Artificial Intelligence

Implications: Chances and Risks

Büroautomationskonferenz der SIK
Luzern, 19. September 2017

Kaspar Etter
www.superintelligence.ch

About Myself: Startup Synacts

Jason Blum
Business

M.A. HSG in Business Mgmt.
B.A. HSG in Business Admin.
Focus on Strategy Execution

Stephanie Stroka
Product

M.Sc. in Computer Science
Form. Lead Dev. at AdNovum
Semantic Web and IAM Expert

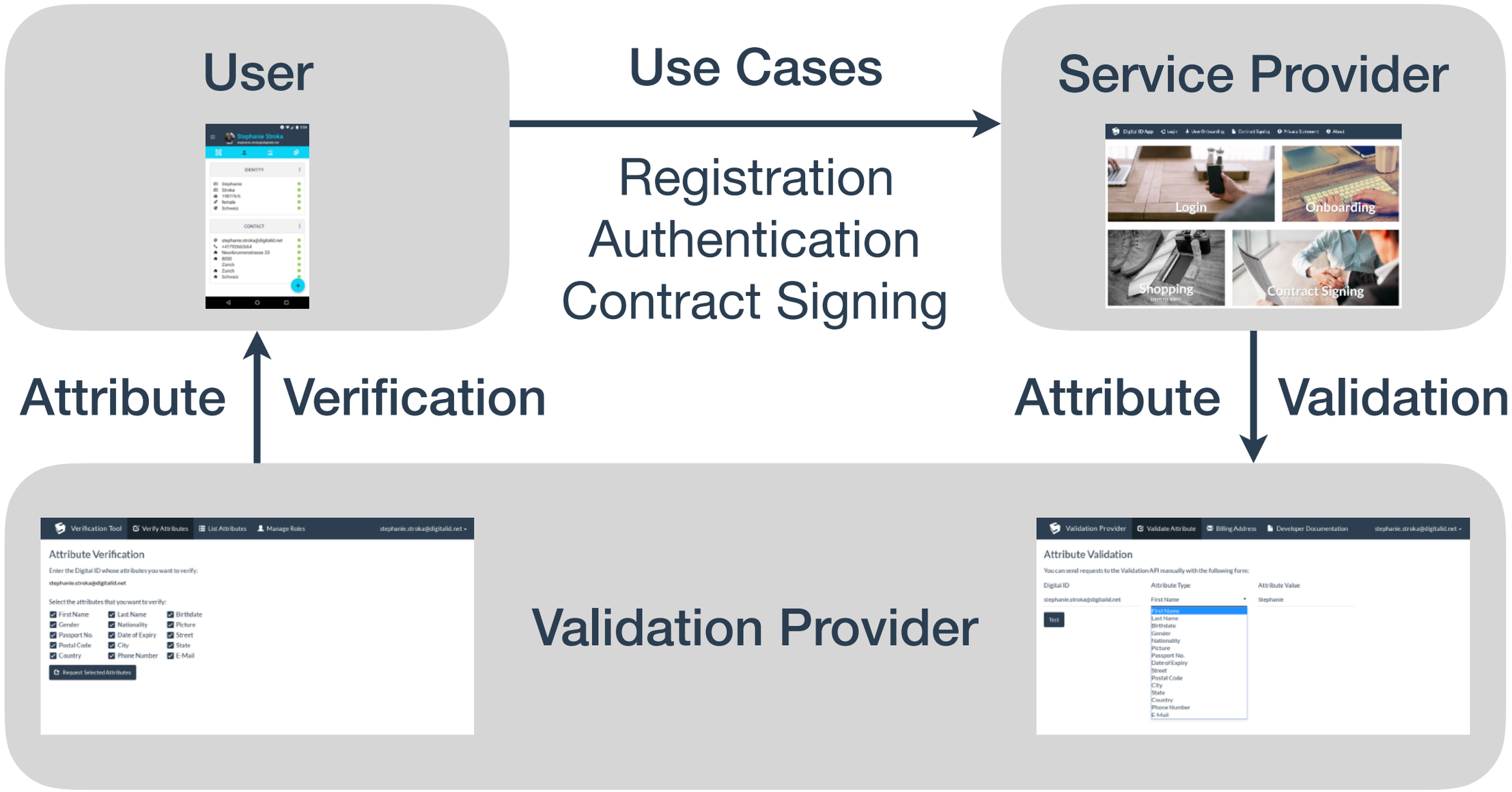
Kaspar Etter
Technology

M.Sc. ETH in Computer Science
Author of the Digital ID Protocol
Identity and Cryptography Expert

Julia Wüthrich
Operations

M.A. HSG in Business Innovation
B.A. in Business Administration
Eidg. dipl. Mediamatician

Our Solution: Verified Identity



Demo Page
app.digitalid.net

Comments on Blockchain (1/2)

Problem: Prevent Double-Spending.

Solution: Agree on the history of transactions – and thus the current state – without a trusted third party.

Issues: No consensus on real world + order of transactions often irrelevant.

Wikipedia: Blockchain
en.wikipedia.org/wiki/Blockchain

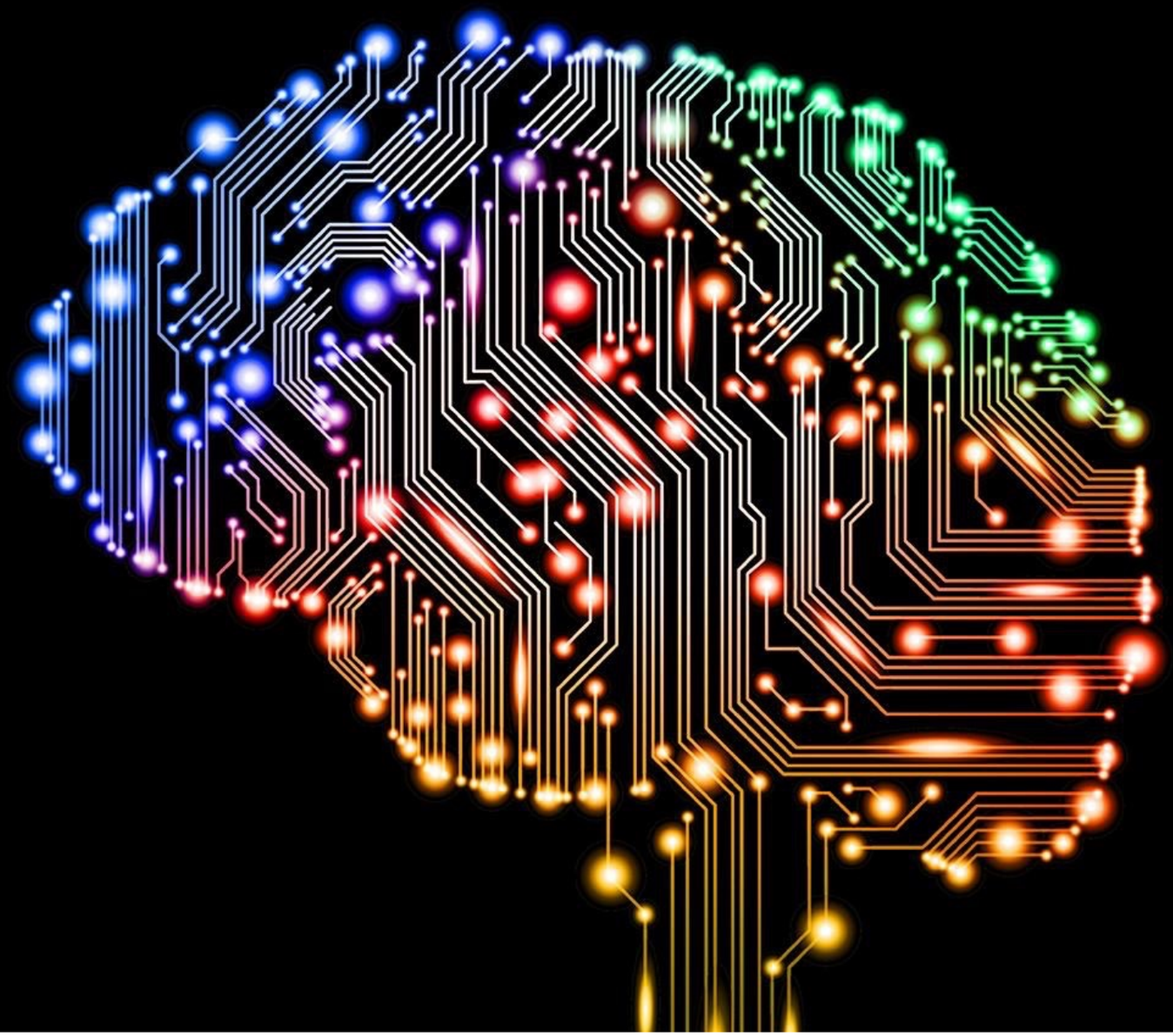
Comments on Blockchain (2/2)

The killer applications of blockchains:

- censorship-resistant payments
- perfect wealth concealment
- timestamping and namespaces

Libertarian crusade against states!

The Top Use Cases For Bitcoin by Tone Vays
libertylifetrail.com/[…]/the-top-use-cases-for-bitcoin/



Introduction

What are we talking about?

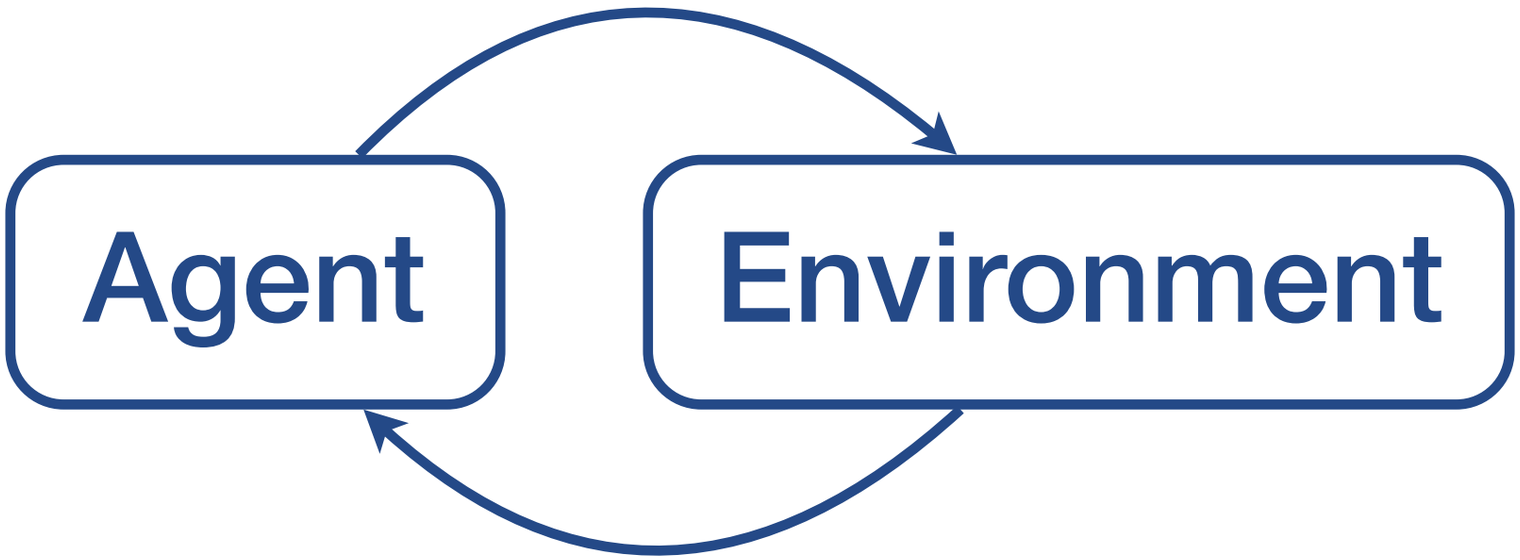
Intelligence

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

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

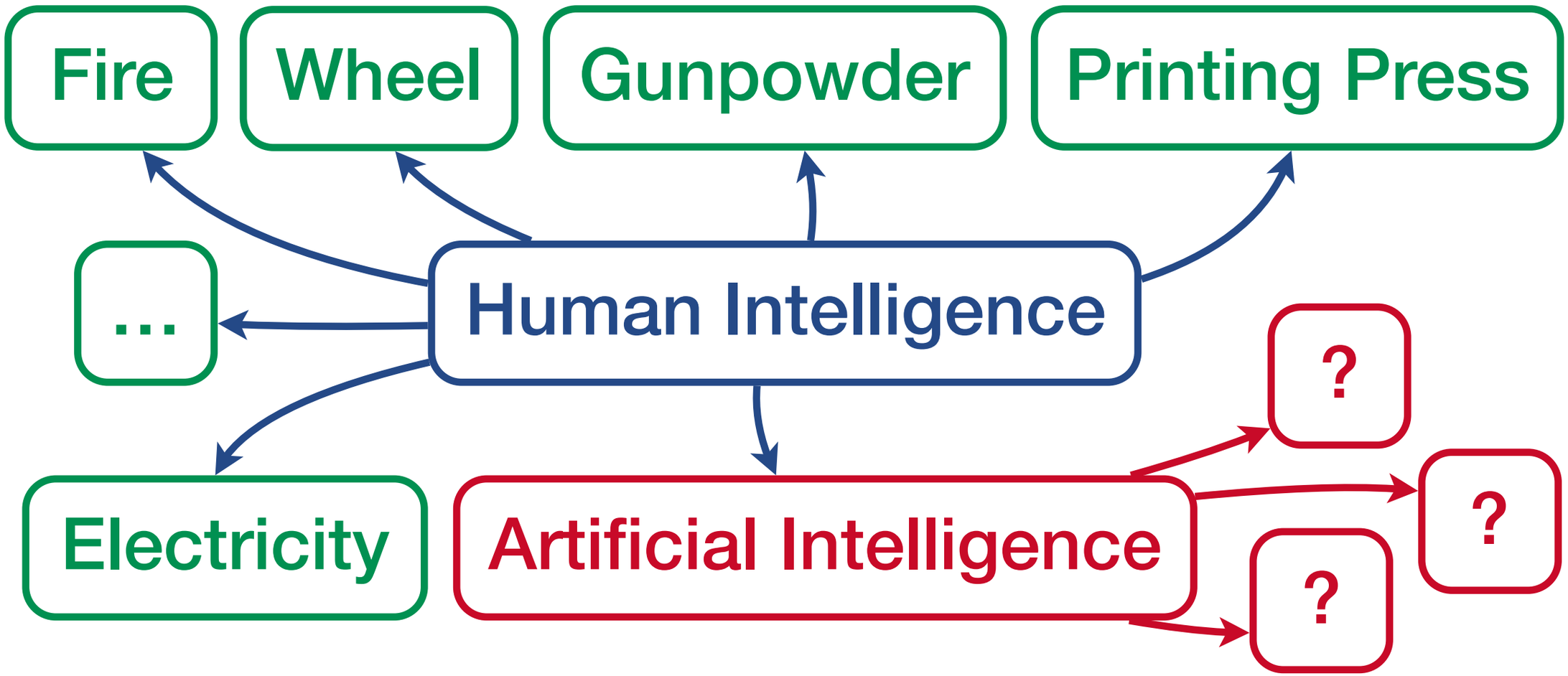
Universal Intelligence
arxiv.org/pdf/0712.3329.pdf

Interaction

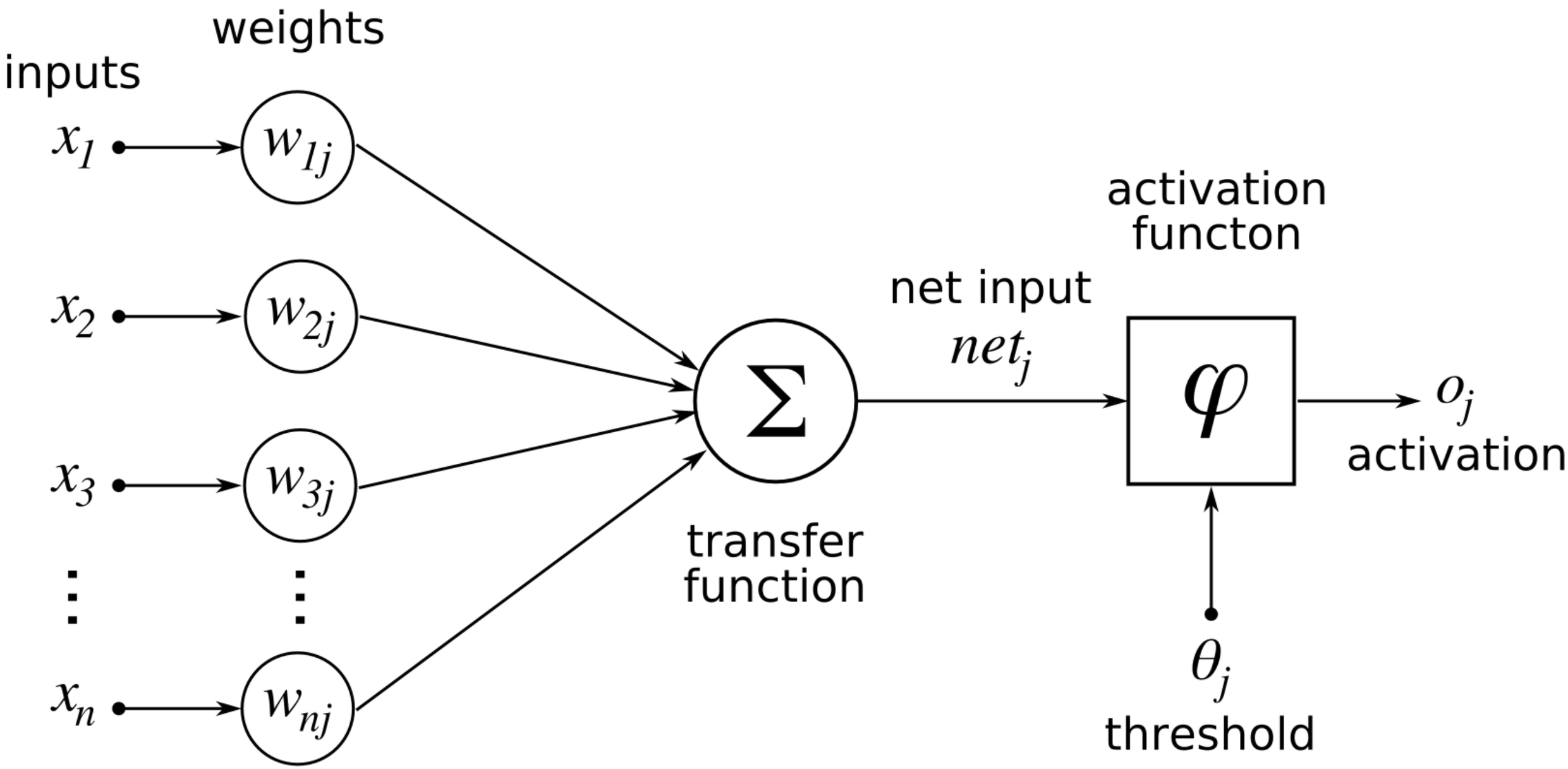


Learn, predict, rate and plan!

Artificial Intelligence



Intelligence is a technology like no other!



Machine Learning

How does it work?

Learning

- Building **predictive** models from data
- Algorithms improve with “experience”
- Useful if you don’t know the solution

- **Machine Learning**: Learn and predict
- **Data Mining**: Discover new properties

Wikipedia: Machine Learning
en.wikipedia.org/wiki/Machine_learning

Feedback Mechanism

Categorization based on feedback:

- **Supervised Learning**: Given a set of inputs & outputs, learn a general rule
- **Unsupervised Learning**: Learn the structure in input without training set
- **Reinforcement Learning**: Achieve a certain goal in dynamic environment

Wikipedia: Supervised Learning
en.wikipedia.org/wiki/Supervised_learning

Tasks and Models

- **Classification**: Assign unseen input to descrete categories (supervised)
- **Regression**: Map the unseen input to continuous values (supervised)
- **Clustering**: Divide set of inputs into a number of groups (unsupervised)
- **Dimensionality Reduction**: Simplify

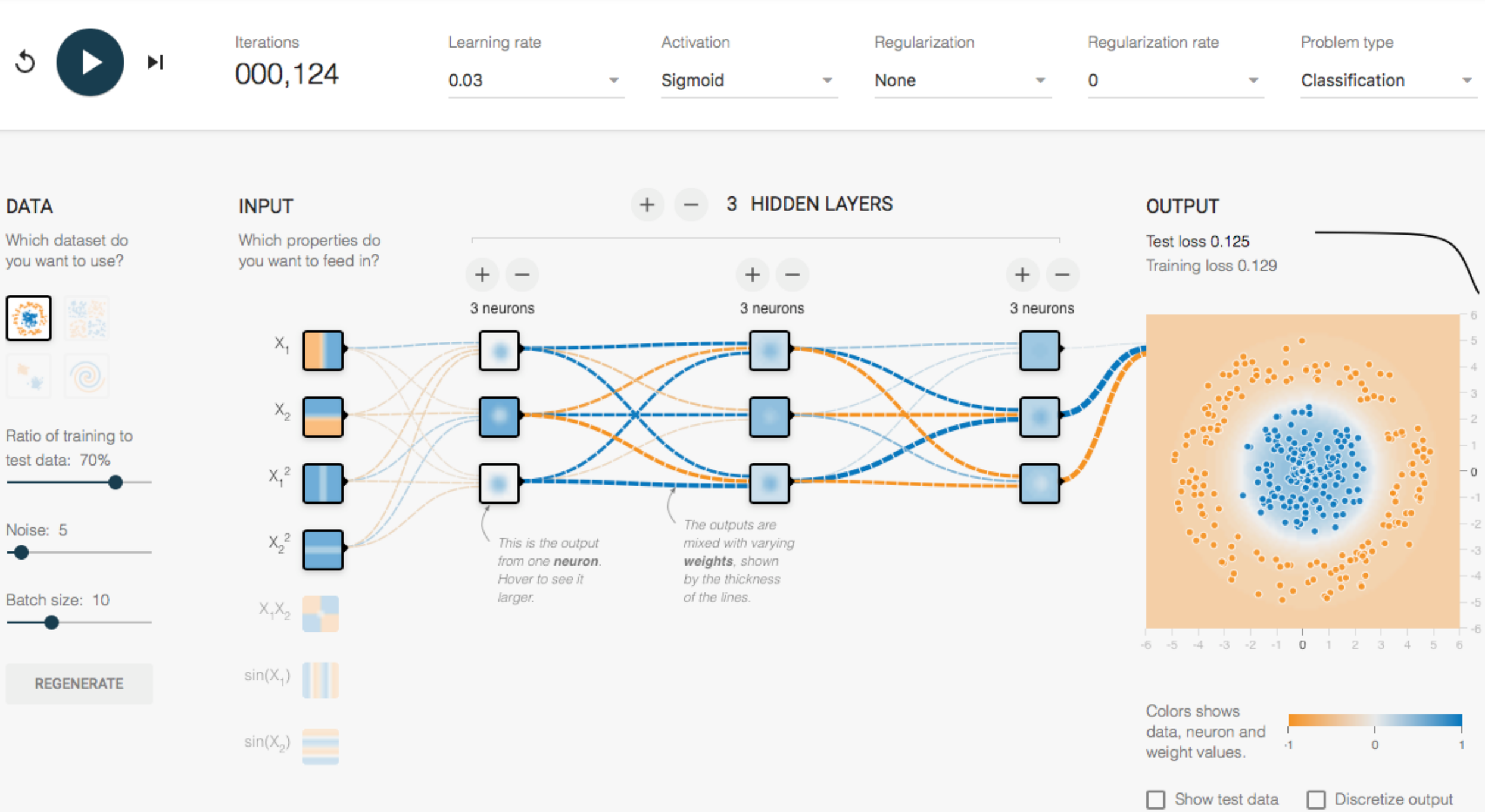
Statistical Classification
en.wikipedia.org/wiki/Statistical_classification

Reinforcement Learning (CNN)

DEEPMIND

Google develops self-learning computer program
www.theguardian.com/technology/2015/feb/25/[…]

Train your own Neural Network

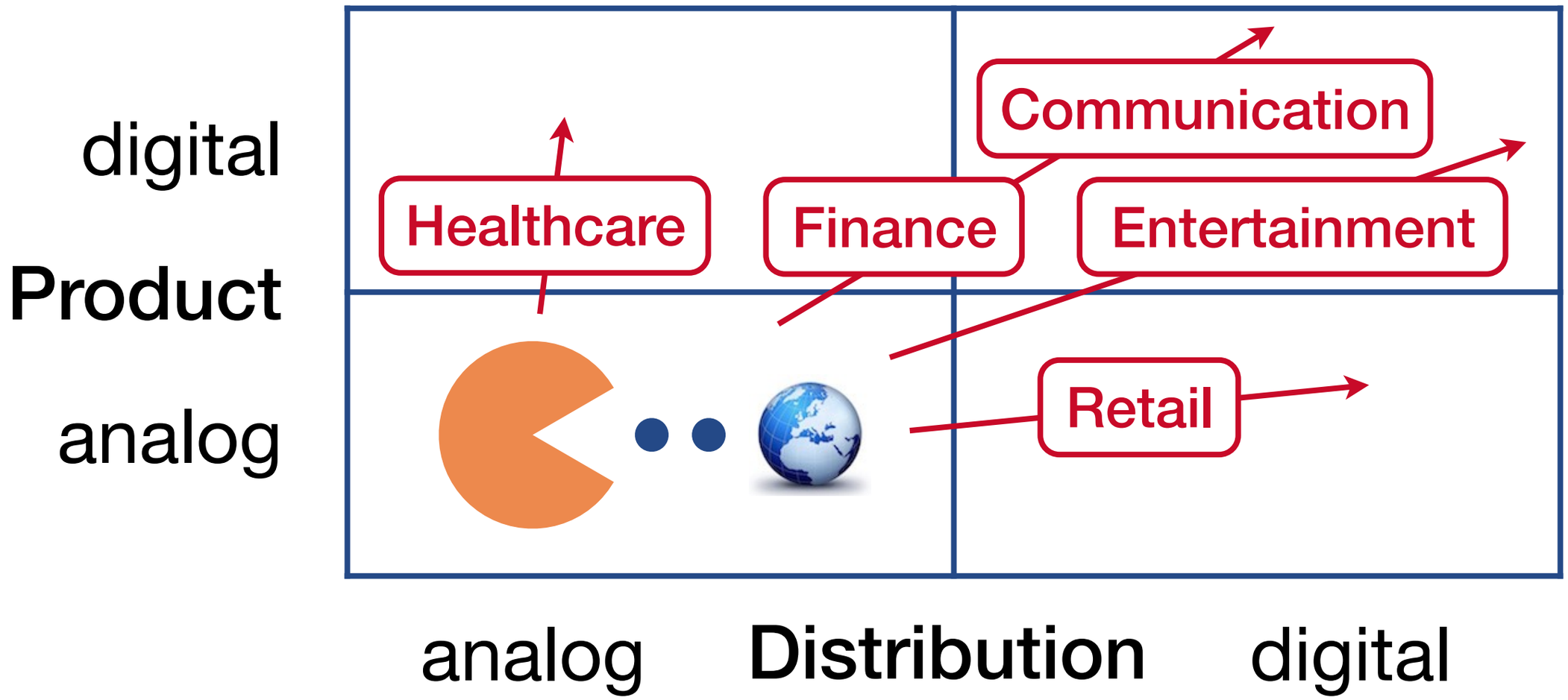


Neural Network in Browser
playground.tensorflow.org

Current Trends

Where are we heading?

Software is eating the world



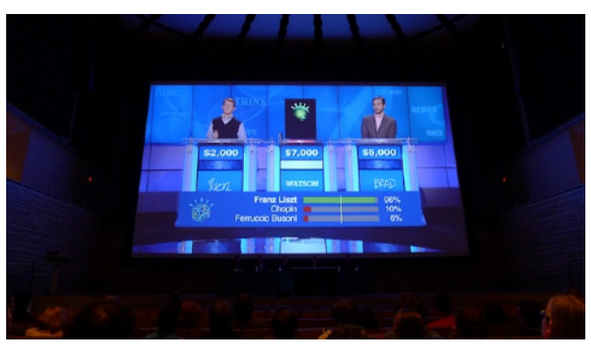
… and AI will digest all this (big) data!

Marc Andreessen
on.wsj.com/1w2FbVs

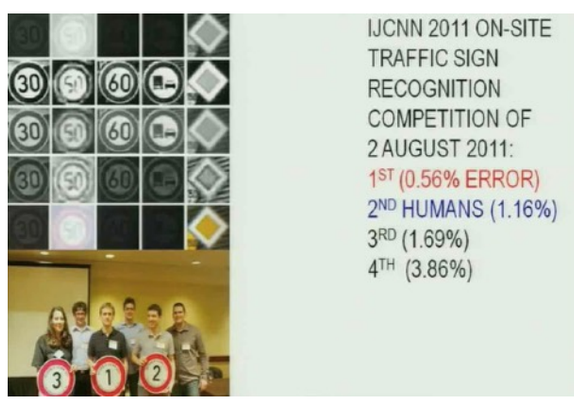
State of the Art

Deep Blue: 1997

Stanley: 2005



IBM Watson: 2011



Schmidhuber: 2011

Checkers	Superhuman
Backgammon	Superhuman
Othello	Superhuman
Chess	Superhuman
Crosswords	Expert Level
Scrabble	Superhuman
Bridge	Equal to Best
Jeopardy!	Superhuman
Poker	Varied
FreeCell	Superhuman
Go	Superhuman

How bio-inspired learning keeps winning competitions
www.kurzweilai.net/how-bio-inspired-deep-learning-[…]

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.

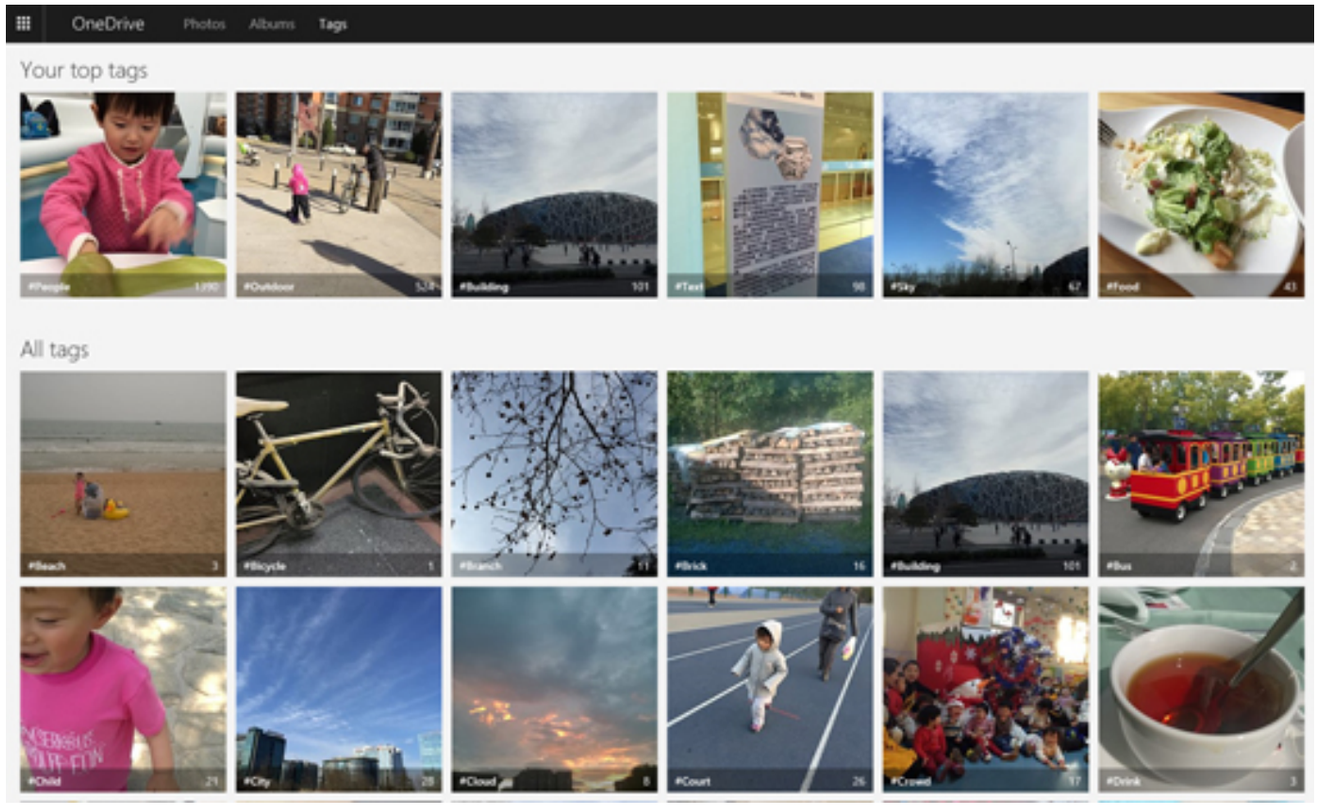
A red motorcycle parked on the side of the road.

A yellow school bus parked in a parking lot.

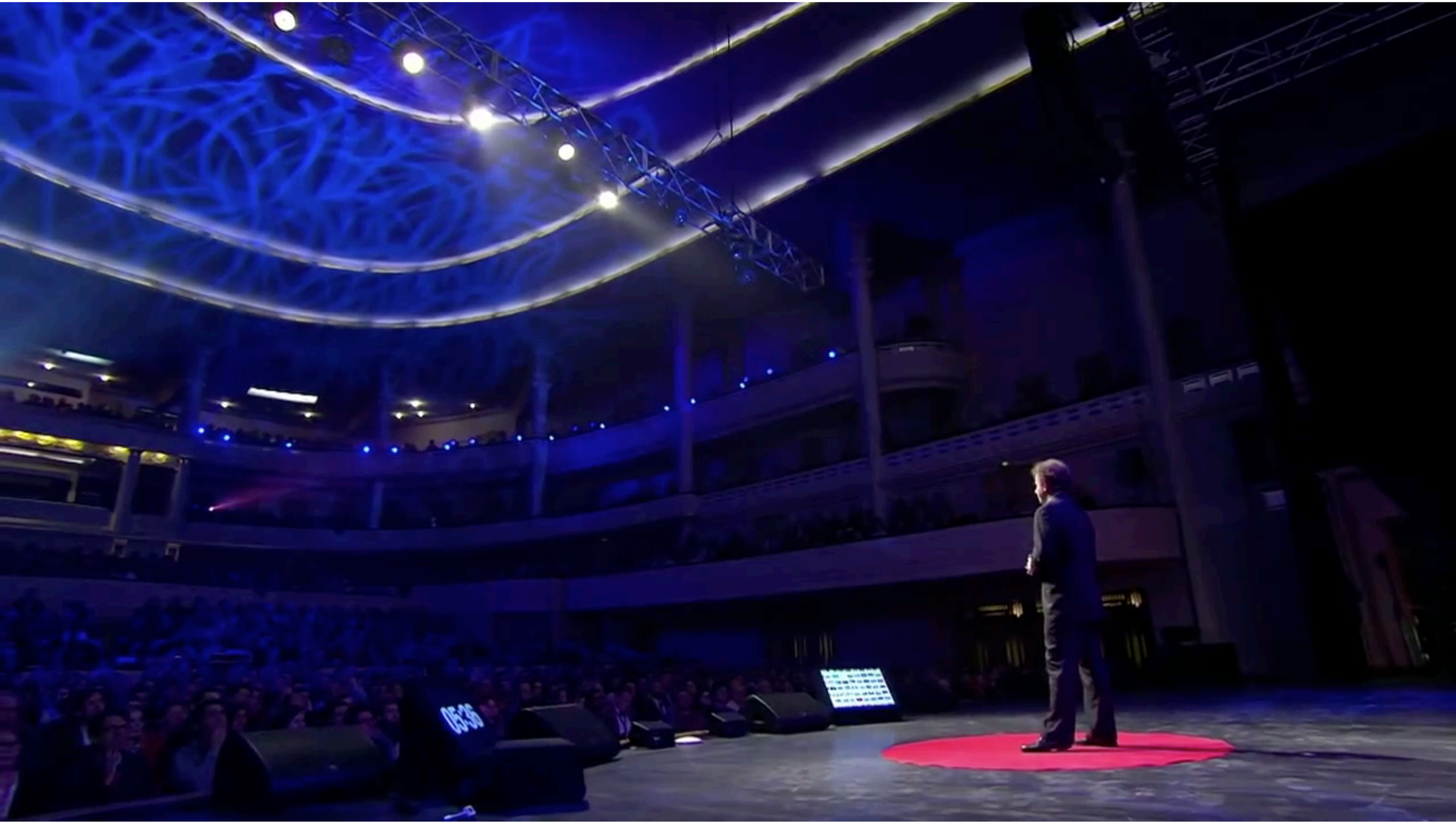
Google Research Blog, November 2014
googleresearch.blogspot.ch/2014/11/[…]

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: The implications of learning computers
www.ted.com/talks/jeremy_howard_the_wonderful_[…]

MACHINE INTELLIGENCE 2.0

AGENTS

PROFESSIONAL
Howdy! x.ai clara KASISTO DigitalGenius OVERLAP.CC meekan fuse|machines PRIMER

PERSONAL
facebook XIAOICE assistant.ai large nestor @wesome Magic

OS INTERFACES
Siri Cortana VIV maluuba api.ai COGNEA Google Now

AUTONOMOUS SYSTEMS

AIR
3DR dji PROJECT LOON Google VERTICAL DroneDeploy AIRDOG SKYCATCH SKYDIO Airware LILY

GROUND
Google UBER TESLA CRUISE MOBILEYE COMMA AdasWorks

SEA
LIQUID ROBOTICS bluefin data OPENROV BluHaptics

INDUSTRIAL
KIVA Systems fetch HARVEST CLEARPATH AVIDBOTS ENERGID rethink robotics GREYORANGE OSARO

ENTERPRISE

SECURITY / FRAUD
Sentinel graphistry BITSIGHT feedzai AREA 1 drawbridge sift science CYLANCE Brighterion

HR / RECRUITING
textio hiQ gild SpringRole entelo unitive GIGSTER

SALES
6sense clari infer Preact Gainsight AVISO salespredict sentient Vidora people pattern Prism

MARKETING
LiftIgniter RADIUS brightfunnel retention science AIRPR

CUSTOMER SUPPORT
CLARABRIDGE QUANTIFIND Wiseio ACTIONIQ FRAMED DigitalGenius

INTERNAL INTEL
Alation ADATAO Palantir sapho lucid Rainbird SKIPFLAG agolo Digital Reasoning Narrative Science

MARKET INTEL
Quid mattermark DataFox bottlenose PREMISE enigma CB INSIGHTS

PLATFORMS

RESEARCH / AGI
OpenAI vicarious Google DeepMind Numenta Cycorp nnaisense SCALED INFERENCE CURIOUS GEOMETRIC INTELLIGENCE

FULL STACK
context relevant CognitiveScale NVIDIA TERADEEP QUALCOMM nervana SYSTEMS

MACHINE LEARNING
Dato rapidminer cortical.io AYASDI amazon web services Azure Machine Learning naralogics PredictionIO SKYTREE bigml blueyonder

INDUSTRIAL IOT
ThingWorx UPTAKE IMUBIT Preferred Networks Alluvium xively PLANET OS

AUDIO
Gridspace TalkIQ nexidia vocaliq NUANCE Expect Labs popUp archive

VISION
ORBITAL INSIGHT Descartes Labs DEXTRO cortica clarifai PLANET LABS MetaMind

DATA ENRICHMENT
diffbot Paxata TRIFACTA iDIBON WorkFusion loop ai labs CrowdFlower

INDUSTRIES

ADTECH
ADTHEORENT dstillery BEYONDVERBAL METAMARKETS TAPAD rocketfuel affectiva

AGRICULTURE
BLUE RIVER tule TerrAvion mavrx THE CLIMATE CORPORATION CERES HONEYCOMB

FOR GOOD
Conservation Metrics DataKind DATA POP thorn BAYES IMPACT

RETAIL FINANCE
inVenture affirm earnest MIRADOR Lendo zest finance LendUp

LEGAL
Everlaw RAVEL LEGAL ROBOT seal BEAGLE ROSS Lex Machina

MATERIALS & MFG
zymergen AUGMATE GINKGO BIOWORKS CITRINE TECHNOLOGIES SIGHT MACHINE CALCULARIO Eigen Innovations

HEALTHCARE
deep genomics 3SCAN enlitic Calico Atomwise Recombine color METABIOTA GRAND ROUNDS Google Life Sciences IBM Watson Health

INDUSTRIES (CONT'D)

EDUCATION
KNEWTON coursera turnitin gradescope UDACITY KHANACADEMY

TRANSPORT & LOGISTICS
NAUTO taleris PRETECKT clearmetal

INVESTMENT FINANCE
Bloomberg Quantopian Dataminr KENSHO iSENTIUM NEURENSIC alphasense

TECH USER TOOLS

DATA SCIENCE
DOMINO kaggle Sentenai sense yseop Outlier yhat DataRobot

MACHINE LEARNING
Cortana Analytics AlchemyAPI glowfi.sh IBM Watson Platform Anodot MonkeyLearn (h [s]) HyperScience fuzzy.io SIGOPT 0xdata H2O SPARKBEYOND indico

OPEN SOURCE
SKYMIND TensorFlow seldon Caffe theano Spark MLlib Microsoft DMTK spaCy DL4J SciKit CGT

Outlook

What is next?

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: Growth? Race with the machines!
www.ted.com/talks/erik_brynjolfsson_the_key_to_[…]

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

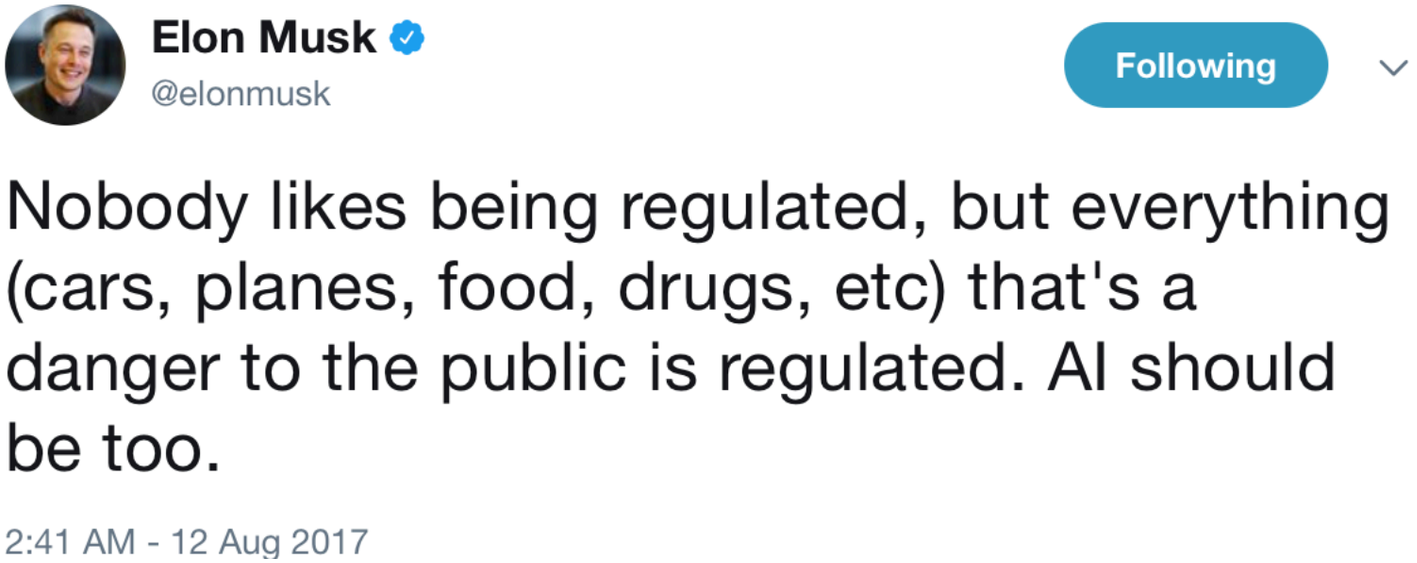
Implications of (advanced) AI

- **Short-term** (now to 15 years): classification/ assistance, surveillance/intelligence, autonomous systems (cars, weapons, etc.) (AIs as moral agents)
- **Mid-term** (15 to 40 years): automation/unemploym., cooperation/arms races, control/principal agent problem, genie/oracle problem (unintended effects)
- **Long-term** (40 to 100 years): scientific breakthroughs by AIs, recursive self-improvement/runaway AI (?), artificial consciousness (?) (AIs as moral patients)

Artificial Intelligence: Opportunities and Risks
ea-foundation.org/files/ai-opportunities-and-risks.pdf

Should we regulate AI?

- Buyer incentives vs. public interests
- Discrimination & profiling of people
- Liability of autonomous systems
- AI misuse for military purposes



Potential Risks from Advanced Artificial Intelligence
www.openphilanthropy.org/research/cause-reports/ai-risk

With great power comes great responsibility!