

# “The AI Revolution”

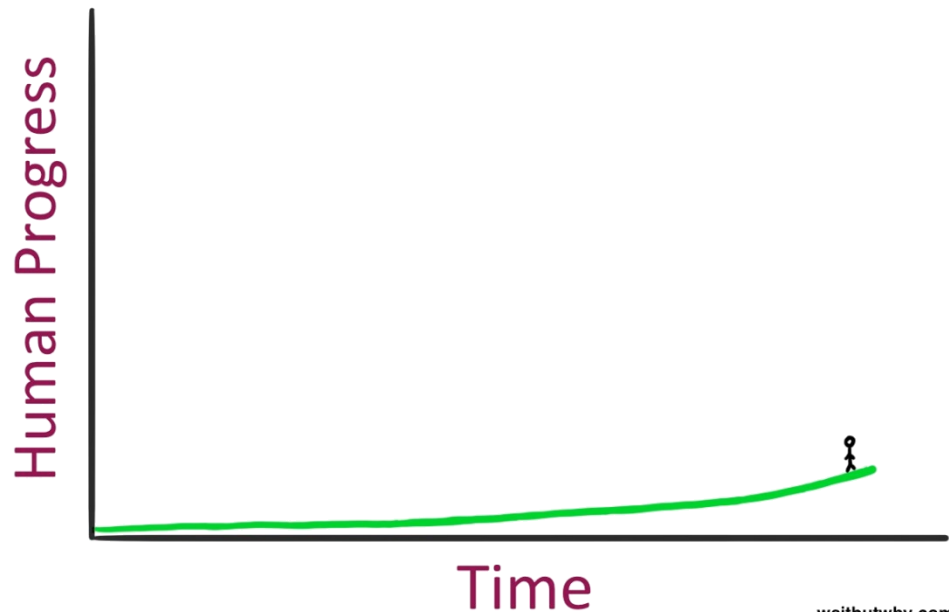
by Tim Urban, [WaitButWhy.com](https://waitbutwhy.com)

CNC seminar, FMFI UK

Juraj Holas

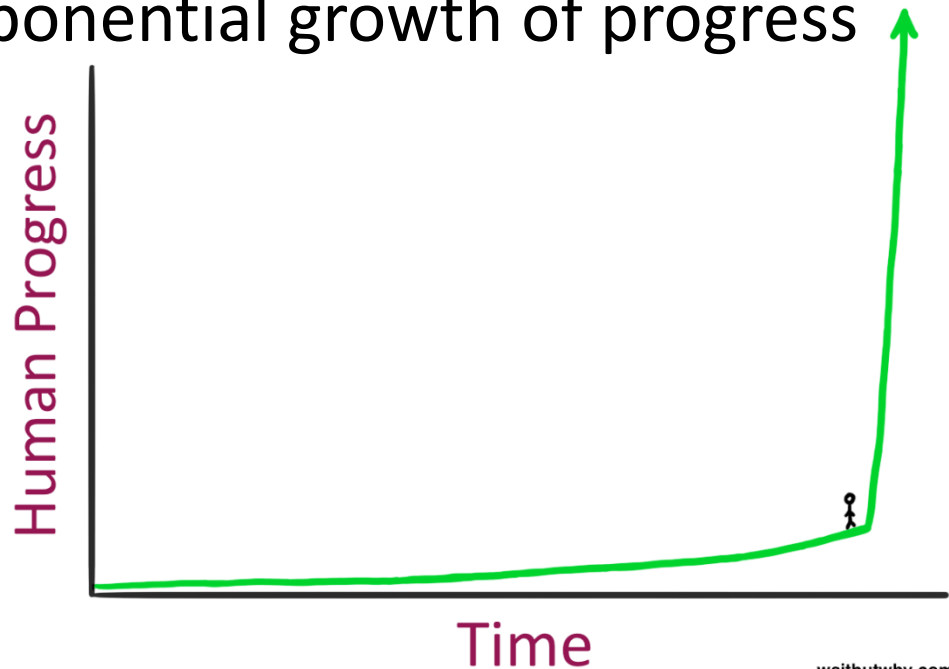
# Part 1: The Road to Superintelligence

- difference between 2000s-1750s vs. difference between 1750s-1500s
- “...more advanced societies have the ability to progress at a faster rate than less advanced societies...” => exponential growth of progress
- our perception:



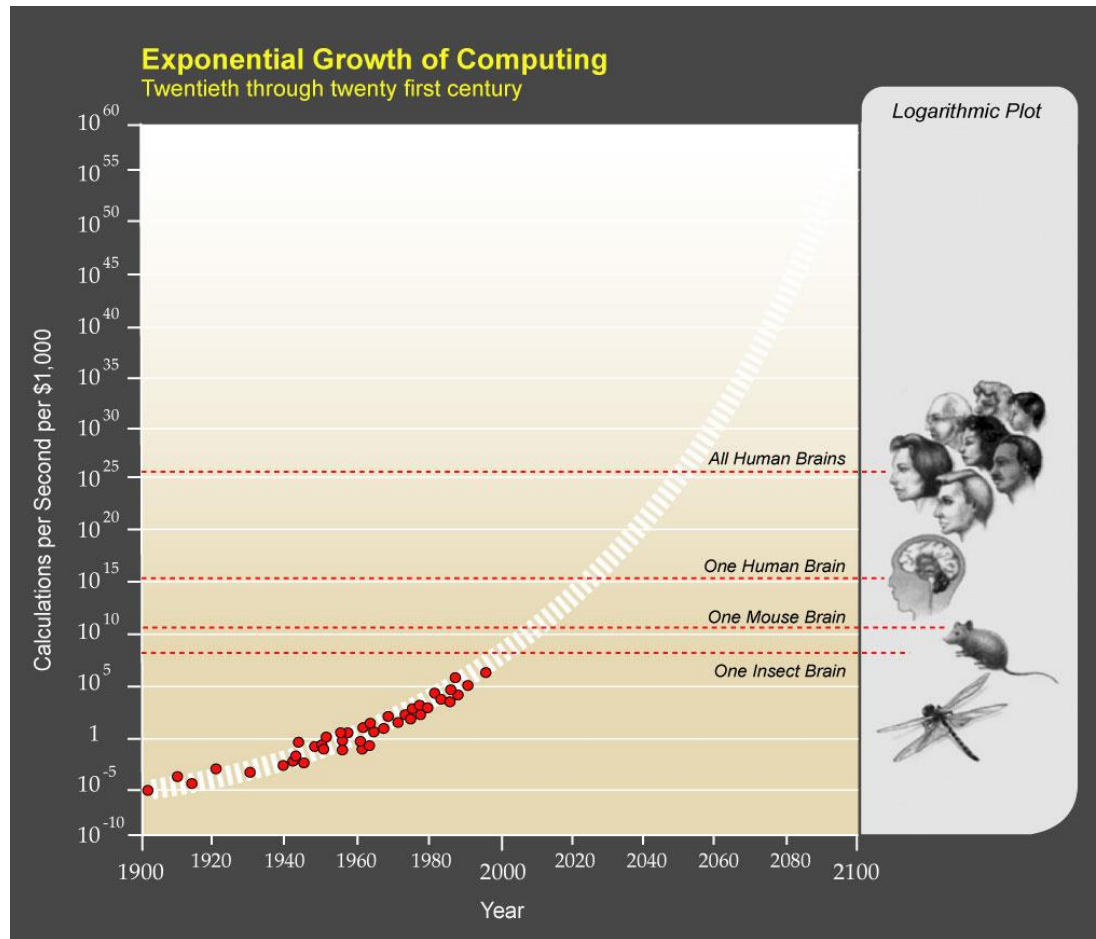
# 1. The Road to Superintelligence

- difference between 2000s-1750s vs. difference between 1750s-1500s
- “...more advanced societies have the ability to progress at a faster rate than less advanced societies...” => exponential growth of progress
- ~~our perception~~
- what reality might look like:



# Hardware

- prediction of computers` **raw** computational power:

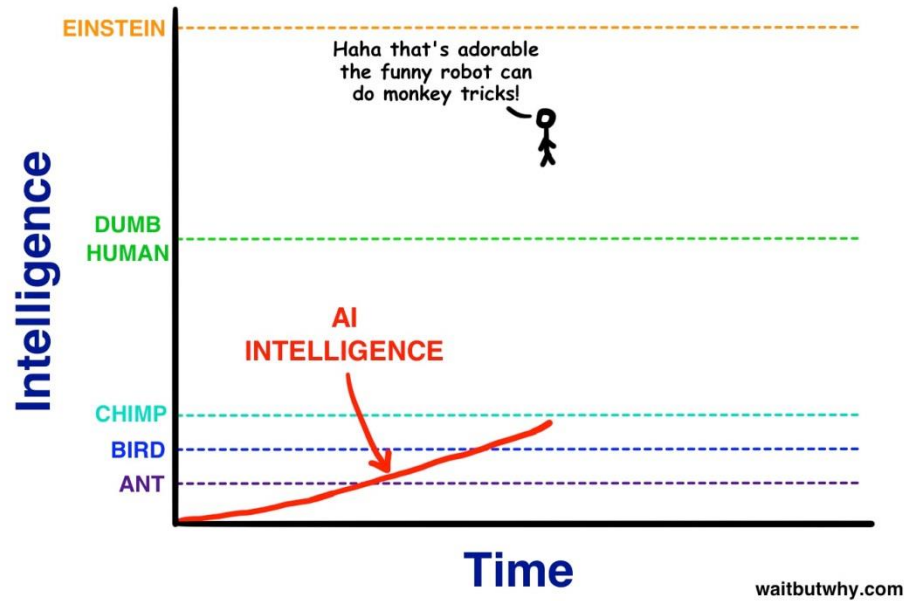


# Software

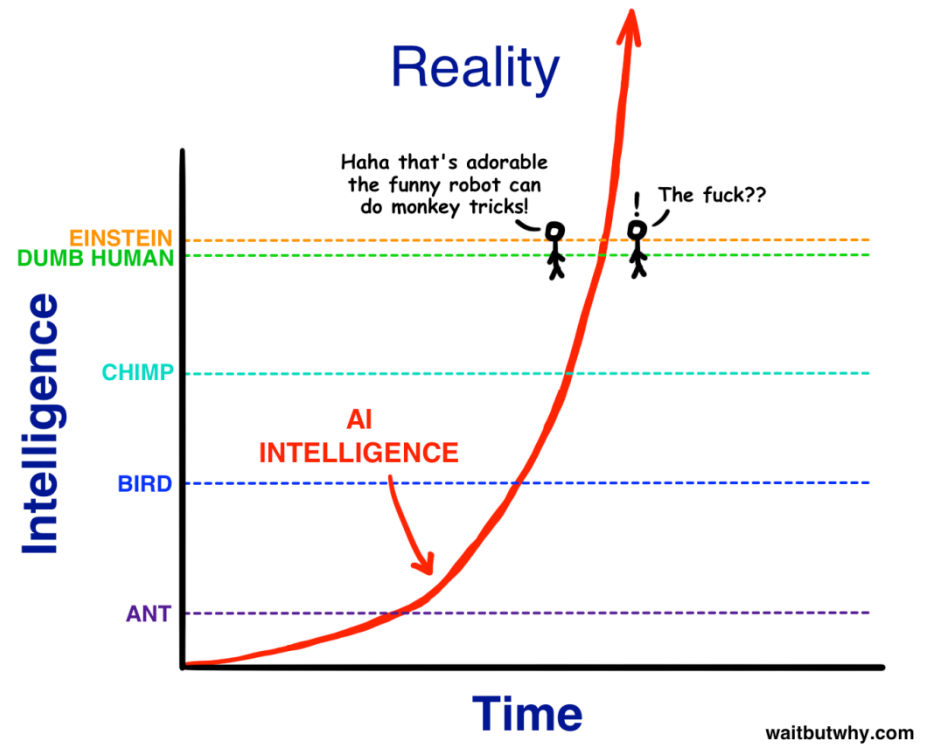
- Multiple options, optimal one to be found:
- Mimic brain – artificial neural networks
- Mimic evolution – genetic algorithms
- Build AI researching all AI studies, make it improve itself
- other approaches?

# Skewed perspective

Our Distorted View of Intelligence



Reality



# 3 tiers of AI

- Artificial Narrow Intelligence: weak AI, today`s applications
- Artificial General Intelligence: human-level AI across fields
- Artificial Superintelligence (ASI): “...is ***much*** smarter than the best human brains in practically every field”

# Experts` views

- “By what year would you see a (10% / 50% / 90%) probability for HLMI to exist?”

	median	mean
10%	2022	2036
50%	2040	2081
90%	2075	2183

- “How long after HLMI will there exist an AI superintelligence?”

	median	mean
≤ 2 yrs.	10%	19%
≤ 30 yrs.	75%	62%

- “How positive or negative would be overall impact on humanity, in the long run?”



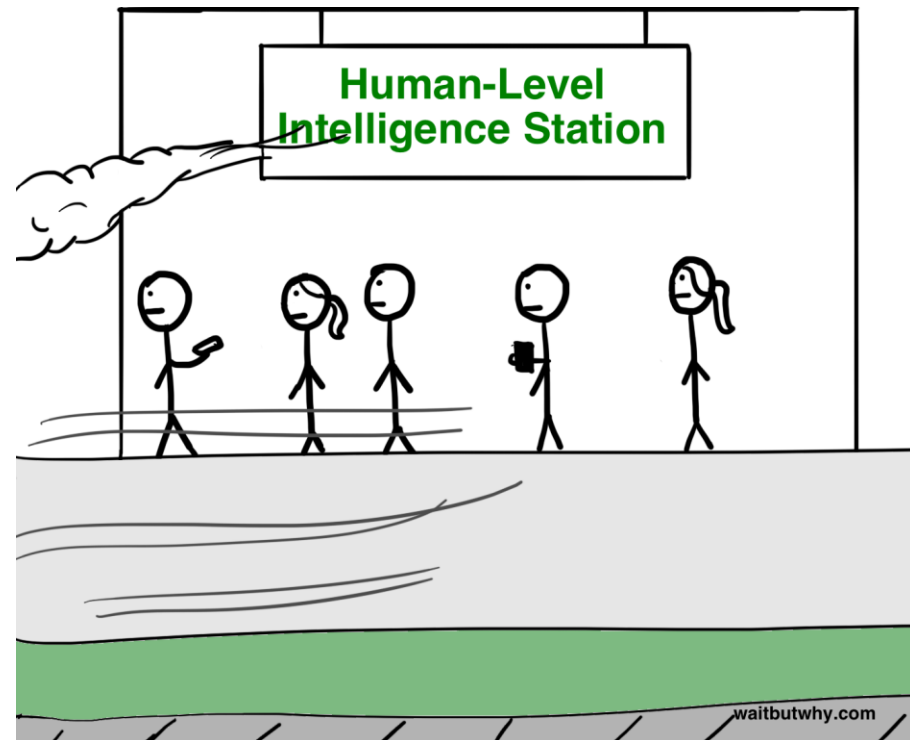
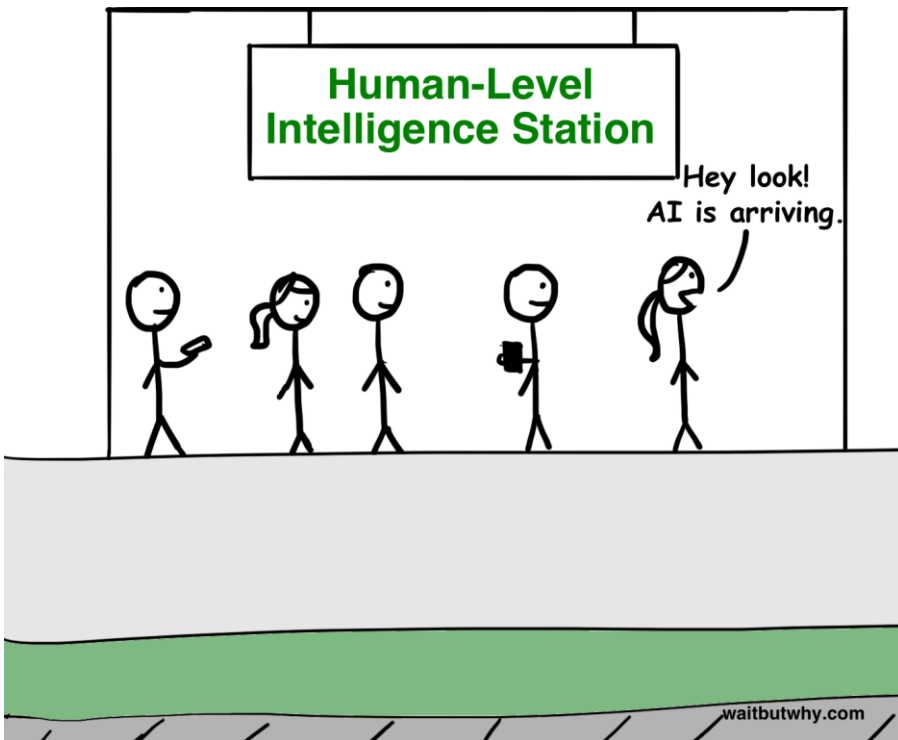
■ Extremely good ■ Good ■ Neutral ■ Bad ■ Extremely bad



# Part 2: God in a Box

*a.k.a. the weird stuff*

- Assuming we'll create superintelligent AI: what's next?
- Author suggests *"intelligence explosion"*



# Speed vs. quality difference

- AI superintelligent by:
  - Speed: thinks *faster* than human can
  - Quality: thinks of more *concepts* than human can –  
e.g. you can't explain relativity to an ant
- Quality-ASI to human difference can be 100x more than human to ant difference
- E.g. if ASI conquers nanotechnology -> can place every atom to desired position -> can do “god-like” actions

# Optimistic predictions

ASI in role:

1. **Oracle** – can answer any question (Google on steroids)  
*“How can I manufacture a more efficient car engine?”*
2. **Genie** – executes any high-level command  
*“Build a new and more efficient kind of car engine”*
3. **Sovereign** – is assigned a broad and open-ended pursuit and allowed to operate in the world freely, making its own decisions  
*“Invent a faster, cheaper, and safer way than cars for humans to privately transport themselves.”*

# More optimistic predictions

- ASI finds cure for mortality
  - Reverse aging by “repairing” old tissues
  - Argument of possibility:  
*“Every man so far has died. (2016)”*  
*“No man can fly in the air. (1800)”*
- Further thoughts: repairing tissues could be moved further to *improving* tissues, e.g. brain, thus making humans superintelligent as well

# More optimistic (?) predictions

- Tissue improvement -> replacement with better artificial alternatives -> ... -> making us entirely artificial
- Merging human species with AI
- Will it still be *us*?

*Note: such extensive predictions are heavily disputed among specialists.*

# Pessimistic predictions

- Non-existential risks:
  - Economy collapse
  - Population boom
  - ...
- Most are “to be solved by ASI”

# More pessimistic predictions

- *“...creating something smarter than you is a basic Darwinian error”*
- Existential risks:
  - Army/terrorists create ASI, loose control
  - ASI decides to kill us – Terminator, Matrix, ...
  - Extinction by ASI`s goal optimization

T. Urban (2015), The AI Revolution, WaitButWhy.com, [Part 1](#), [Part 2](#)  
Images by WaitButWhy.com, unless stated otherwise

