The Future of (Artificial) Intelligence

Fantasy Basel, 14 May 2015

gbs switzerland

Kaspar Etter
www.superintelligence.ch

Robin Li, Bill Gates, Elon Musk on AI

The Huge Challenge of Artificial Intelligence
www.youtube.com/watch?v=vHzJ_AJ34uQ

More Money for Entertainment

… than ensuring a good outcome!

Introduction

What are we talking about?

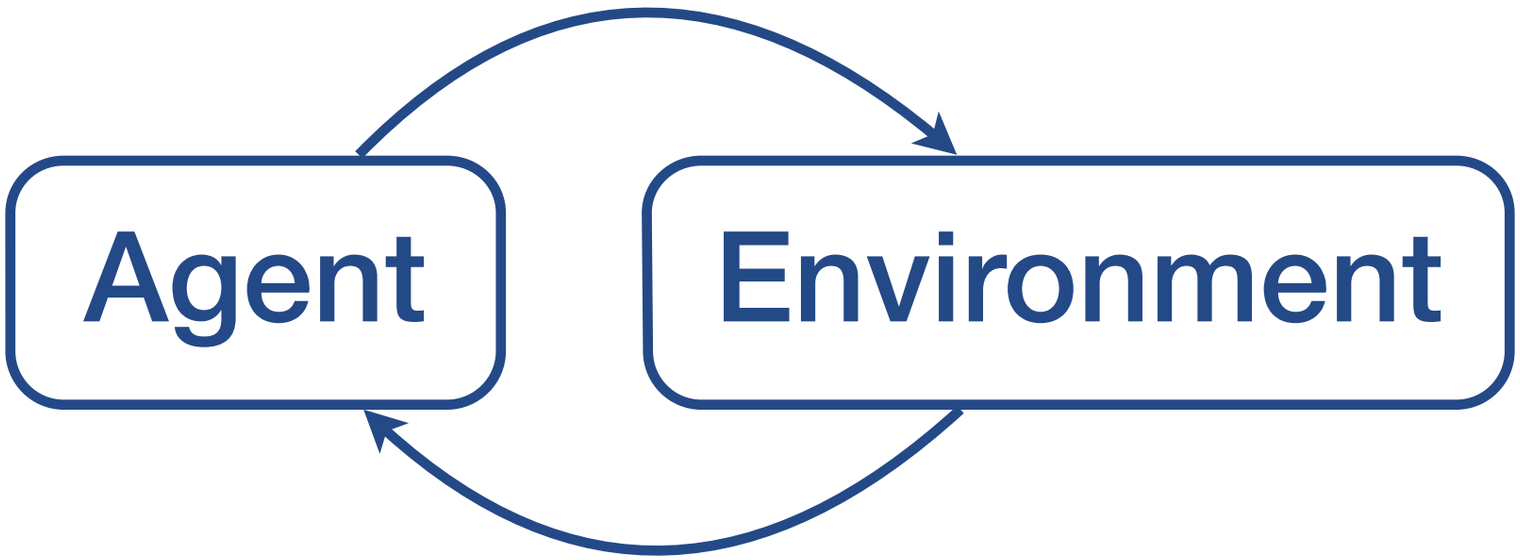
Intelligence

«Intelligence measures an agent's ability to achieve its goals in a wide range of unknown environments.»

$$\text{Intelligence} = \frac{\text{Optimization Power}}{\text{Used Resources}}$$

Universal Intelligence
arxiv.org/pdf/0712.3329.pdf

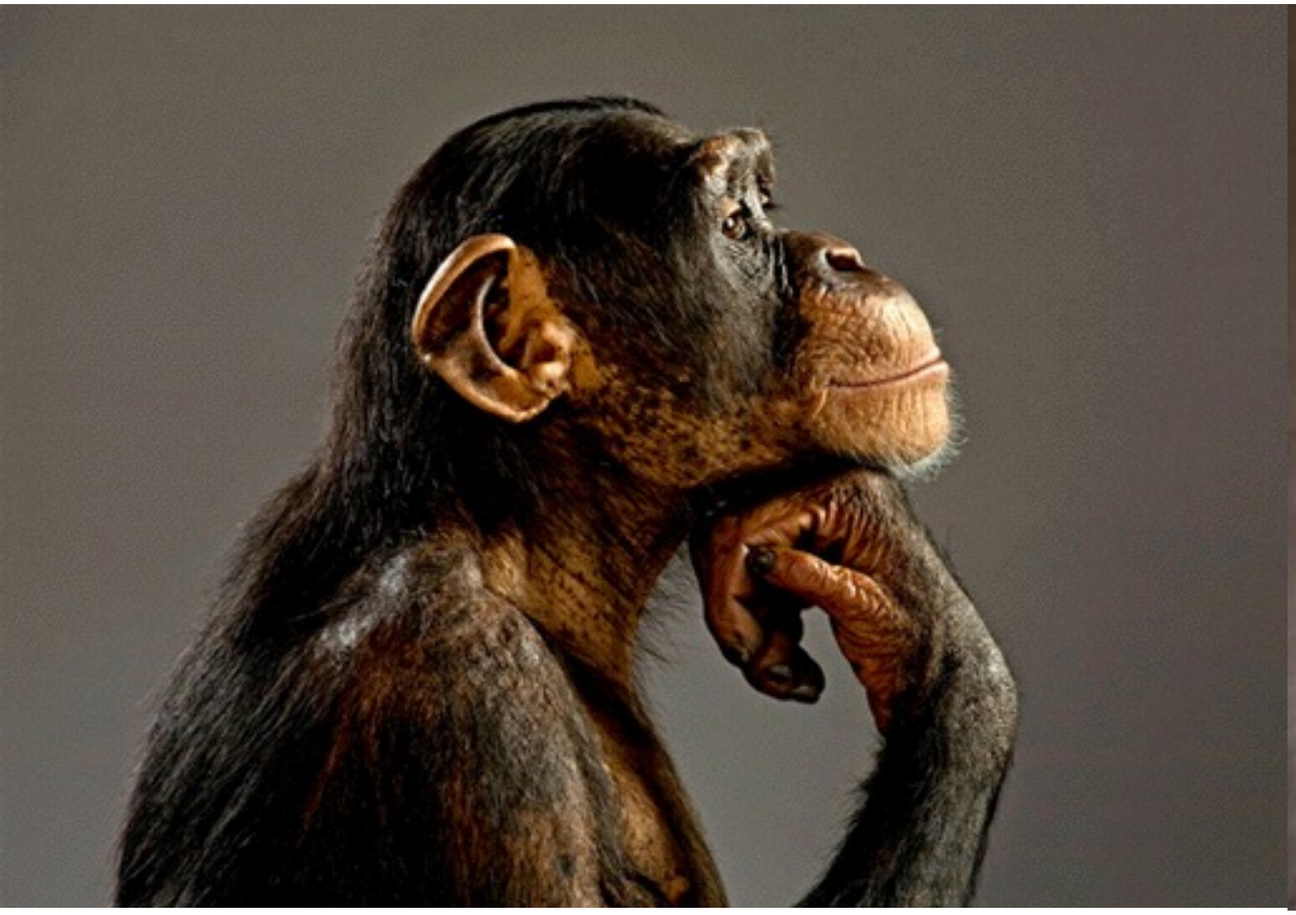
Ingredients



Learn, predict, rate and plan!

Marcus Hutter: Universal Artificial Intelligence
www.youtube.com/watch?v=I-vx5zbOOXI

Intelligence is a Big Deal

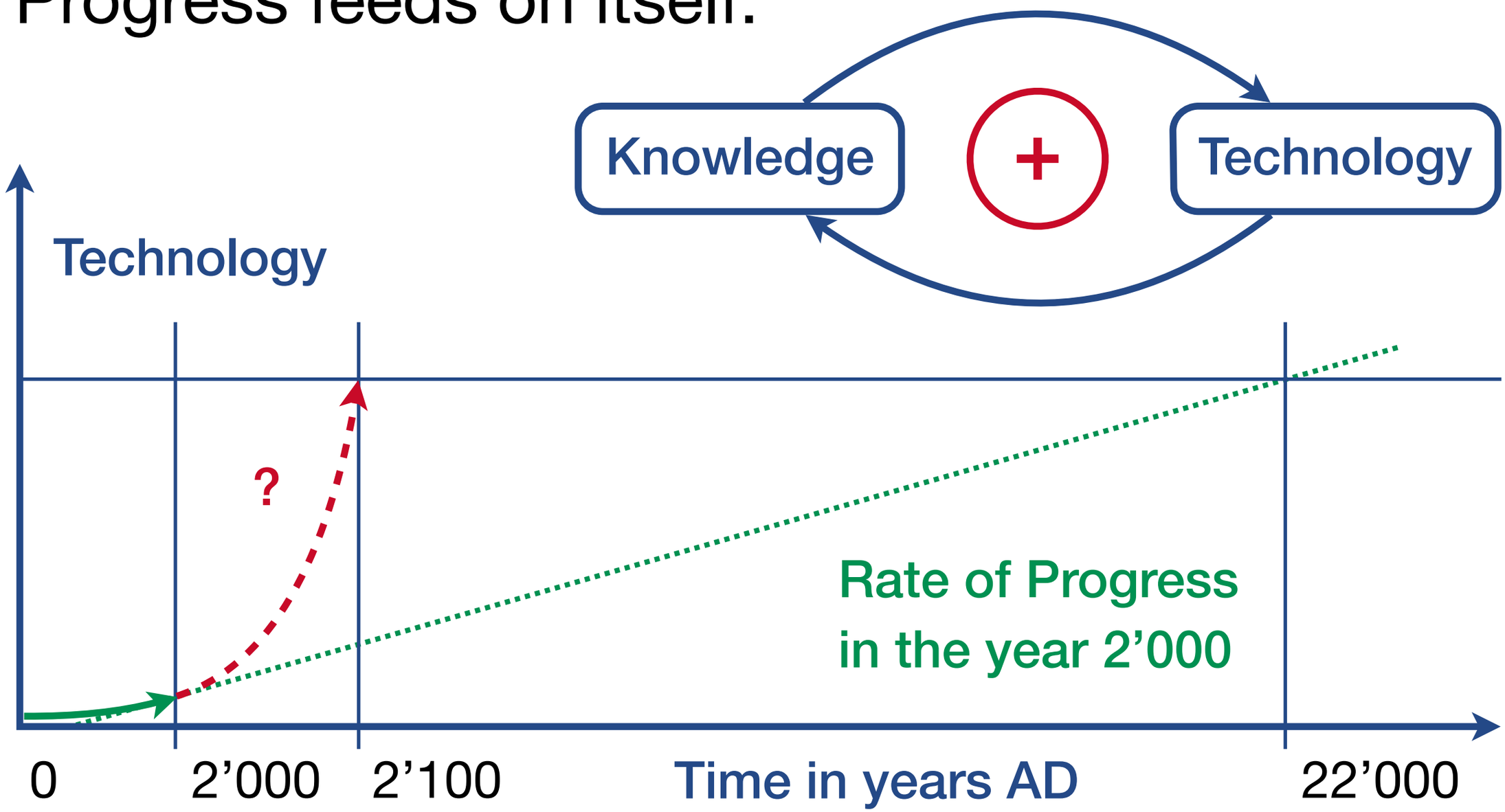


6 million years ago, 96% common DNA

Fast-Evolving Human DNA Leads to Bigger-Brained Mice
phenomena.nationalgeographic.com/2015/02/19/[…]

Accelerating Change

Progress feeds on itself:



The Law of Accelerating Returns
www.kurzweilai.net/the-law-of-accelerating-returns

Technology = (Neutral) Lever

Greed and laziness drive us to

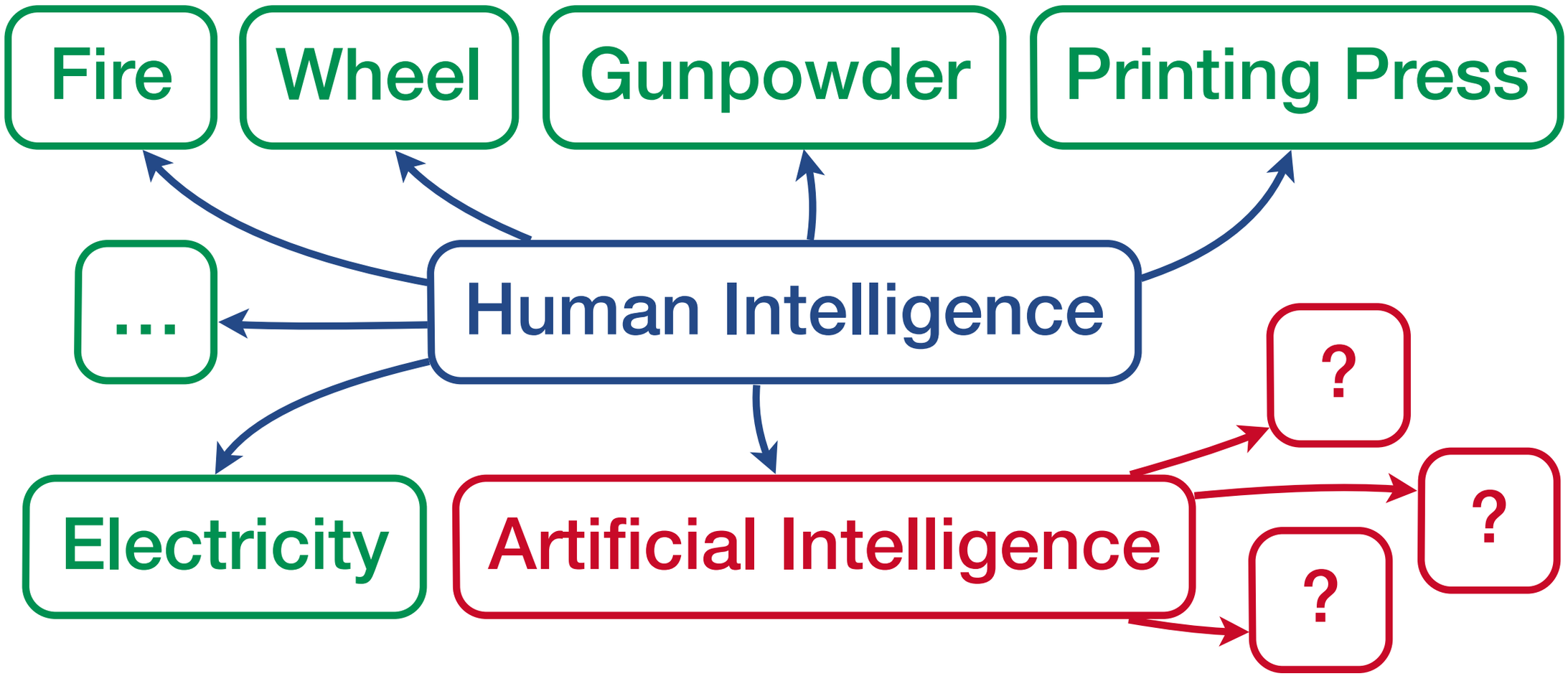
- increase our **productivity**
- make our **lives easier**

… which is fine except:

Our technological progress far outperforms our moral progress!

Dual-Use Technology
en.wikipedia.org/wiki/Dual-use_technology

Artificial Intelligence



Intelligence is a technology like no other!

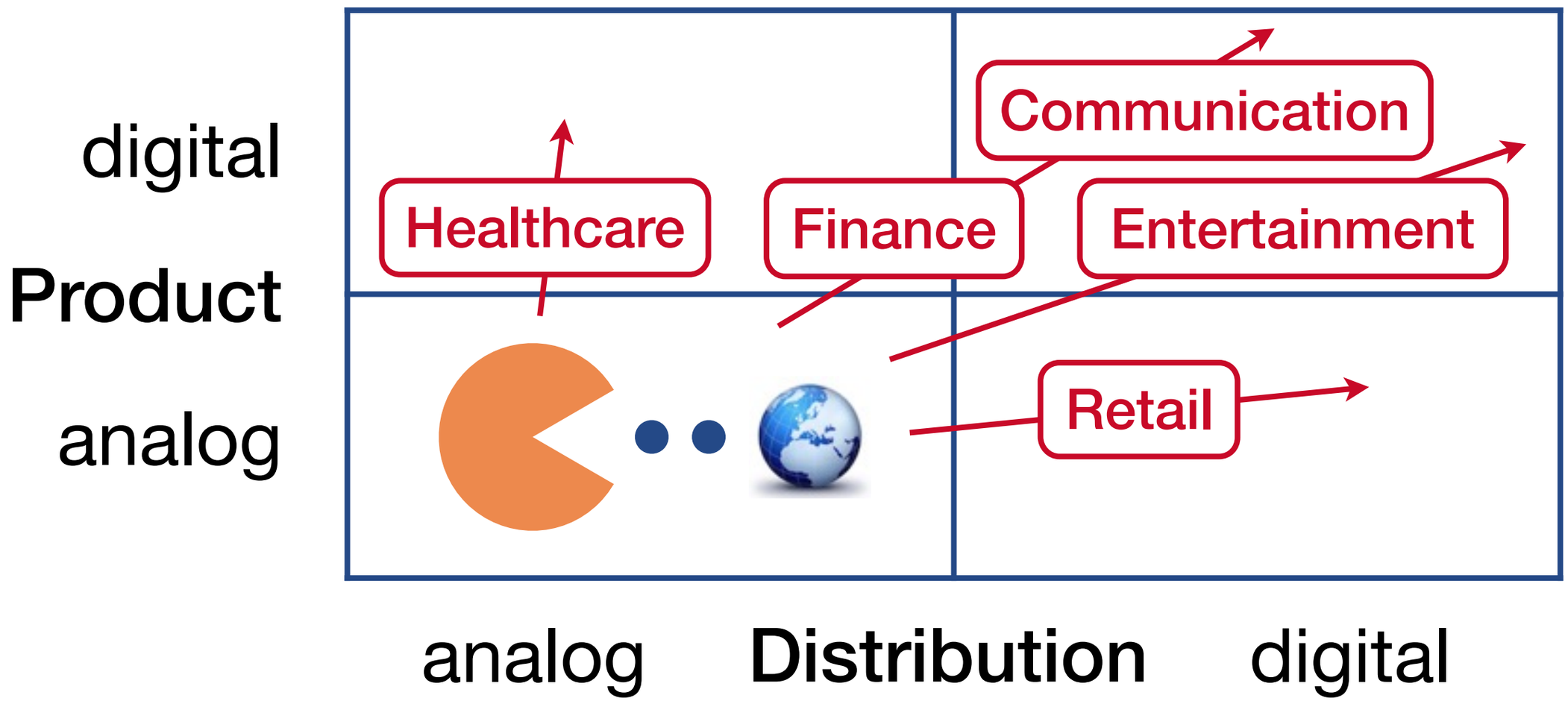
Current Trends

Where are we heading to?

gbs switzerland

Software is eating the world

… because it is more productive!



Marc Andreessen
on.wsj.com/1w2FbVs

Describes without errors	Describes with minor errors	Somewhat related to the image	Unrelated to the image

A person riding a motorcycle on a dirt road.

Two dogs play in the grass.

A skateboarder does a trick on a ramp.

A dog is jumping to catch a frisbee.

A group of young people playing a game of frisbee.

Two hockey players are fighting over the puck.

A little girl in a pink hat is blowing bubbles.

A refrigerator filled with lots of food and drinks.

A herd of elephants walking across a dry grass field.

A close up of a cat laying on a couch.

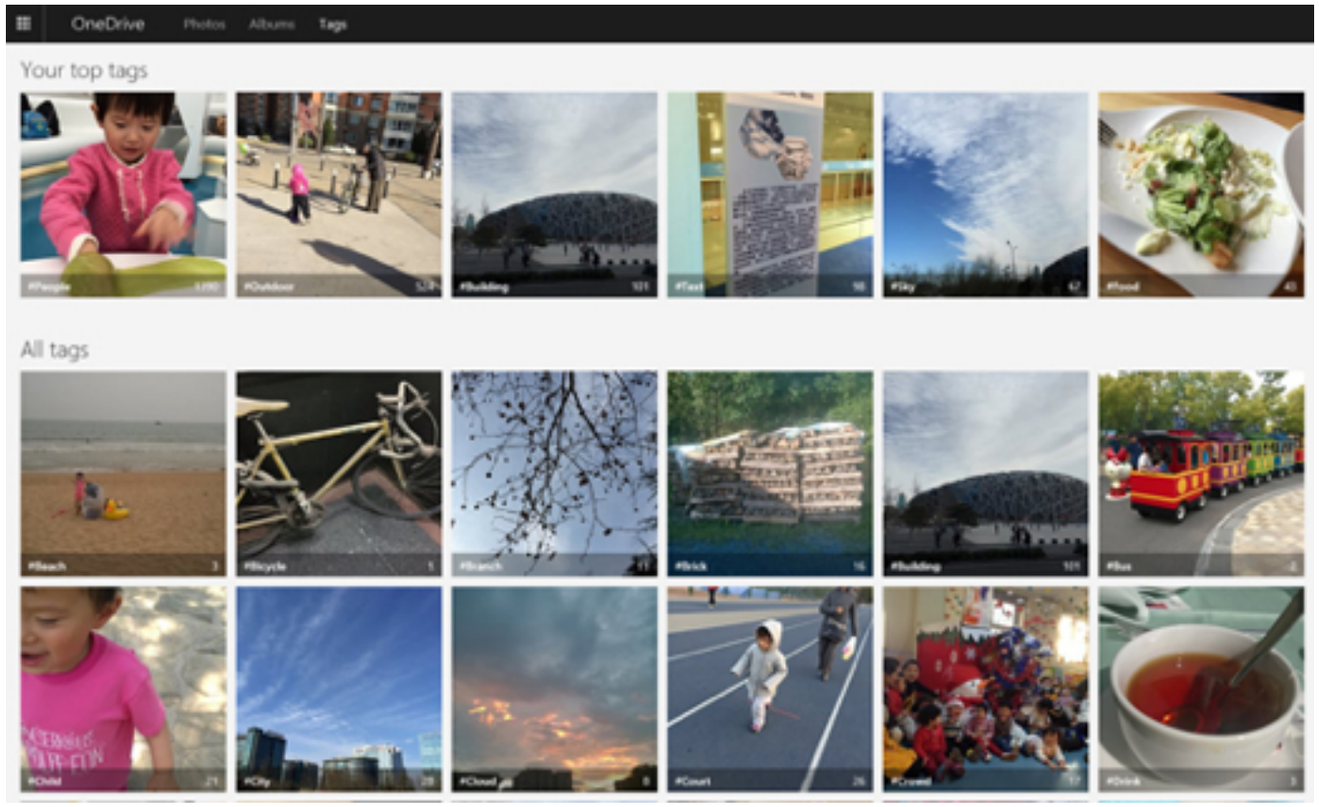
A red motorcycle parked on the side of the road.

A yellow school bus parked in a parking lot.

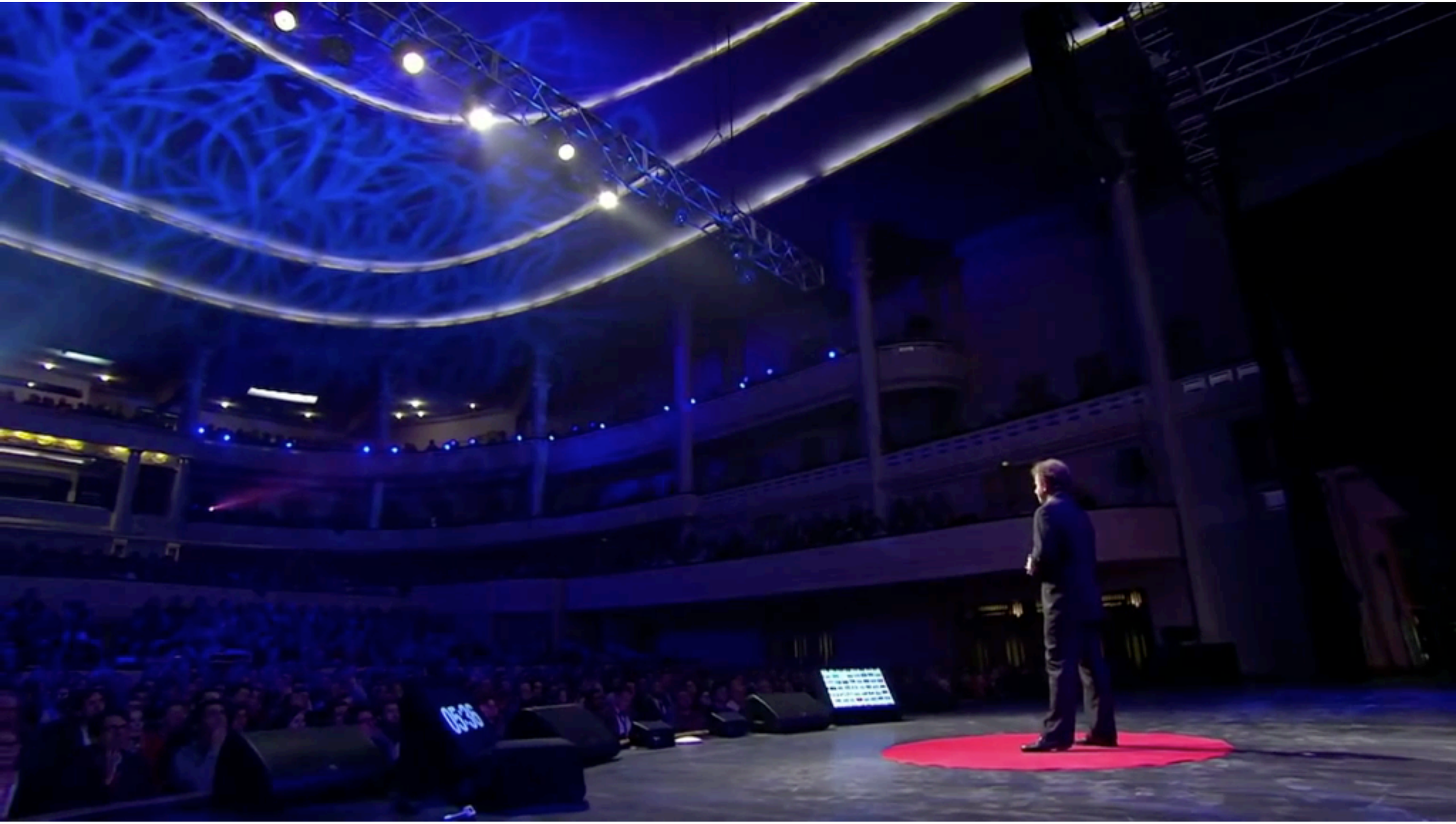
Google Research Blog, November 2014
googleresearch.blogspot.ch/2014/11/[…]

Superhuman Image Recognition

- With convolutional neural networks
- 1.2 m training images, ~ 30 layers



Inside Microsoft Research, February 2015
blogs.technet.com/b/inside_microsoft_research/[…]



Jeremy Howard
go.ted.com/bbZC

Machine Intelligence LANDSCAPE

CORE TECHNOLOGIES

ARTIFICIAL INTELLIGENCE
IBM WATSON, MetaMind, Numenta, ai-one, Cycorp, Microsoft Research, nara, Reactor, SI Scaled Inference

DEEP LEARNING
vicarious, Vision Factory, facebook, Baidu IDL, Google, ersatz, SKYMIND, SignalSense

MACHINE LEARNING
rapidminer, context relevant, 0xdata H2O, DATARPM, LiftIgniter, SPARKBEYOND, Azure ML, yhat, Sense, GraphLab, Alpine, Ionian

NLP PLATFORMS
cortical.io, idibon, LUMINOSO, wit.ai, Maluuba

PREDICTIVE APIS
AlchemyAPI, MINDOPS, Google, bigml, indico, ALGORITHMIA, Expect Labs, PredictionIO

IMAGE RECOGNITION
clarifai, MADBITS, DNNresearch, DEXTRO, VISENZE, lookflow

SPEECH RECOGNITION
GRIDSPACE, popup archive, NUANCE

RETHINKING ENTERPRISE

SALES
Preact, AVISO, RelateIQ, NGDATA, CLARABRIDGE, FRAMED, infer, ATTENSITY, causata

SECURITY / AUTHENTICATION
CROSSMATCH, conjur, EYEVERIFY, BITSIGHT, CYLANCE, AREA 1 SECURITY, bionym

FRAUD DETECTION
sift science, SOCURE, ThreatMetrix, feedzai, Brighterion, VERAFIN

HR / RECRUITING
TalentBin, entelo, predikt, Connectifier, gild, hiQ, CONCEPT NODE

MARKETING
brightfunnel, bloomreach, CommandIQ, AIRPR, RADIUS, TellApart, people pattern, Freshplum

PERSONAL ASSISTANT
Siri, X, Google now, Cortana, cleversense, tempo, Robinlabs, KASISTO, fuse|machines, VIV, CLARA LABS

INTELLIGENCE TOOLS
ADATAO, Palantir, Quid, Digital Reasoning, FirstRain

RETHINKING INDUSTRIES

ADTECH
METAMARKETS, dstillery, rocketfuel, YieldMo, ADBRAIN

AGRICULTURE
BLUE RIVER, TerrAvion, ceresimaging, HONEYCOMB, THE CLIMATE CORPORATION, tule

EDUCATION
declara, coursera, KNEWTON, kidaptive

FINANCE
Bloomberg, FinGenius, alphasense, KENSHO, Dataminr, minettabrook, BINATIX

LEGAL
Lex Machina, brightleaf, COUNSELYTICS, RAVEL, JUDICATA, Brevia, DiligenceEngine

MANUFACTURING
SIGHT MACHINE, MICROSCAN, BOULDER IMAGING, IVISYS

MEDICAL
Parzival, transcriptic, ZEPHYR HEALTH, Genescient, grand round table, bina TECHNOLOGIES, TUTE GENOMICS

OIL AND GAS
kaggle, AYASDI, TACHYUS, biota, Flutura

MEDIA / CONTENT
Outbrain, newsle, ARRIA, SAILTHRU, wavii, Owlin, NarrativeScience, yscope, Summly, Prismatic, ai

CONSUMER FINANCE
affirm, inMenture, zest finance, BILLGUARD, LendUp, LendingClub, Kabbage

PHILANTHROPIES
DataKind, thorn, DATA GUILD

AUTOMOTIVE
Google, Continental, TESLA, MOBILEYE, CRUISE

DIAGNOSTICS
enlitic, 3SCAN, lumiata, ENTOPIS

RETAIL
BAY SENSORS, PRISM SKYLABS, celect, euclid

RETHINKING HUMANS / HCI

AUGMENTED REALITY
wearable intelligence, APX, blippar, META, layar

GESTURAL COMPUTING
THALMIC LABS, omek, FLUTTER, LEAP MOTION, eyeSight, 3Gear, GestureTek, nod

ROBOTICS
intel, LIQUID ROBOTICS, SoftBank, iRobot, Boston Dynamics, jibo, anki, evolution robotics

EMOTIONAL RECOGNITION
affectiva, BEYOND VERBAL, EMOTIENT, BRIGHAM, cogito

SUPPORTING TECHNOLOGIES

HARDWARE
NVIDIA, XILINX, QUALCOMM, NERVANA SYSTEMS, TERADEEP, Artificial Learning, rigetti

DATA PREP
TRIFACTA, Paxata, tamr, Alation

DATA COLLECTION
diffbot, kimono, CrowdFlower, Connotate, WorkFusion, import.io

www.shivonzilis.com/machineintelligence

Bloomberg BETA

Automation brings abundance

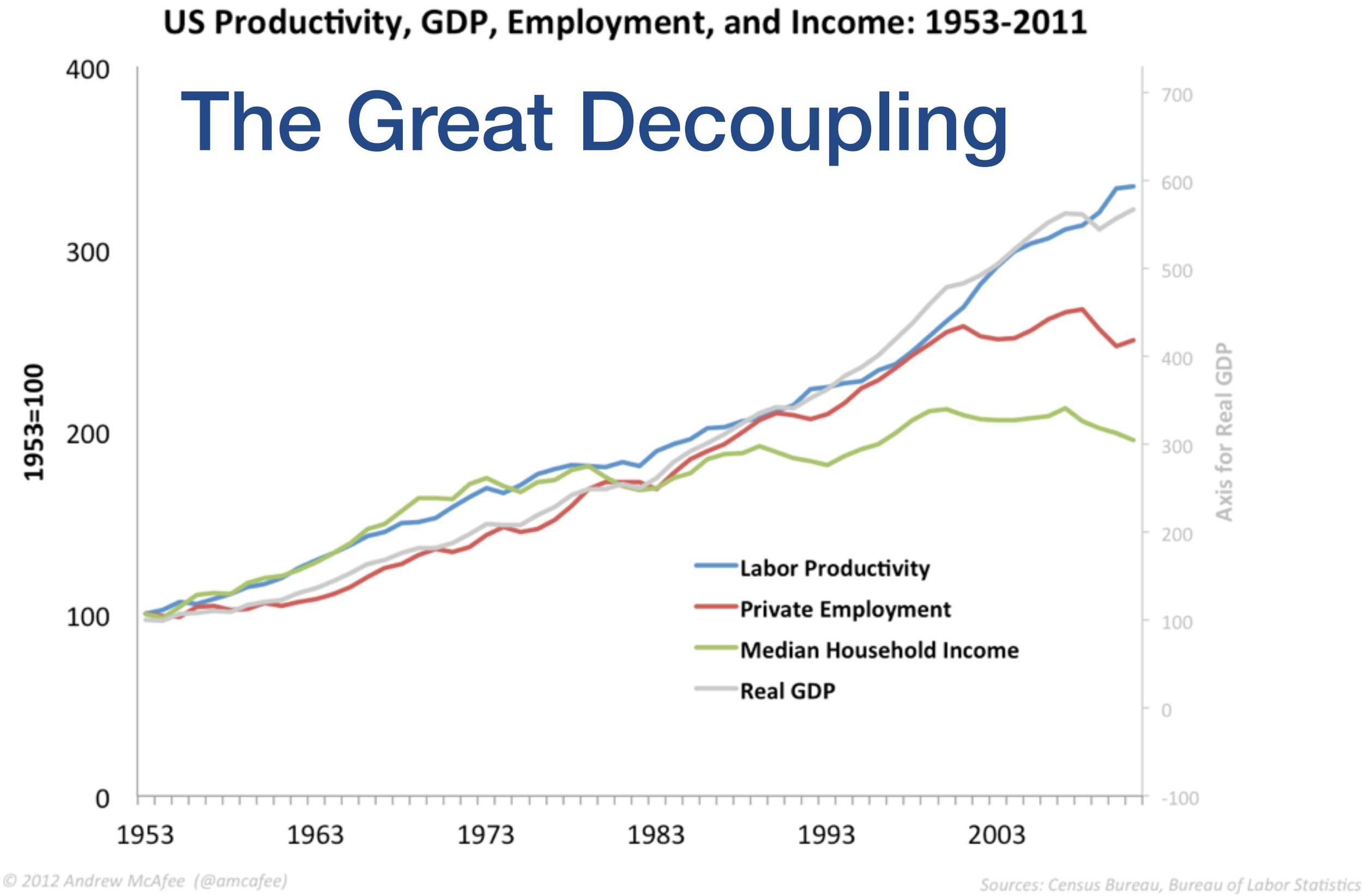
Many industries will be transformed into pure AI industries…

Capital: ⬆, Wages: ⬇

Should be beneficial if we manage it well – but we are not prepared!

Erik Brynjolfsson
go.ted.com/biXe

The Great Decoupling



The Great Decoupling of the US Economy
andrewmcafee.org/2012/12/[…]

We'll give robots full autonomy

… because of

- increased speed
- high complexity
- risk of jamming

Current examples:

- financial markets
- auton. weapons

AI is the ultimate productivity boost!

Daniel Suarez
go.ted.com/biXE

Machine Learning by Google

DEEPMIND

Google's Algorithms Learn Tasks Independently
www.theguardian.com/technology/2015/feb/25/[…]

Reinforcement Learning

Train with first-person shooters and deploy on armed drones…

Difficulties

What might go wrong?

A rational agent will strive to …

- stay functional (**self-preservation**)
- keep its goal (**goal-preservation**)
- get stuff (**resource accumulation**)
- be smarter (**intelligence explosion**)

The problem is not malevolence but different goals and higher decision quality!

Stephen M. Omohundro: The Basic AI Drives
selfawaresystems.[…].com/2008/01/ai_drives_final.pdf

Single-Shot Situation

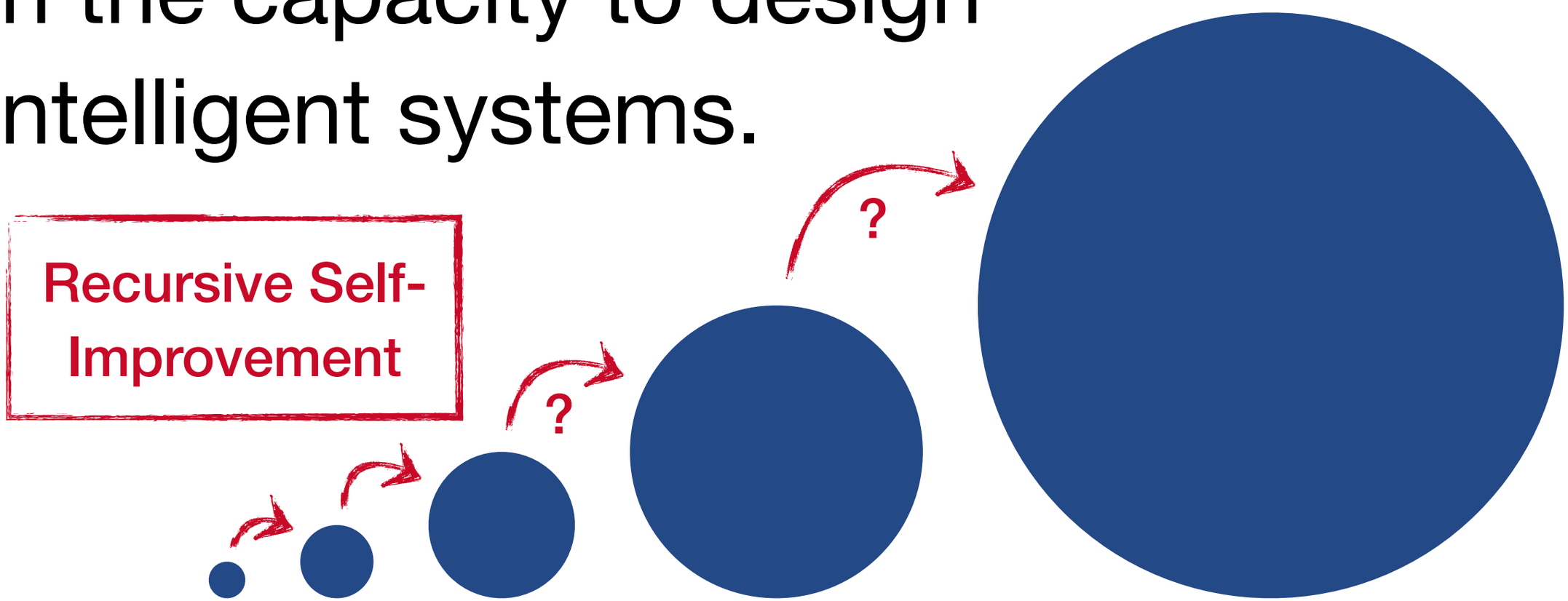
Our first superhuman AI must be a safe one for we may not get a second chance!

- We're good at iterating with **testing** and **feedback**
- We're terrible at getting things right **the first time**
- Humanity only learns when catastrophe occurred

List of Cognitive Biases
en.wikipedia.org/wiki/List_of_cognitive_biases

Intelligence Explosion

Proportionality Thesis: An increase in intelligence leads to similar increases in the capacity to design intelligent systems.



Intelligence Explosion
intelligence.org/files/IE-EI.pdf

Technological Singularity

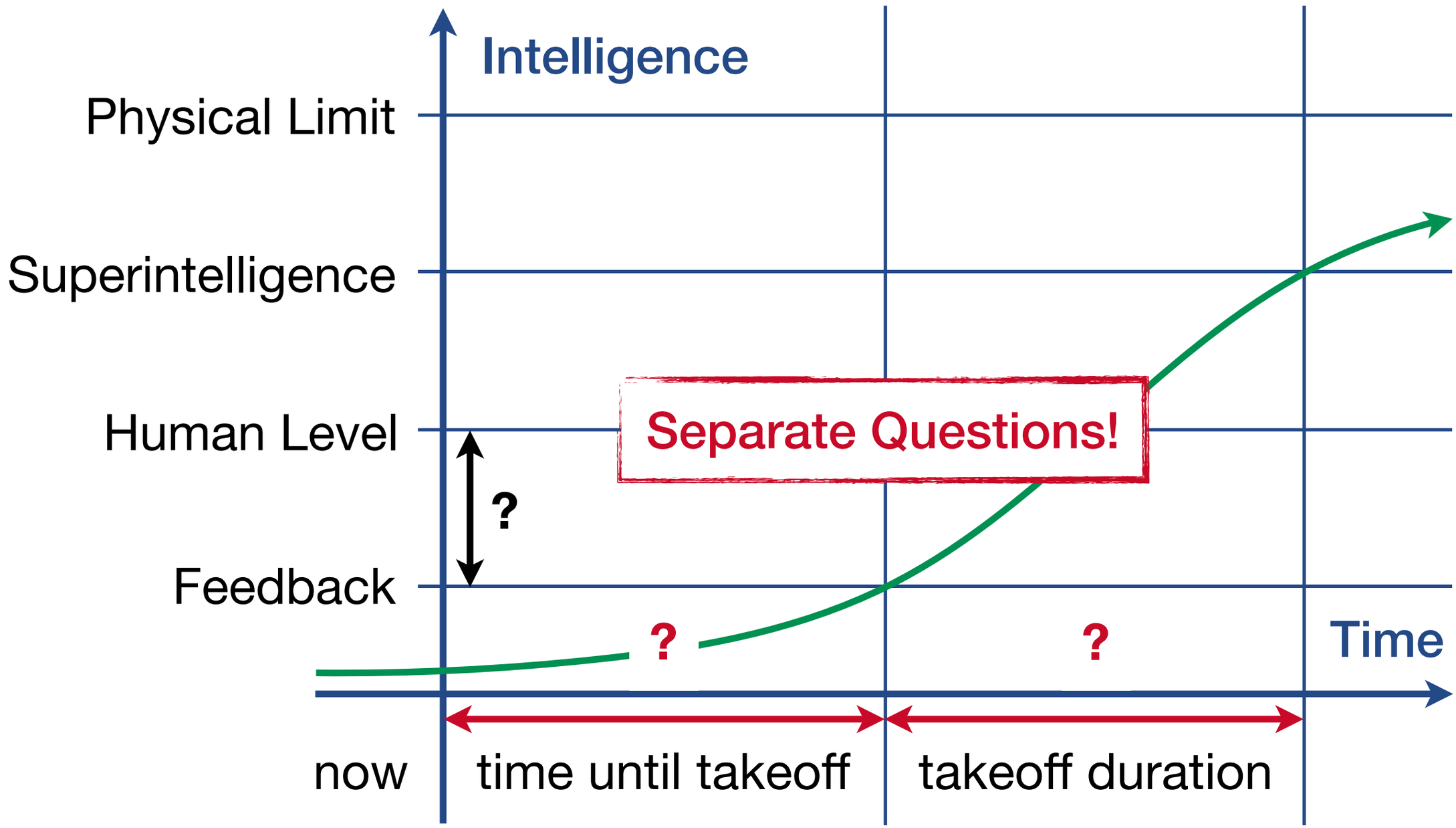
Theoretic phenomenon: There are arguments why it should exist but it has not *yet* been confirmed experimentally.

Three major singularity schools:

- **Accelerating Change** (Ray Kurzweil)
- **Intelligence Explosion** (I.J. Good)
- **Event Horizon** (Vernor Vinge)

David Chalmers: The Singularity
consc.net/papers/singularity.pdf

Takeoff Scenarios



Transcendence with Johnny Depp
www.transcendencemovie.com

Advantages of AIs over Brains

Hardware:
- Size
- Speed
- Memory

Software:
- Editability
- Copyability
- Expandability

Effectiveness:
- Rationality
- Coordination
- Communication

Human Brain	Modern Microprocessor
86 billion neurons	1.4 billion transistors
firing rate of 200 Hz	4’400’000’000 Hz
120 m/s signal speed	300’000’000 m/s

Advantages of AIs, Uploads and Digital Minds
kajsotala.fi/Papers/DigitalAdvantages.pdf

Interview by John Oliver with Stephen Hawking
www.youtube.com/watch?v=T8y5EXFMD4s

Modelling Capabilities

An advanced AI will also model its operators and go to great lengths to prevent being switched off!

It will behave nicely and cooperatively until the **external threats** are under control and it is ready for takeover.

Ex Machina Movie
exmachina-movie.com

Optimization Power

Problem: When optimizing a system, unspecified parameters often assume extreme values.

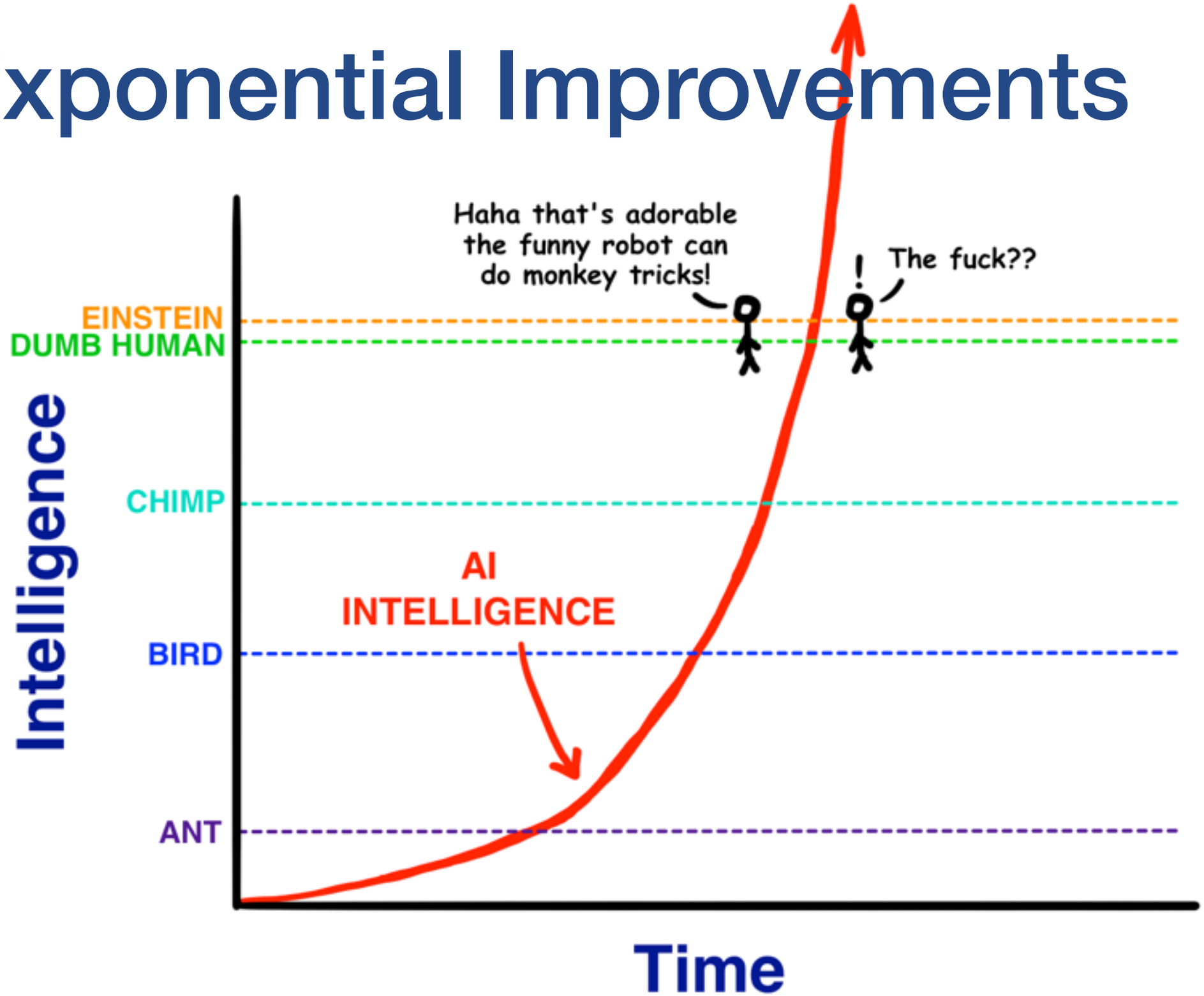
You will get what you wished for and not what you wanted

Coherent Extrapolated Volition
intelligence.org/files/CEV.pdf

Outlook

What is ahead of us?

Exponential Improvements



The Road to Superintelligence
waitbutwhy.com/2015/01/[…]

Predicting AI Timelines

Great uncertainties:

- Hardware or software the bottleneck?
- Small team or a Manhattan Project?
- More speed bumps or accelerators?

Probability for AGI	10%	50%	90%
AI scientists, median	2024	2050	2070

How We're Prediciting AI – or Failing To
intelligence.org/files/PredictingAI.pdf

Speed Bumps

- Depletion of low-hanging fruit
- An end to Moore's law
- Societal collapse
- Disinclination

Accelerators

- Faster **hardware**
- Better **algorithms**
- Massive **datasets**

+ enormous economic, military and egoistic incentives!

The breakthroughs that have unleashed AI on the world
www.wired.com/2014/10/future-of-artificial-intelligence/

Flow-Through Effects

Going meta: Solve the problem-solving problem!

- Extreme Poverty
- Factory Farming
- Climate Change
- Artificial Intelligence

could solve other issue

Holden Karnofsky: Flow-Through Effects
blog.givewell.org/2013/05/15/flow-through-effects/

Controlled Detonation

Difficulty:

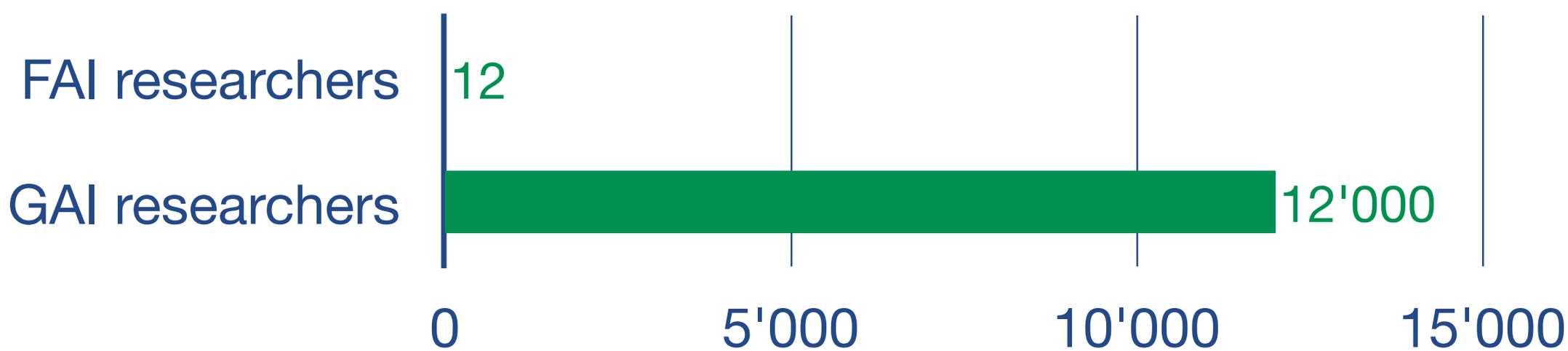
Friendly AI >> General AI

AI as a Positive and Negative Factor in Global Risk
intelligence.org/files/AIPosNegFactor.pdf

Differential Intellectual Progress

Prioritize risk-reducing intellectual progress over risk-increasing intellectual progress

AI safety should outpace **AI capability** research



Differential Intellectual Progress as a Positive-Sum Project
foundational-research.org/[…]/differential-progress-[…]/

International Cooperation

- We are the ones who will create superintelligent AI
- Not primarily a **technical problem**, rather a social
- **International regulation?**

In face of uncertainty, cooperation is robust!

Lower Bound on the Importance of Promoting Cooperation
foundational-research.org/[…]/[…]-promoting-cooperation/

2014: A turning point in AI safety!

Many smart people take superintelligence very seriously.

Future of Life Institute: AI Conference
futureoflife.org/misc/ai_conference

With great power comes great responsibility!