necessarily
- Dangers posed by poor human decisions about deploying AI

# Example of Such a Danger

Conviction that political community  
can usefully outsource political tasks  
to AI

toward

greater efficiency

greater fairness

a more rational administration of society

# Political Responsibility

- To outsource the tasks of political community to AI would undermine core features of liberal democratic community:
  - political deliberation
  - making decisions in the public sphere
  - taking responsibility for those decisions & actions
  - while other participants take responsibility for their decisions & actions

*Danke für Ihre  
Aufmerksamkeit!*

# CREATING HUMAN NATURE

The Political Challenges of  
Genetic Engineering

Benjamin Gregg

Cambridge University Press  
2022

