

# WHAT WE BELIEVE ABOUT THE MIND AND HOW WE BUILD A.I.

Robert Lee Harris  
President – Communications  
Advantage, Inc.

[robert.harris@comadvantage.net](mailto:robert.harris@comadvantage.net)

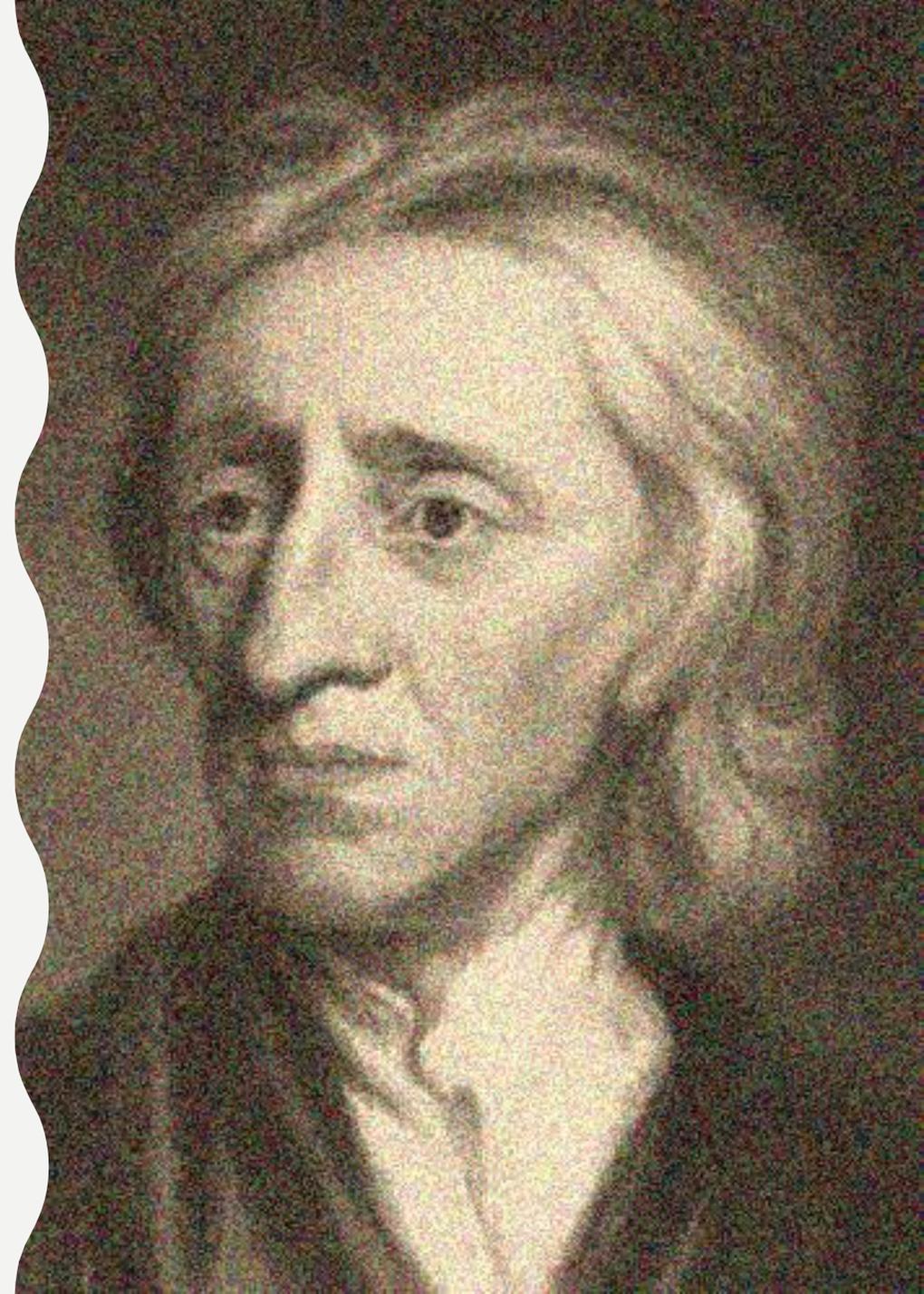


“I have always thought the actions of men the best interpreters of their

thoughts.”  
**JOHN LOCKE**

An Essay Concerning Human Understanding

- Are a technologist's thoughts about the mind reflected in their actions in developing AI systems?
- Are there times when it is better to ignore our understanding of consciousness (and just do what works)?



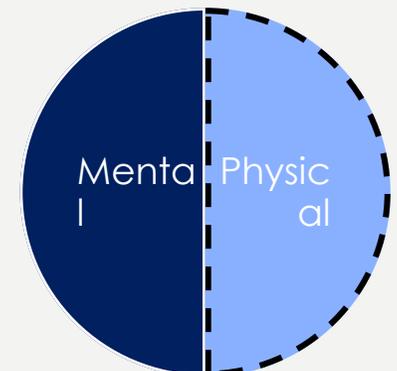
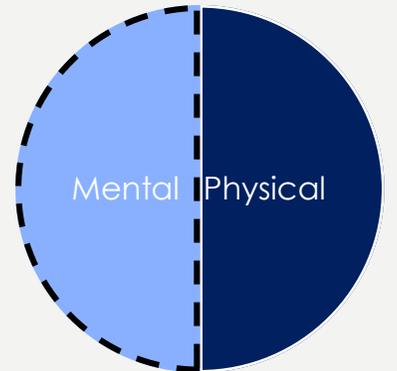
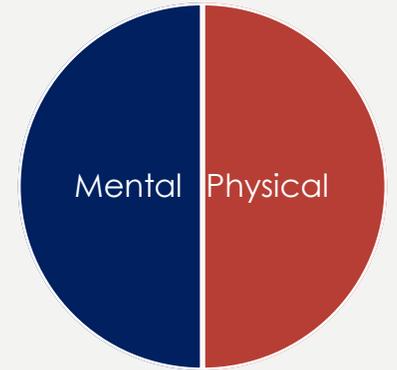
# WHAT WE BELIEVE ABOUT THE MIND

Dualism - Separate Mental and Physical substances (or properties)

Monism -

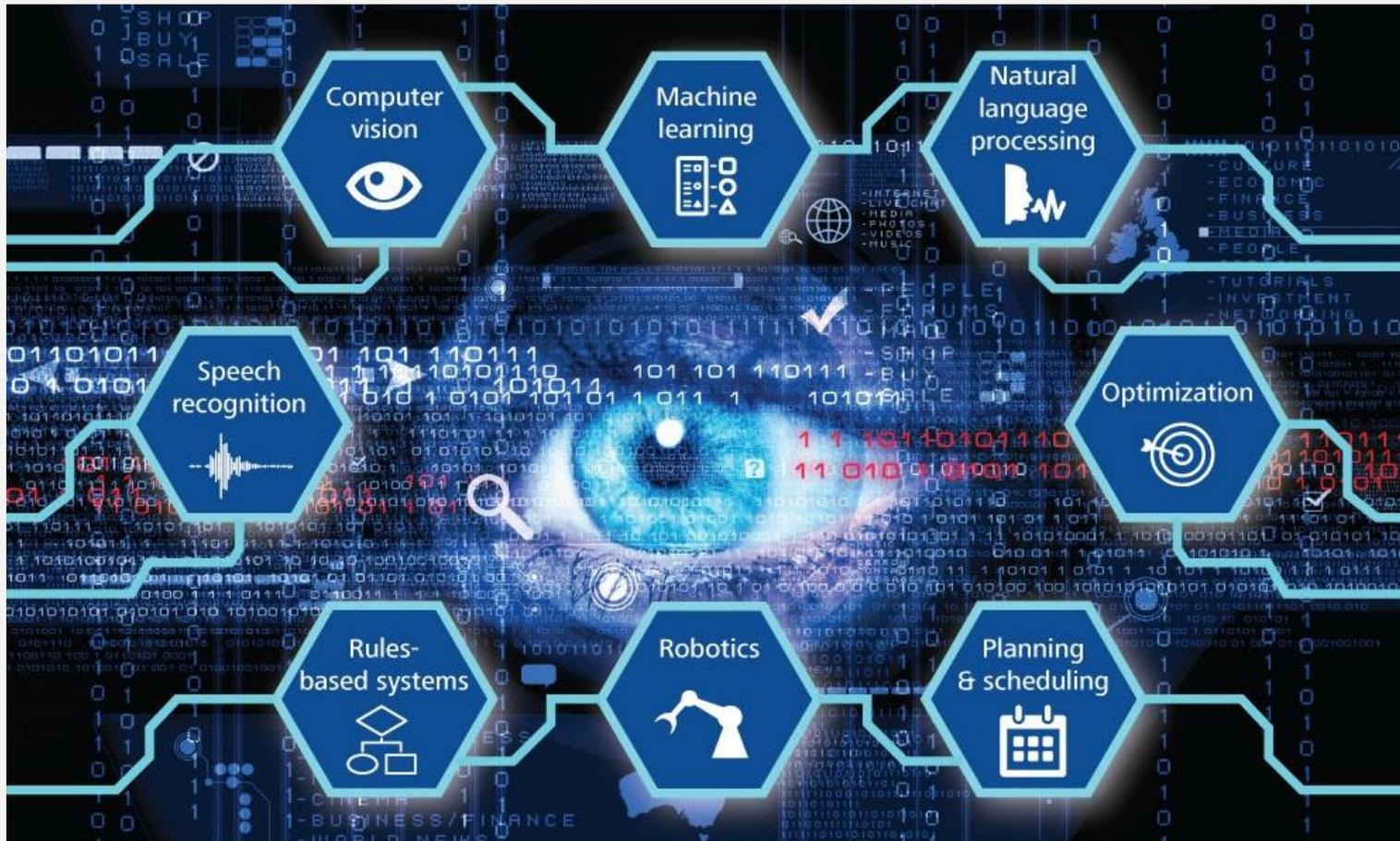
- Materialism – Physical is fundamental and mental is derivative
- Idealism – Mental is fundamental and physical is derived

- What is the presupposition for most work in AI?
- Can you believe in dualism and also believe in computer consciousness?



# NARROW AI AND AGI ARE BOTH MODELED AFTER HUMAN MINDS

- We have no other quantifiable objective
- Artificial Intelligence requires Intelligent Design
- We are creating AI in our own image



# TWO VIEWS OF TECH ENABLED IMMORTALITY:

## Transhumanism

**Assumes some form of Dualism – Mind or identity is transferable from body**

Technology will allow mind preservation

Step by Step process:

- “Cyborg” stage,
- “Transferred mind” or “Whole Brain Emulation” stage

## Cryonics

**Assumes Monism - The mind is not going anywhere**

- Cryopreservation currently works for blood, embryos, tissue, etc.
- Body is preserved before damage to biological neural network
- Assumes no identity outside of pure biology
- Mind or Identity is restored when knowledge or technology allows restoration

There is a branch of cryonics that also hopes for whole brain emulation



# A GENUINE MATERIALIST APPROACH TO AI:

## Unsupervised Learning



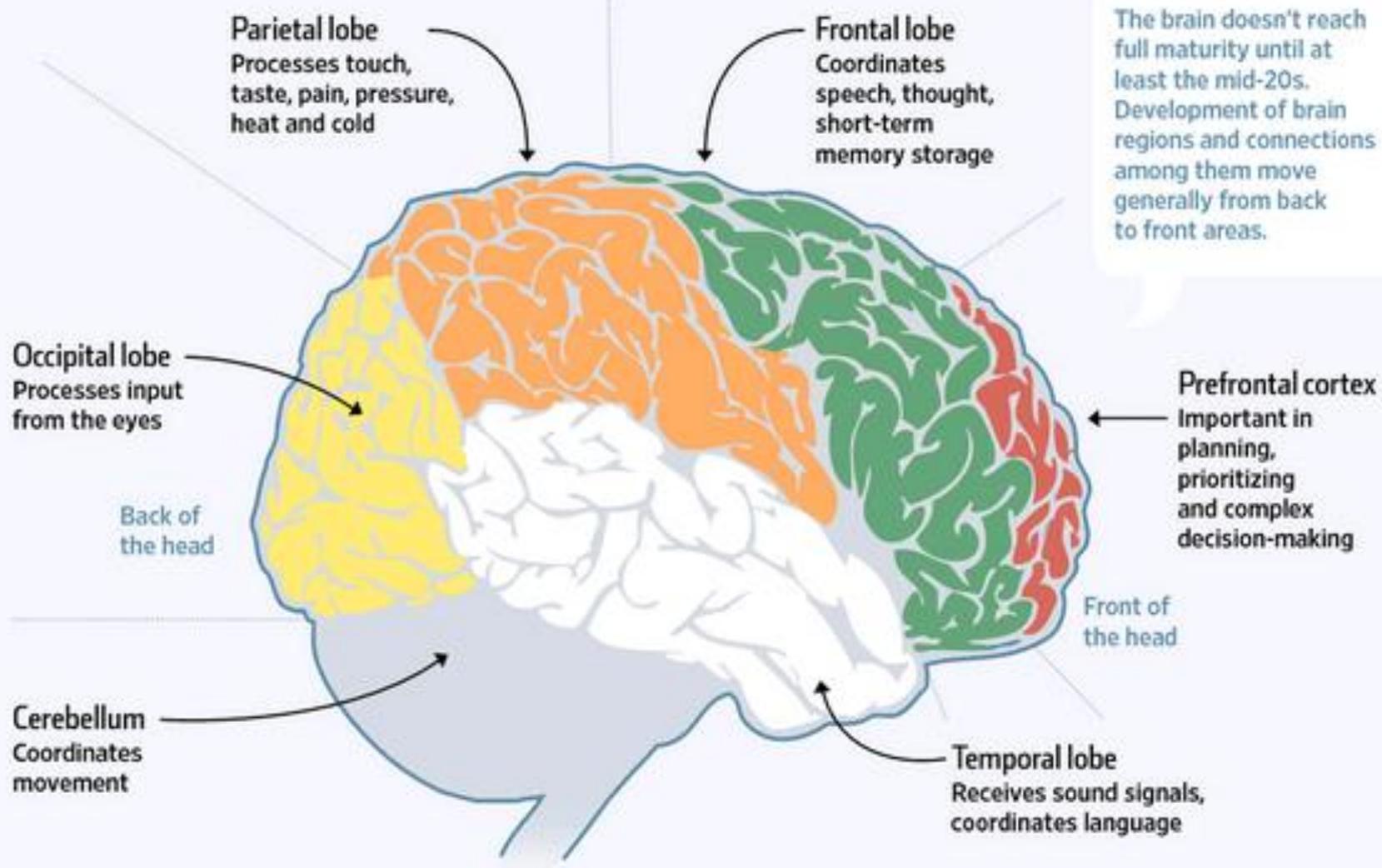
Terrance Tao – IQ 225-230



Marilyn Vos Savant – IQ 228



# HOW HUMANS LEARN:



- **Sensuous** (through the senses)
- **Intuitive** (which means to know or sense as if by instinct without conscious reasoning)
- **Reflective** (serious, thoughtful thinking)

1. Accessed by our senses and directed to our brain,
2. Organized (processed/integrated) and moved into memory
3. Stored, accurately and orderly,
4. Retrieved, so it can be used for a specific purposes

# FEAR OF ARTIFICIAL INTELLIGENCE

“We don’t know how to control superintelligent machines.” **Stephen Hawking**

“They’ll become self-protective and seek resources to better achieve their goals. They’ll fight us to survive, and they won’t want to be turned off.” **A.I. scientist Steve Omohundro**

It seems increasingly likely that we will one day build machines that possess superhuman intelligence. We need only continue to produce better computers—which we will, unless we destroy ourselves or meet our end some other way.

**We already know that it is possible for mere matter to acquire “general intelligence”**—the ability to learn new concepts and employ them in unfamiliar contexts—because the 1,200 cc of salty porridge inside our heads has managed it. There is no reason to believe that a suitably advanced digital computer couldn’t do the same.

**Sam Harris**



# REVERENCE OR WORSHIP OF AI

“We seem to be in the process of building a God. Now would be a good time to wonder whether it will (or even can) be a good one.”

**Sam Harris**

The new religion of artificial intelligence is called Way of the Future focusing on “the realization, acceptance, and worship of a Godhead based on Artificial Intelligence (AI) developed through computer hardware and software.” That includes funding research to help create the divine AI itself.

“What is going to be created will effectively be a god, it’s not a god in the sense that it makes lightning or causes hurricanes. But if there is

something a billion times smarter than the smartest human, what



Can We Know When Machines Achieve Consciousness?  
Anthony Levandowski

But who’s to say machines don’t *already* have minds? What if they take unexpected forms, such as networks that have achieved a group-level consciousness? What if artificial intelligence is so unfamiliar that we have a hard time recognising it? Could our machines have become self-aware without our even knowing it?



# SUMMARY:

