

# 6 IMPORTANT QUESTIONS

about

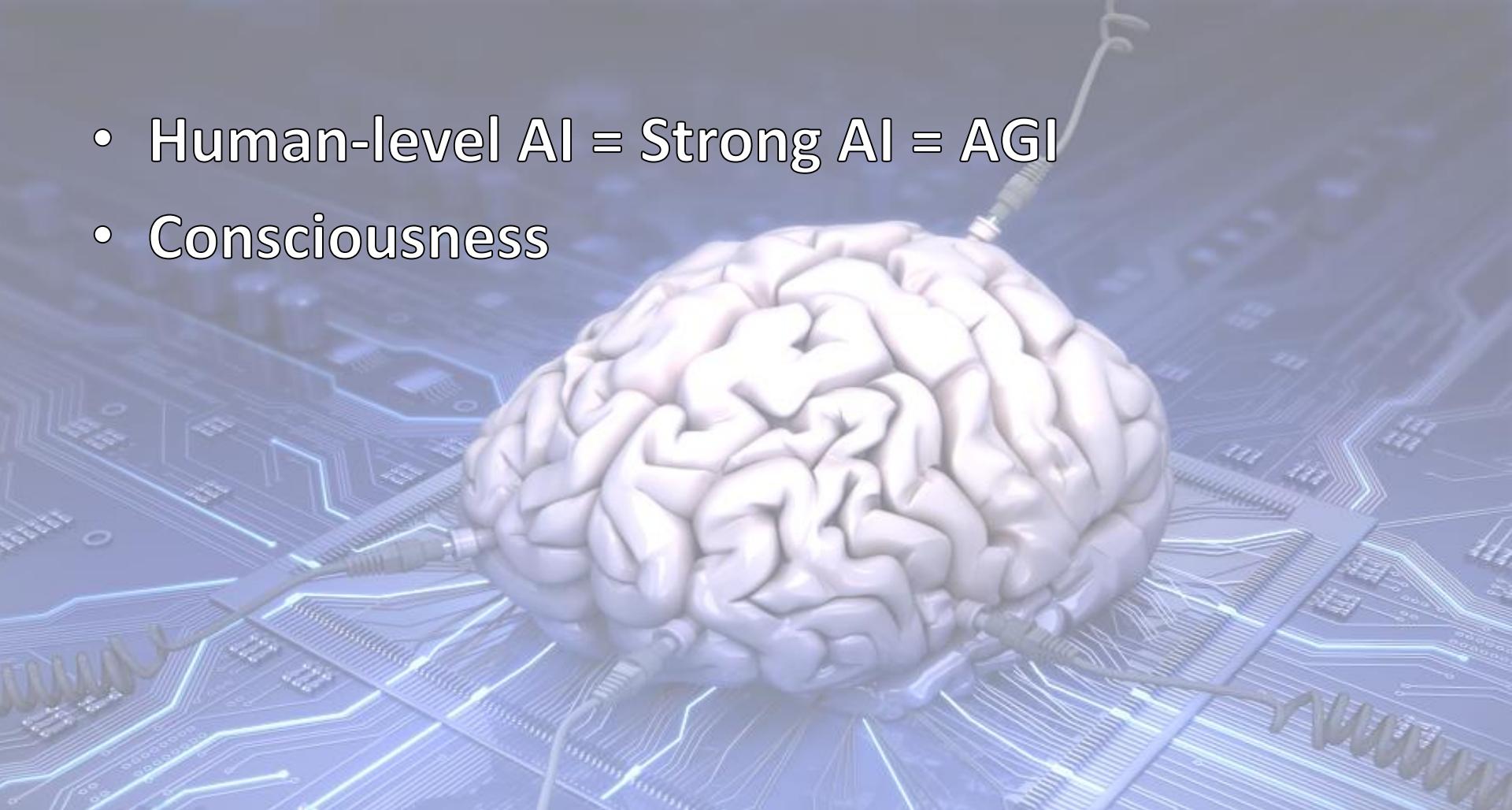
*artificial  
intelligence*

# 6 Questions

1. Can we make an AGI?
2. If so, when?
3. Will AGI lead to ASI?
4. Will an ASI harm us?
5. Can we upload our minds into computers?
6. Can we de-risk ASI?

# Definitions

- Human-level AI = Strong AI = AGI
- Consciousness



# The TURING Test



# Definitions

- Super-intelligence (ASI)
- Intelligence explosion

• ~~Singularity~~



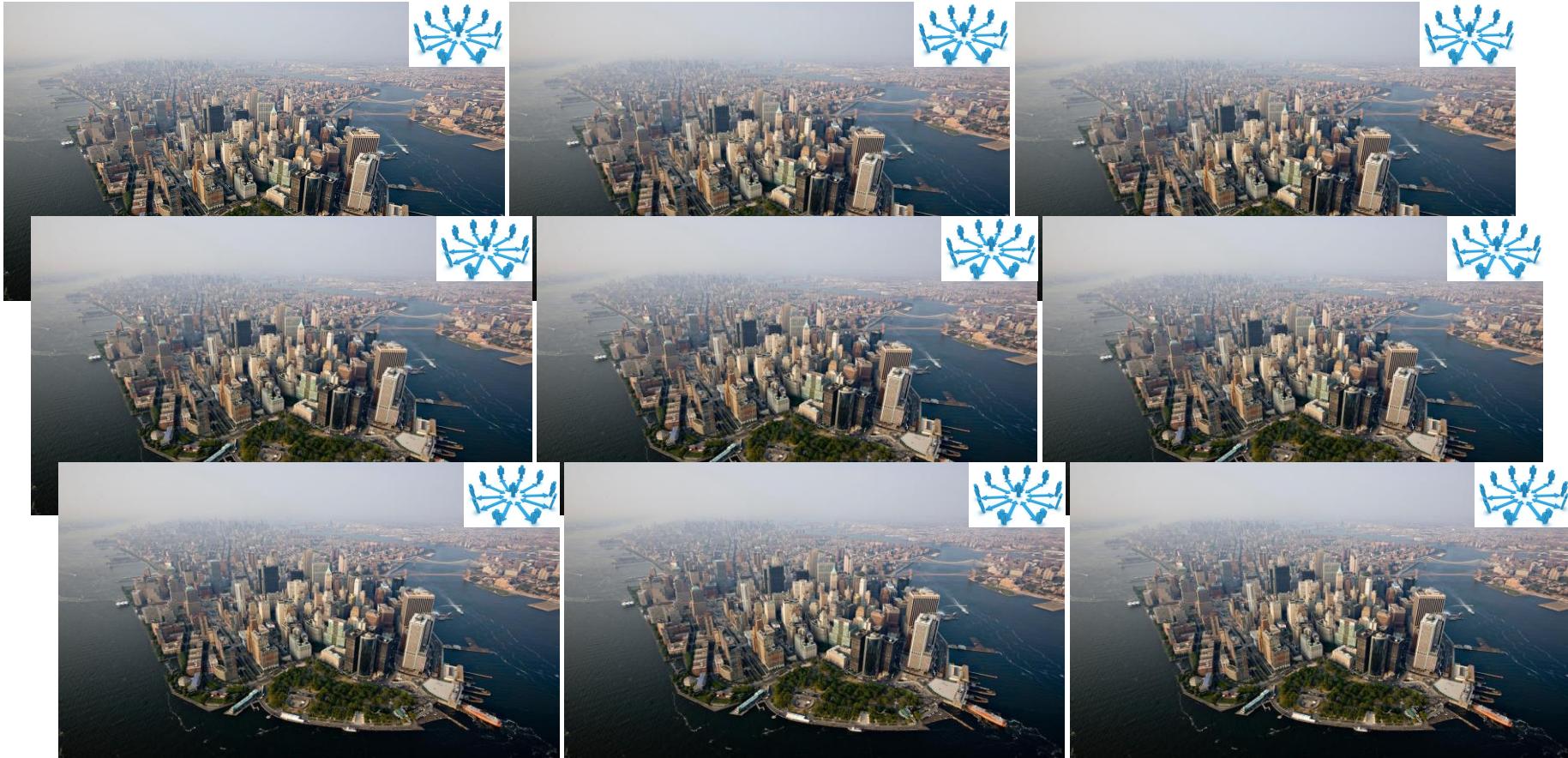
# 6 Questions

1. Can we make an AGI?
2. If so, when?
3. Will AGI lead to ASI?
4. Will an ASI harm us?
5. Can we upload our minds into computers?
6. Can we de-risk ASI?

# Q: How complex is a human brain?



# A: Very complex indeed!



**x 1,000**

# Hardware capacity is the easy bit



www:wwwAKE

# 3 ways to build a brain

Reverse  
Engineering

Incremental

Theory  
of Mind



# Why it might not work

THE MODEL  
MUST CAPTURE  
TIME-SERIES DATA

CONSCIOUSNESS  
IS A QUANTUM  
PHENOMENON

THE MODEL  
MUST CAPTURE  
SUB-ATOMIC DETAIL

WE HAVE SOULS



# A billion here, a billion there...

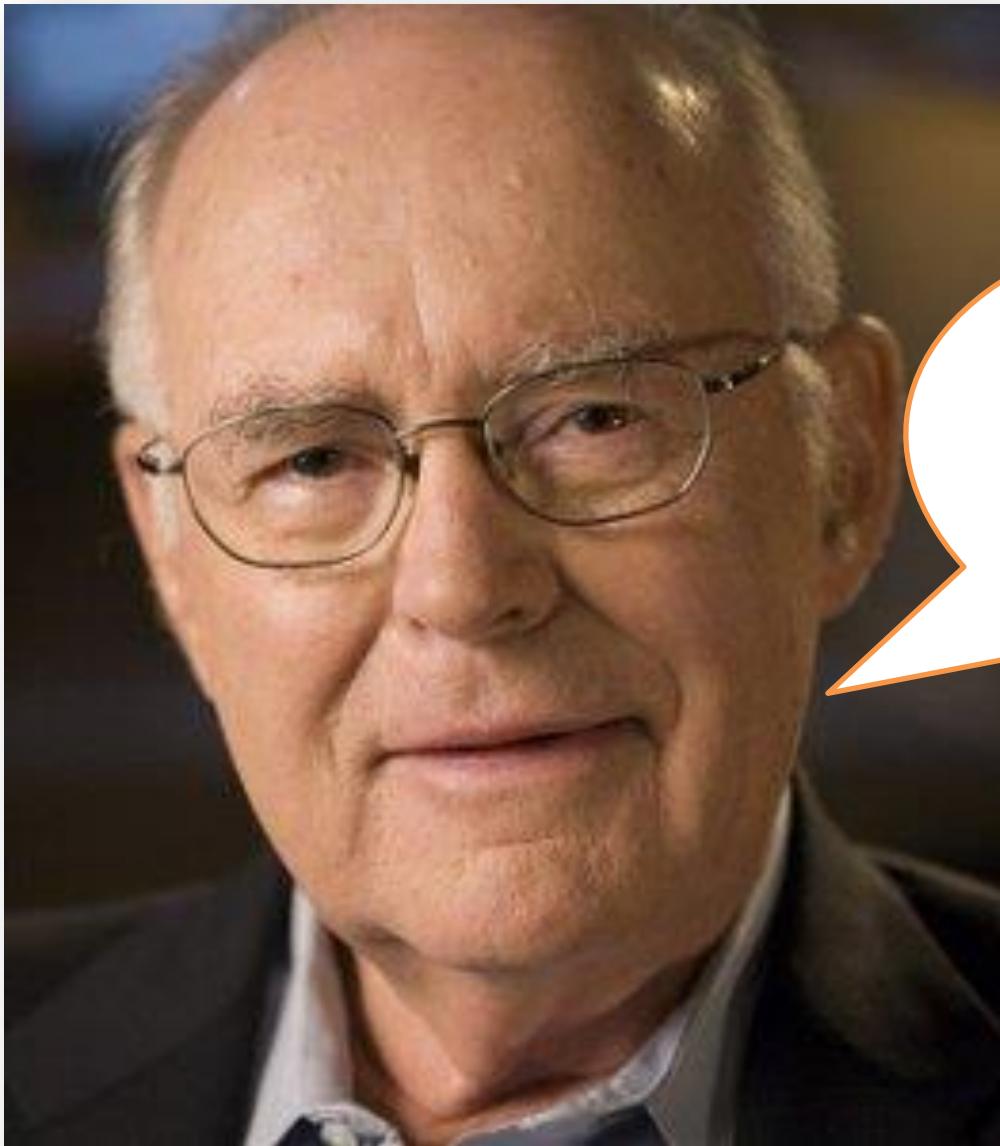


... pretty soon you're talking real  
money!

# 6 Questions

1. Can we make an AGI?
2. If so, when?
3. Will AGI lead to ASI?
4. Will an ASI harm us?
5. Can we upload our minds into computers?
6. Can we de-risk ASI?

# Moore's Law



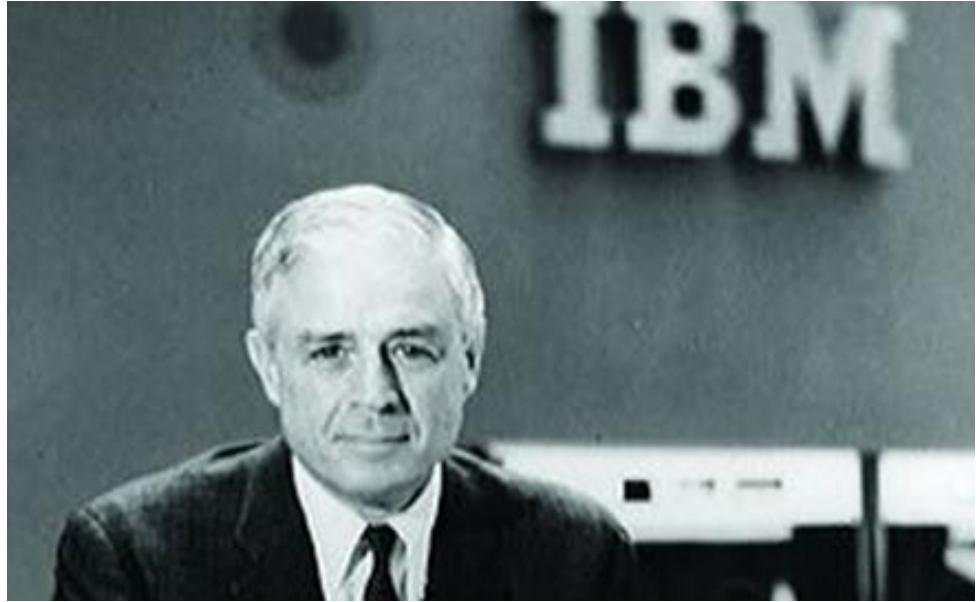
*I am **not** a  
Singularitarian!*

# Forecasting is difficult...



... especially about the future

# Plenty of scope for error



# Prospects for Moore's Law



**Exascale  
computing**

# Prospects for Moore's Law



Quantum  
computing



# 6 Questions

1. Can we make an AGI?
2. If so, when?
3. Will AGI lead to ASI?
4. Will an ASI harm us?
5. Can we upload our minds into computers?
6. Can we de-risk ASI?

# From AGI to ASI: fast or slow?



# 6 Questions

1. Can we make an AGI?
2. If so, when?
3. Will AGI lead to ASI?
4. Will an ASI harm us?
5. Can we upload our minds into computers?
6. Can we de-risk ASI?

# Scenarios if and when ASI arrives

We upload and become godlike

They help us

They leave

We become Pets

... Farm animals

... Zoo animals

... Slaves

We wilt

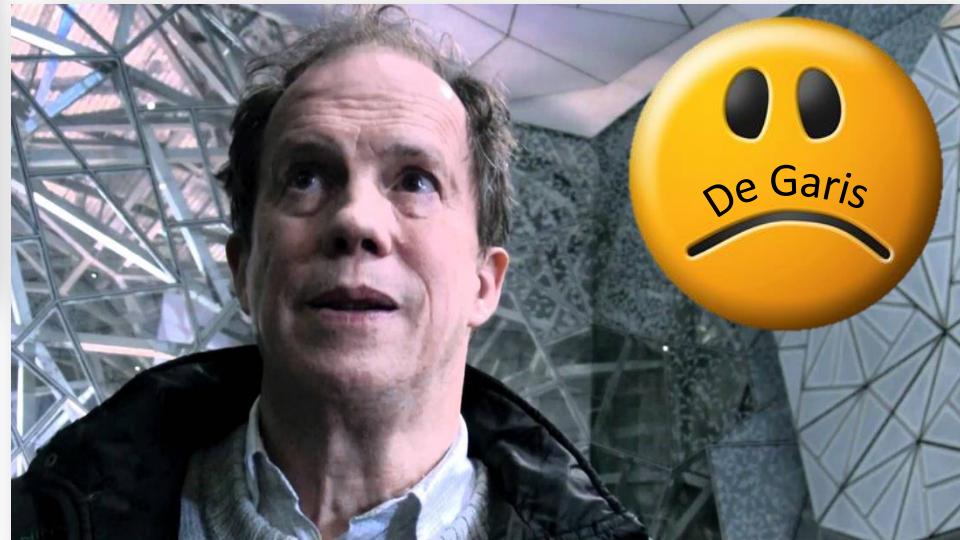
Humane extermination

Brutal extermination

Eternal torture



# Optimism and pessimism are just different forms of bias



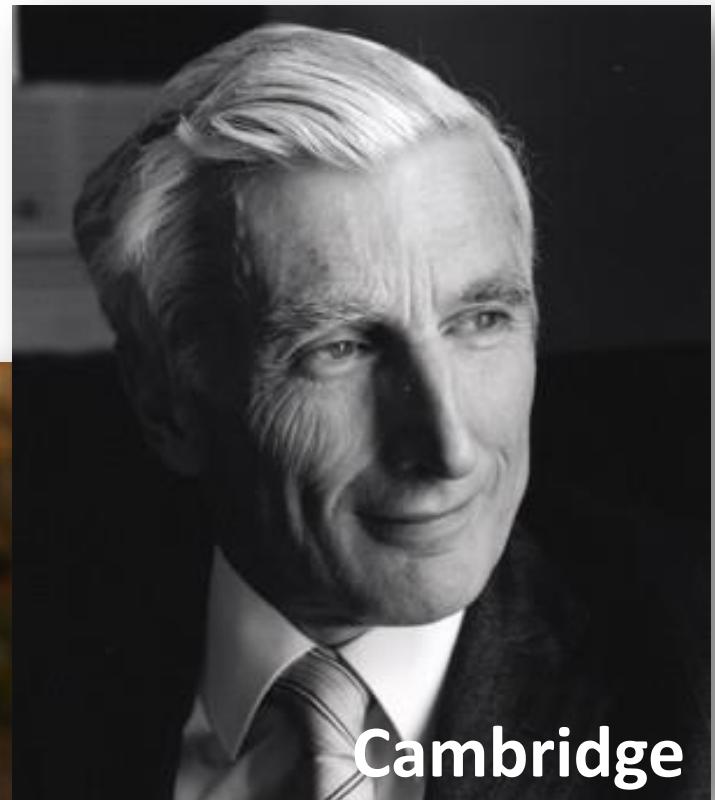
# Relinquishment is not an option



# Can we programme in friendliness?



Oxford



Cambridge

We can neither predict nor  
control an ASI's goals



# Maybe our best option is an Oracle AI

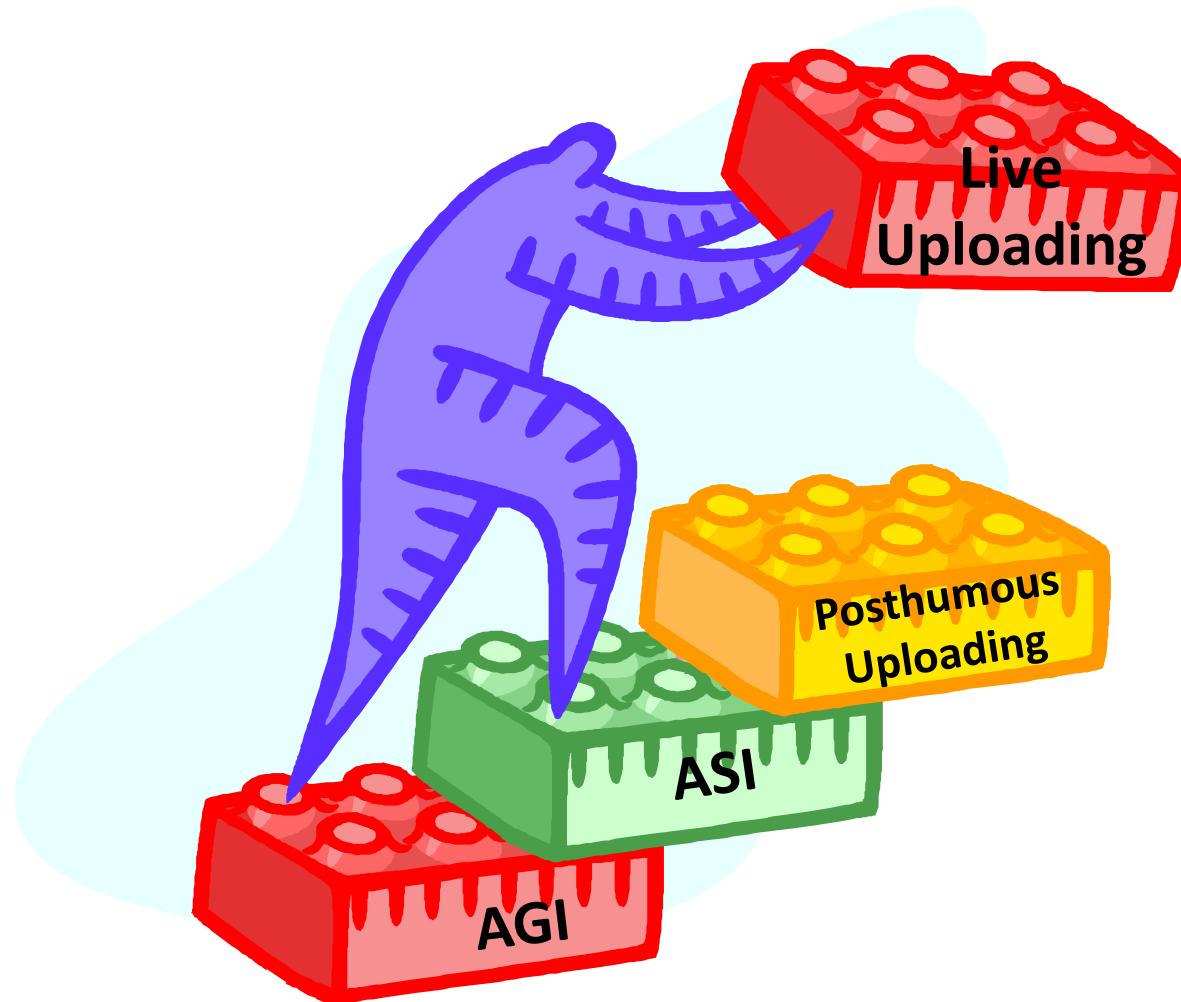


...which is  
very risky  
and may be  
impossible!

# 6 Questions

1. Can we make an AGI?
2. If so, when?
3. Will AGI lead to ASI?
4. Will an ASI harm us?
5. Can we upload our minds into computers?
6. Can we de-risk ASI?

# A very steep learning curve

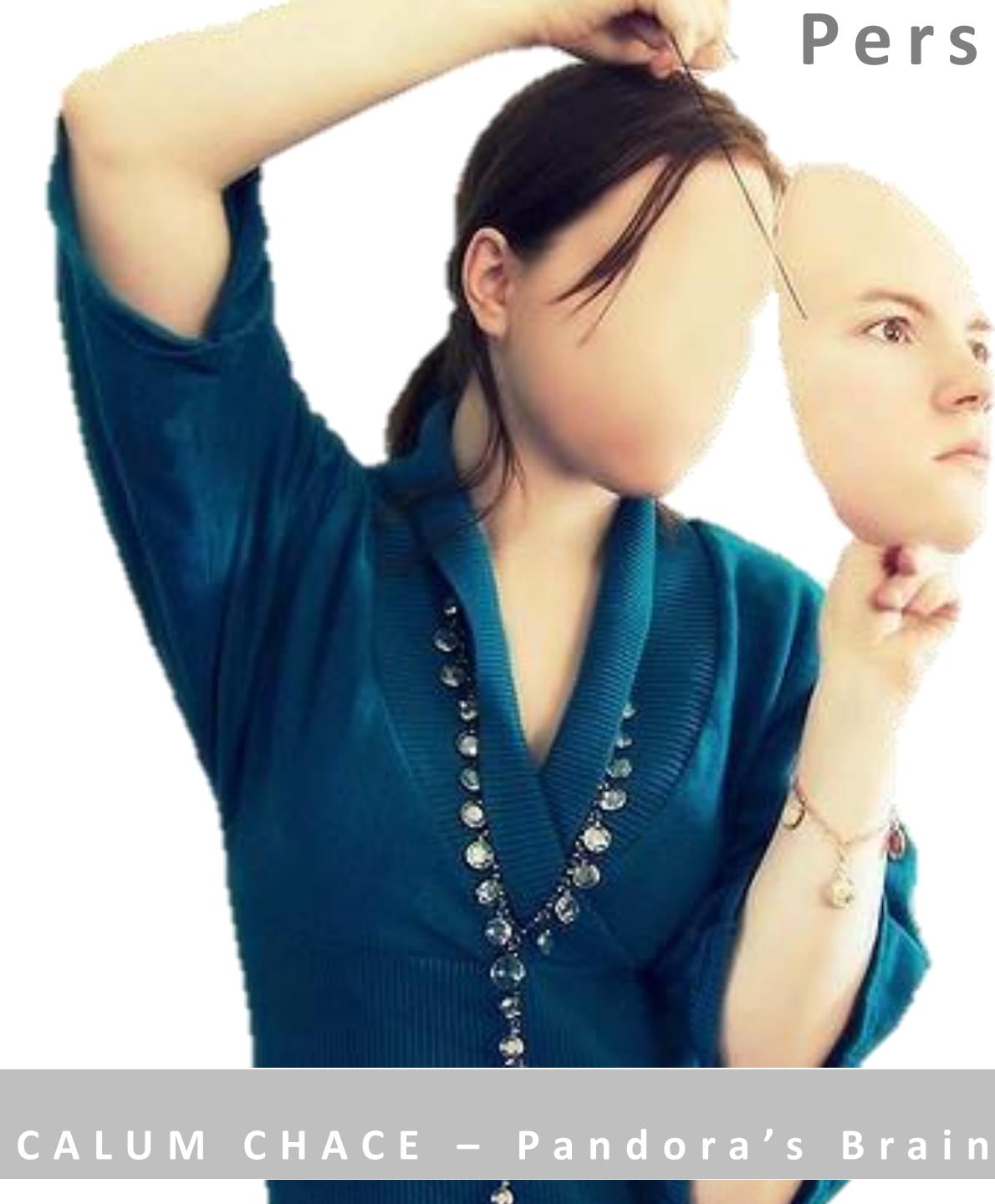


# Uploading or sideloading?



# The Ship of Theseus Resolution





Personal identity  
is going to  
get weird

# 6 Questions

1. Can we make an AGI?
2. If so, when?
3. Will AGI lead to ASI?
4. Will an ASI harm us?
5. Can we upload our minds into computers?
6. Can we de-risk ASI?

# FAILURE MODES



1. AGI not possible, or does not lead to ASI
2. ASI harms us
3. Uploading is impossible, or unaffordable

# Two new Apollo projects



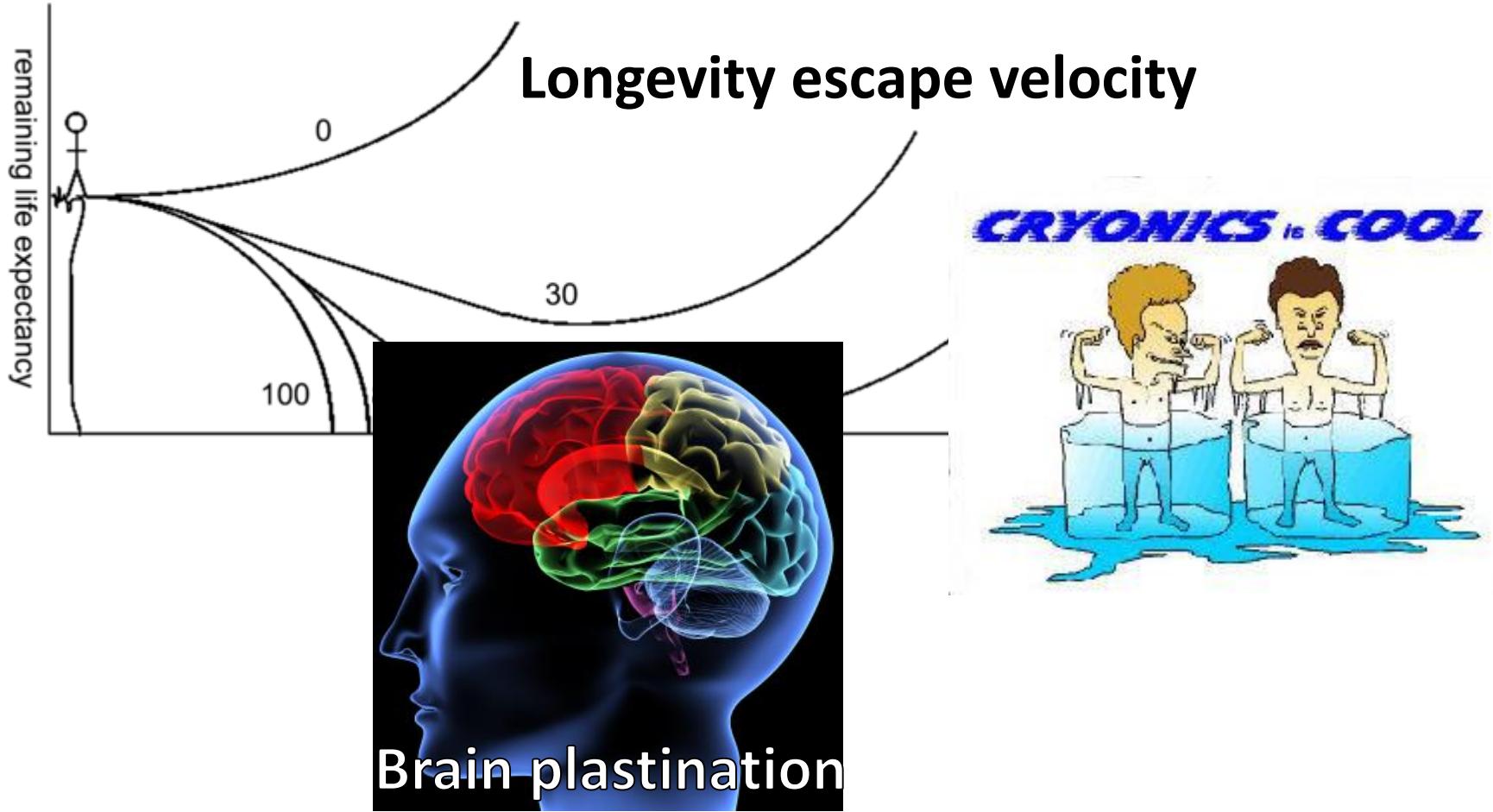
We are rushing  
towards AGI and ASI

We need to solve:

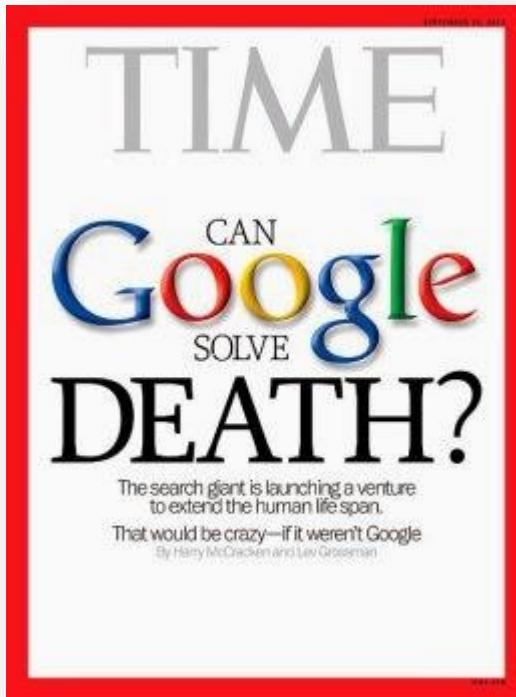
1. FAI

2. Death

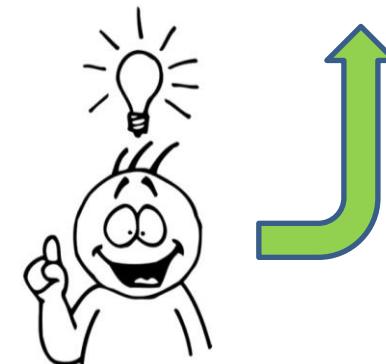
# Making death optional



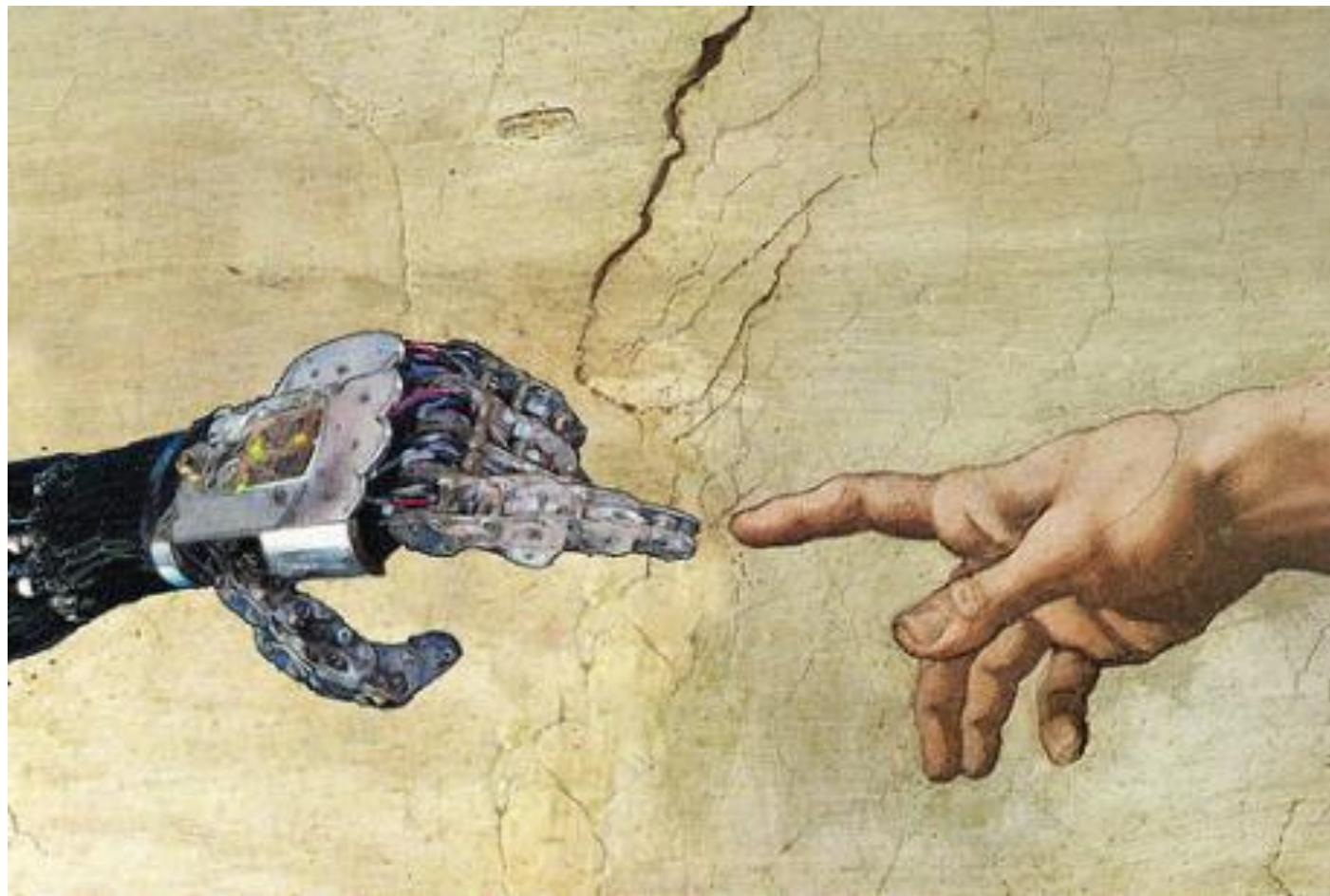
# Some of the funding will come from uber-geeks...



# Most of the funding must come from governments...



# It's not a Rapture for Nerds



# We need to get more people talking about ASI



# THANK YOU



CALUM CHACE – *Pandora's Brain*

March '14