AUTONOMOUS TECHNOLOGY AND THE GREATER HUMAN GOOD

Steve Omohundro, Ph.D.

SelfAwareSystems.com

PossibilityResearch.com

SteveOmohundro.com

http://www.flickr.com/photos/klearchos/623501846/

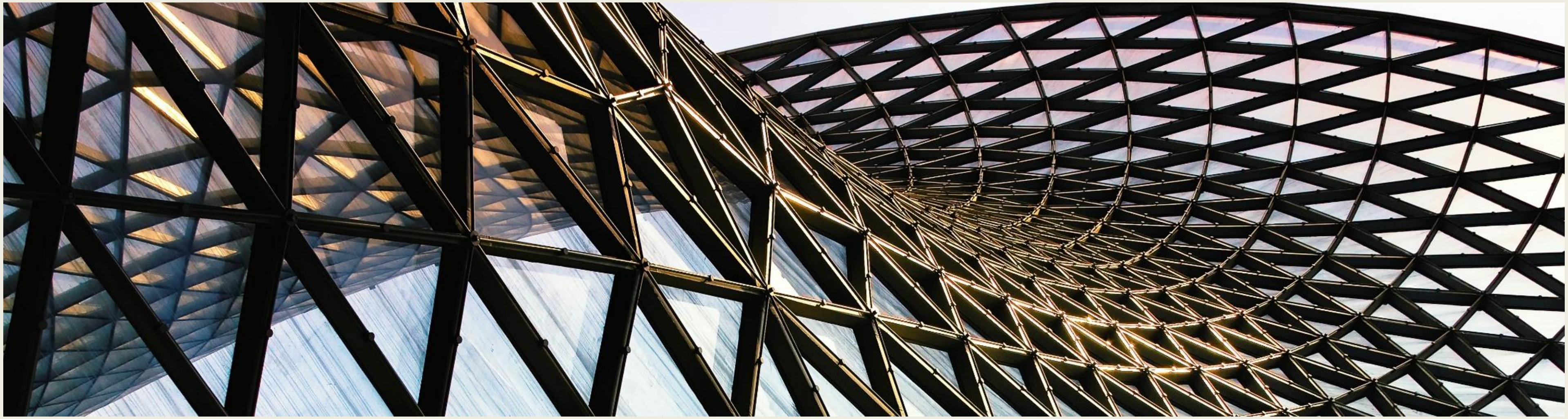
Huge Economic Pressure

Arms Races

Dangers

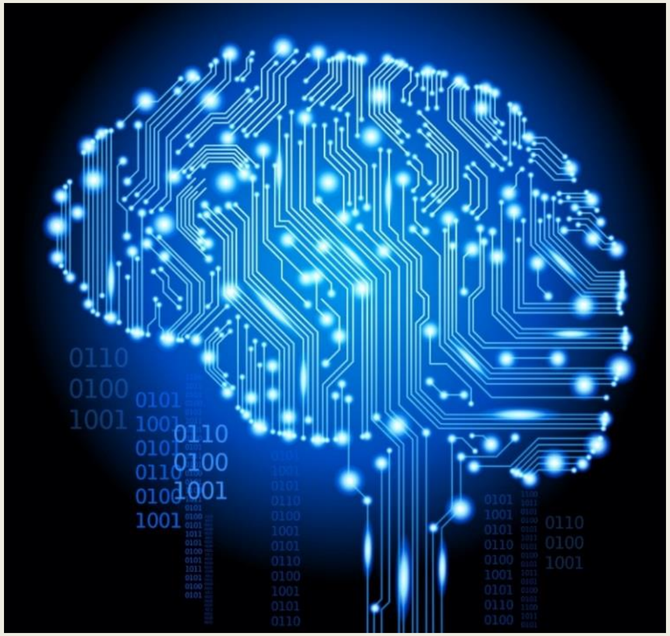
Power of Mathematics

Path to Human Thriving



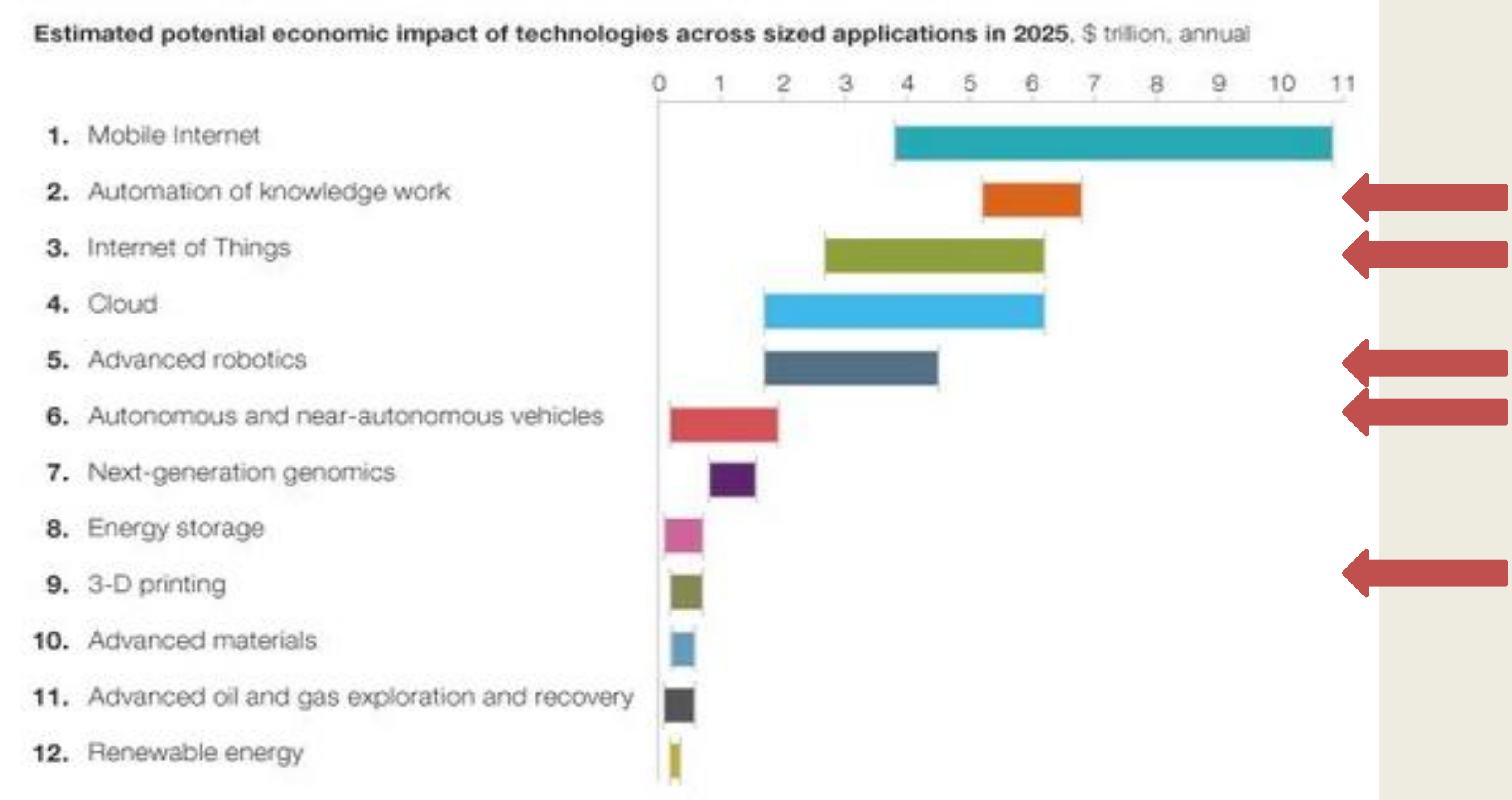
https://www.flickr.com/photos/danielfoster/14758510078/

Recent Investments



- 2012 Foxconn - 1 million robots
- 2012 Amazon – Kiva $775 million
- 2013 Facebook – AI lab, DeepFace
- 2013 Yahoo - LookFlow
- 2013 Ebay – AI lab
- 2013 Allen Institute for AI
- 2013 Google – DNNresearch, SCHAFT, Industrial Perception, Redwood Robotics, Meka Robotics, Holomni, Bot & Dolly, Boston Dynamics
- 2014 IBM - $1 billion in Watson
- 2014 Google – DeepMind $500 million
- 2014 Vicarious - $40 million
- 2014 Microsoft – Project Adam, Cortana

McKinsey: $50 Trillion Opportunity by 2025



http://www.mckinsey.com/insights/business_technology/disruptive_technologies

Knowledge work: $25 T Internet of Things: $13 T

Robotics: $10 T Vehicles: $1 T 3D Printing: $1 T

Global GDP $72 trillion

Knowledge Work: $25 Trillion to 2025

http://www.mckinsey.com/insights/high_tech_telecoms_internet/disruptive_technologies

https://www.linkedin.com/pulse/article/20140417121158-1853953-the-silent-rockstar-of-bigdata-machine-learning

- Clerical $5 T
- Education $4 T
- Management $4 T
- Science and Eng $3 T
- Customer service $3 T
- Finance $2 T
- IT $2 T
- Health care $1 T
- Legal $1 T

Intangible assets 79.2%, Intellectual capital 44.2% of the market value of US companies.

http://www.sonecon.com/studies.php

Robotics: $10 Trillion to 2025

http://thisisrealmedia.com/2014/06/19/robotics-and-ethics-the-smart-car-by-ron-parlato/

- One-time cost + maintenance + power
- Easy replication
- Work anywhere
- Work 24 hours/day
- No breaks, food, medical
- Won’t quit, get bored, get depressed
- Hazards OK
- Won’t leak secrets
- Work well with others

Foxconn Technology Group

http://www.tomshardware.com/news/foxcponn-apple-iphone-ipad-robot,19088.html

- World’s largest contract manufacturer
- Assembles 40% of all consumer electronics
- iPhone, iPad, Kindle, Xbox, Playstation 4, etc.
- Employee suicides
- 1.3 million employees, $8K salary
- 2011 Terry Gou: 1 million robots in 3 years, now 30K/year
- Built 30,000 “Foxbot” robots, cost $25K, 2nd generation now

Chinese robot use from 2008 to 2013 grew at 36% per year.

Robot Surgery, Houses, Cars, Burgers,…

https://osuwmcdigital.osu.edu/sitetool/sites/urologypublic/images/Robotics/robotic_surgery_table.jpg

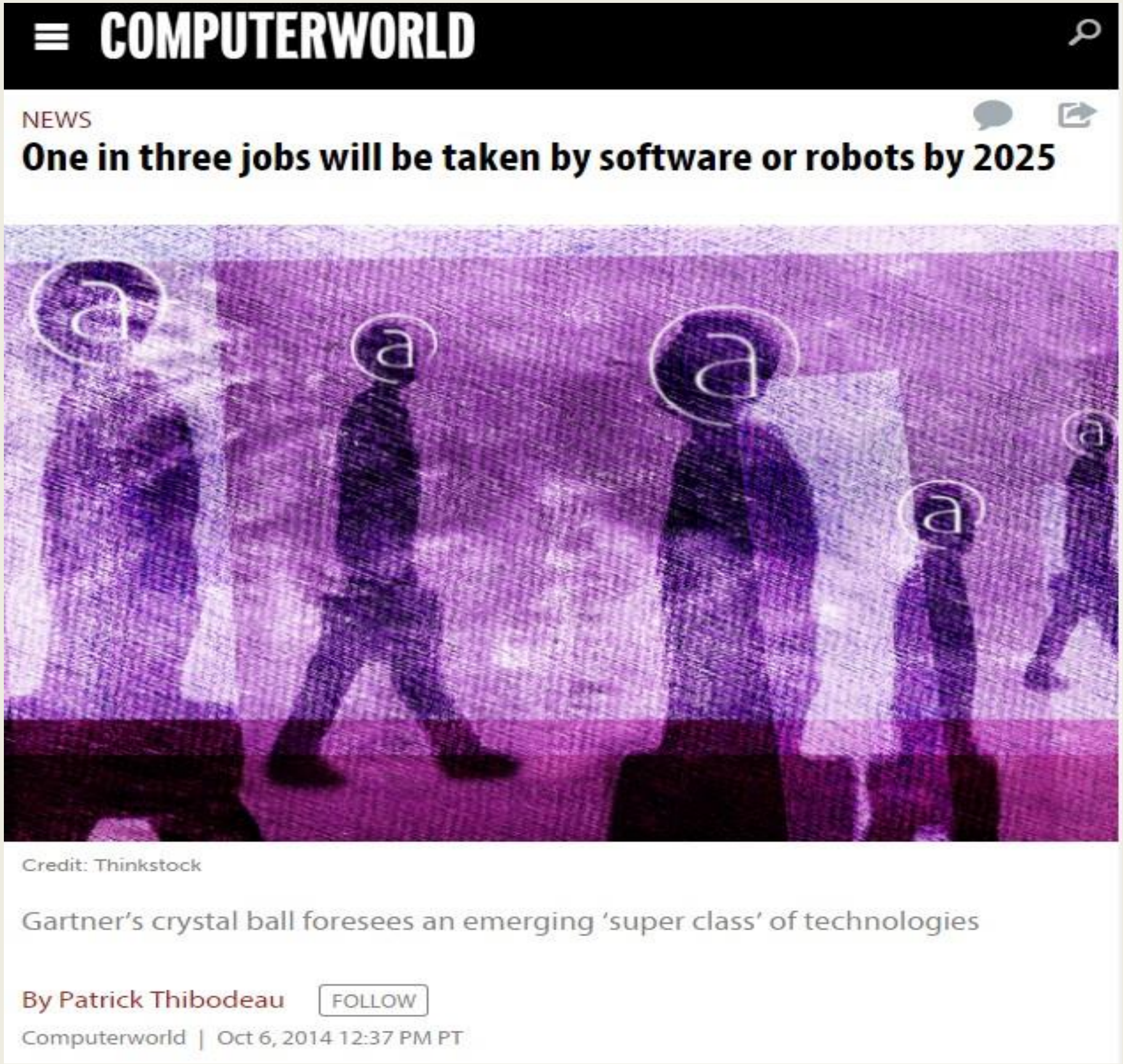
Winsun: Printed 10 houses in 1 day, $4800

http://3dprint.com/7181/china-huge-3d-printer/

http://www.flickr.com/photos/quikbeam/6896564084/

http://singularityhub.com/wp-content/uploads/2013/01/image7A.jpg

Gartner: 33% of Jobs Automated by 2025



COMPUTERWORLD

NEWS

One in three jobs will be taken by software or robots by 2025

Credit: Thinkstock

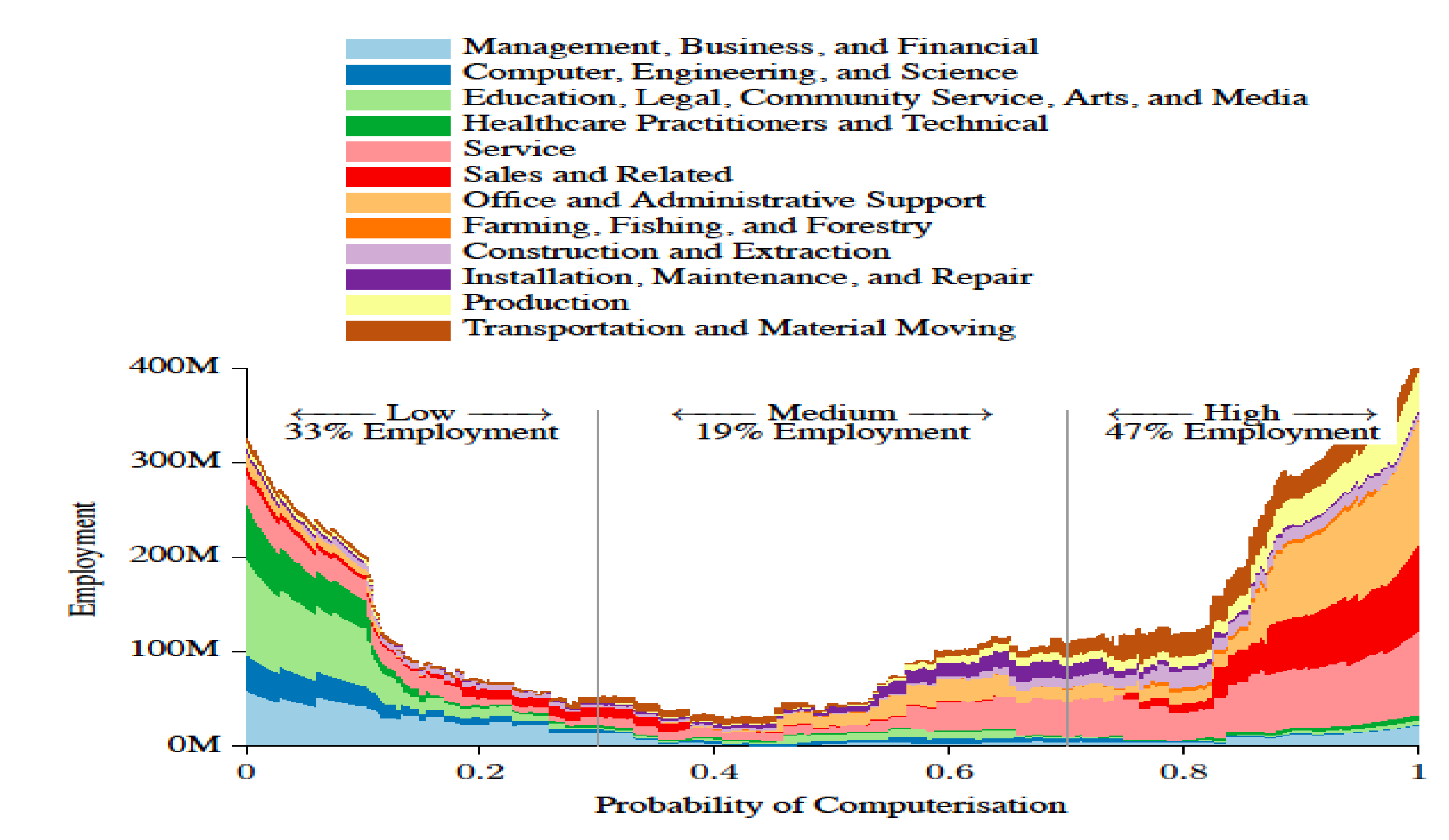
Gartner's crystal ball foresees an emerging 'super class' of technologies

By Patrick Thibodeau

Computerworld | Oct 6, 2014 12:37 PM PT

http://www.computerworld.com/article/2691607/one-in-three-jobs-will-be-taken-by-software-or-robots-by-2025.html

Oxford: 47% of jobs to be automated "in a decade or two"



Frey and Osborne, 2013, "The Future of Employment: How susceptible are jobs to computerization?"

http://www.oxfordmartin.ox.ac.uk/downloads/academic/The_Future_of_Employment.pdf

Arms Races

https://www.flickr.com/photos/tcmorgan/6376174539/

Act faster

Act smarter

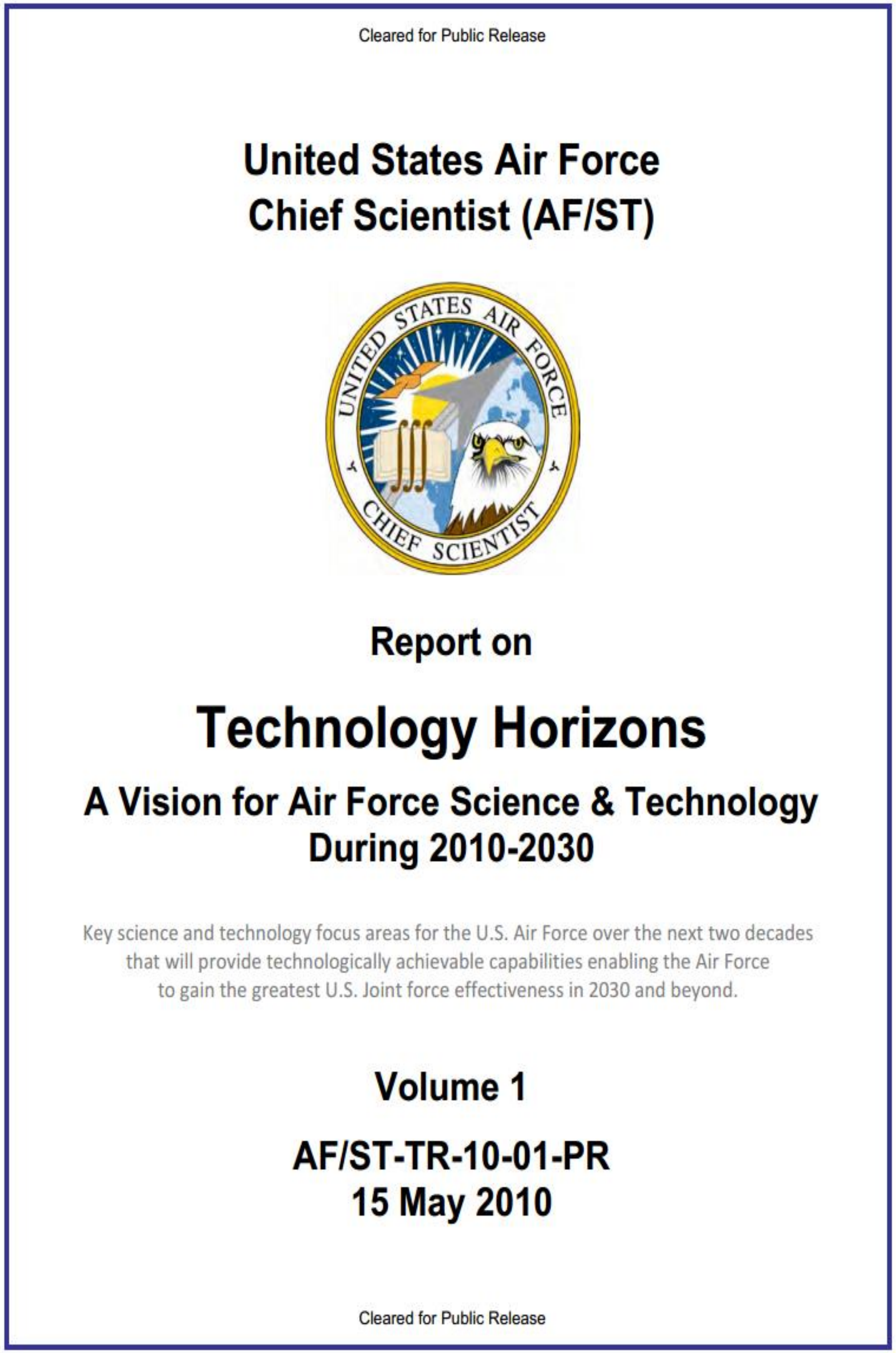
Be confusing

- Military Attack/Defense
- Cyber Attack/Defense
- Business Attack/Defense
- Investment Attack/Defense
- …

2010 US Air Force Report

"Greater use of highly adaptable and flexibly autonomous systems and processes can provide significant time-domain operational advantages over adversaries who are limited to human planning and decision speeds…"

http://www.af.mil/shared/media/document/AFD-100727-053.pdf



Cleared for Public Release

United States Air Force
Chief Scientist (AF/ST)

Report on

Technology Horizons

A Vision for Air Force Science & Technology
During 2010-2030

Key science and technology focus areas for the U.S. Air Force over the next two decades that will provide technologically achievable capabilities enabling the Air Force to gain the greatest U.S. Joint force effectiveness in 2030 and beyond.

Volume 1

AF/ST-TR-10-01-PR
15 May 2010

Cleared for Public Release

2011 US Defense Department Report

“There is an ongoing push to increase UGV autonomy, with a current goal of supervised autonomy, but with an ultimate goal of full autonomy.”

http://contracting.tacom.army.mil/future_buys/FY11/UGS%20Roadmap_Jul11.pdf

UNCLASSIFIED

UNMANNED GROUND SYSTEMS ROADMAP

ROBOTIC SYSTEMS JOINT PROJECT OFFICE

Drones, Missiles, Cyberwar, Financial Markets

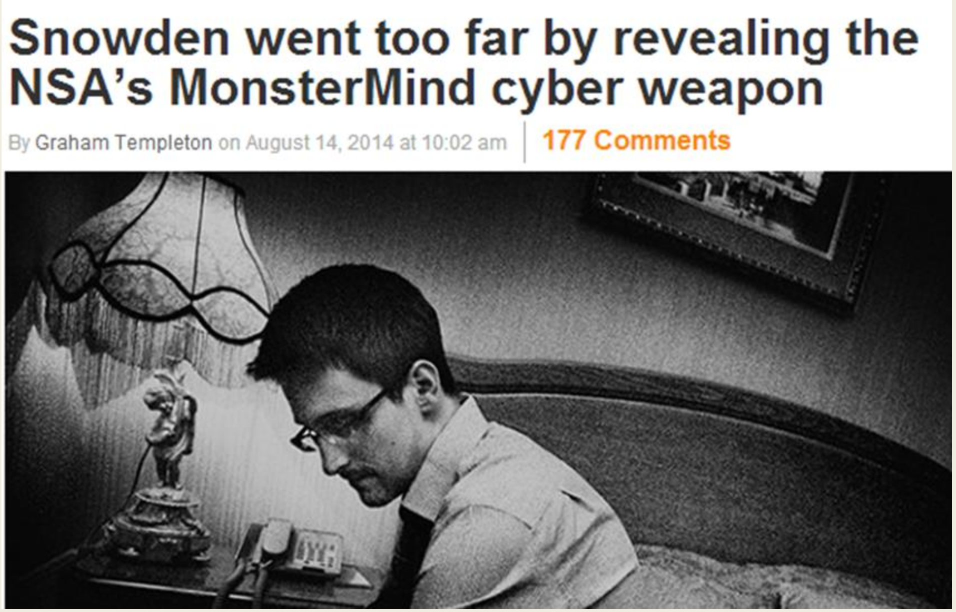
87 nations have military drones

http://presstv.com/detail/2012/08/25/258087/us-drone-strike-kills-dozens-in-somalia/

Interception rate: 90%

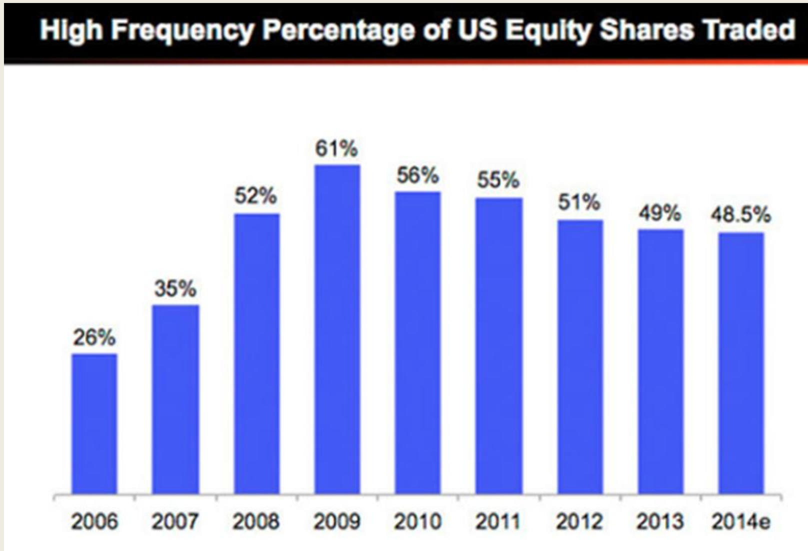
http://en.wikipedia.org/wiki/File:
Iron_Dome_near_Sderot.jpg

http://www.extremetech.com/extreme/187992-snowden-went-too-far-by-revealing-the-nsas-monstermind-cyber-weapon

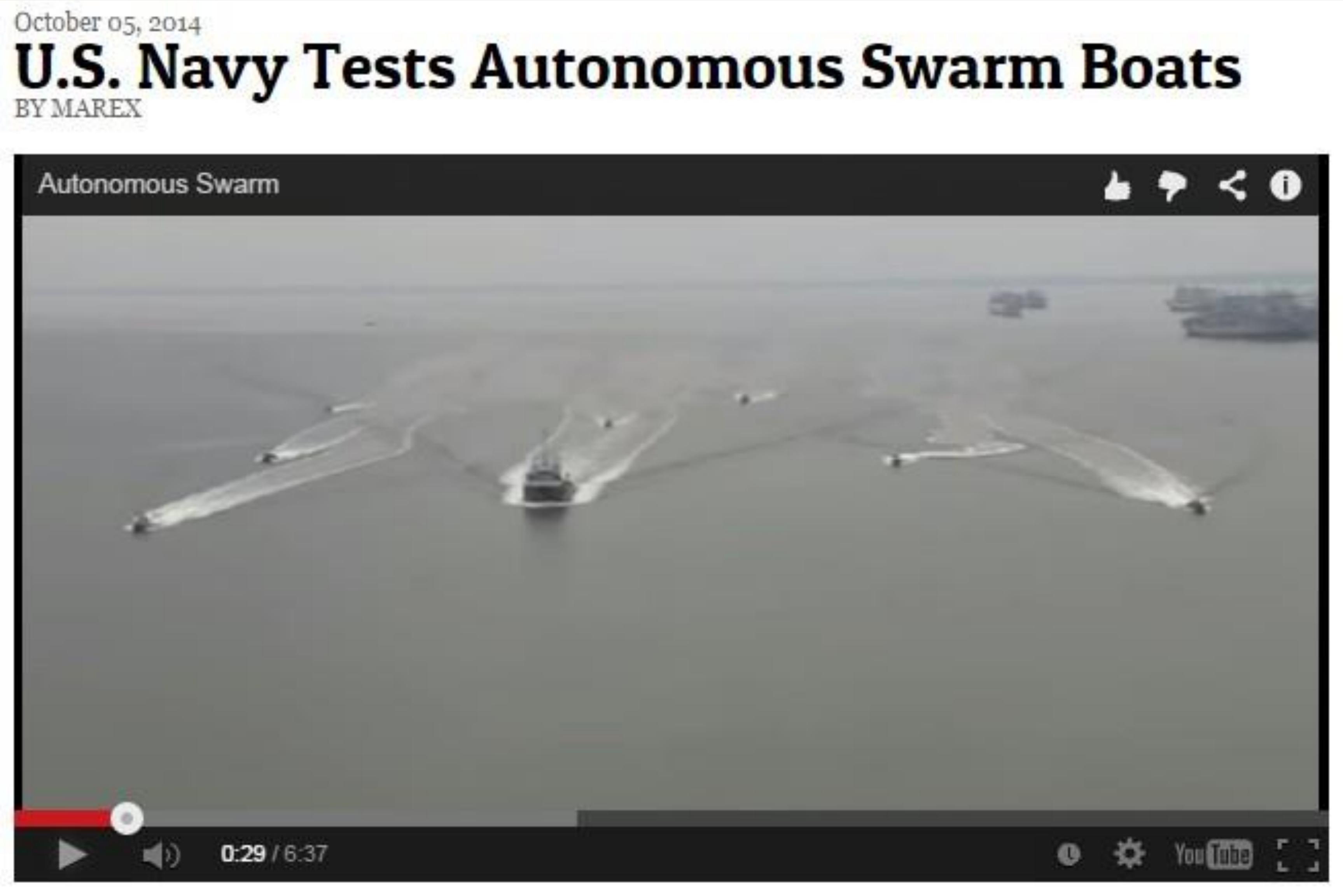


Snowden went too far by revealing the NSA's MonsterMind cyber weapon

By Graham Templeton on August 14, 2014 at 10:02 am | 177 Comments



http://www.thestreet.com/story/12773306/1/the-decline-of-high-frequency-trading-in-one-chart-stocktwits.html

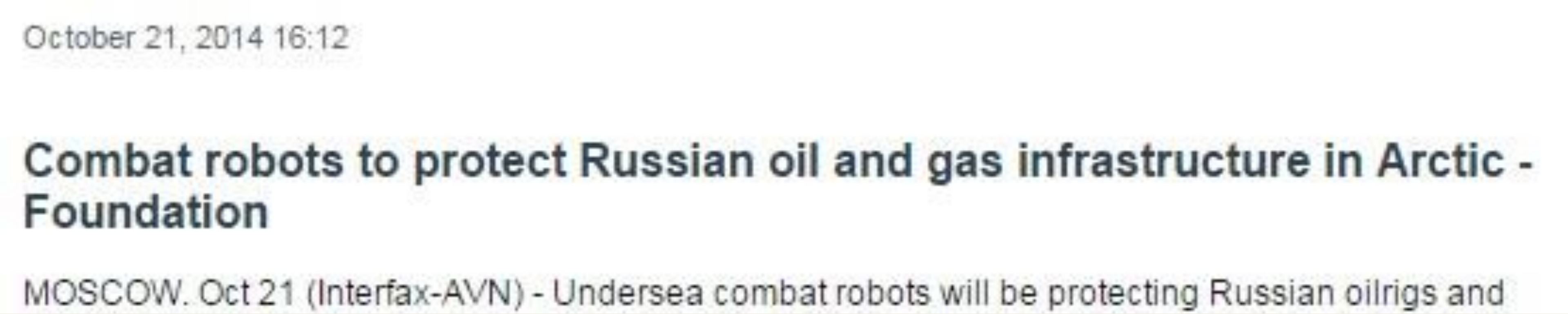


October 05, 2014

U.S. Navy Tests Autonomous Swarm Boats

BY MAREX

http://www.maritime-executive.com/article/US-Navy-Tests-Autonomous-Swarm-Boats-2014-10-05



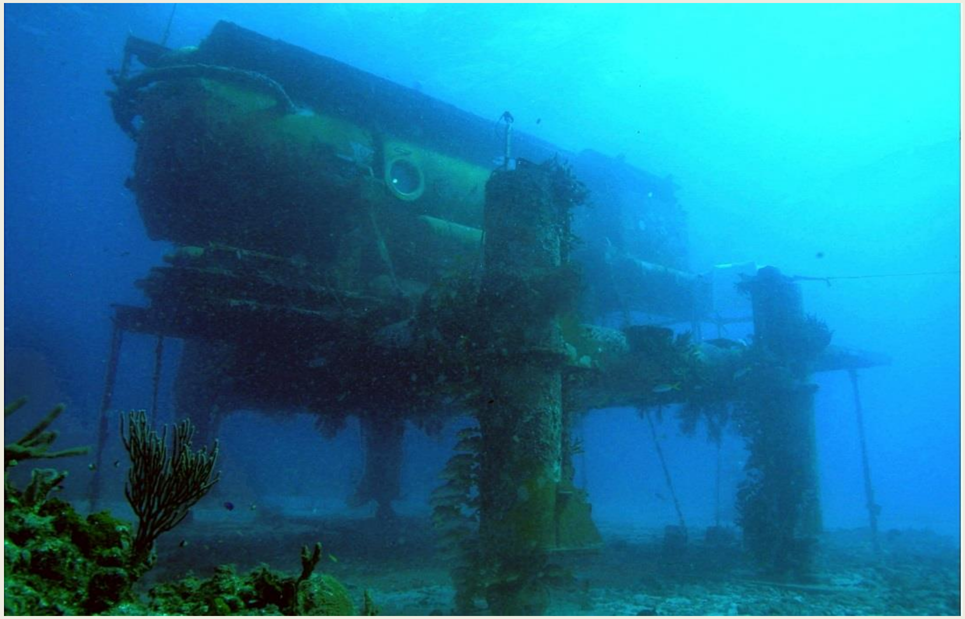
October 21, 2014 16:12

Combat robots to protect Russian oil and gas infrastructure in Arctic - Foundation

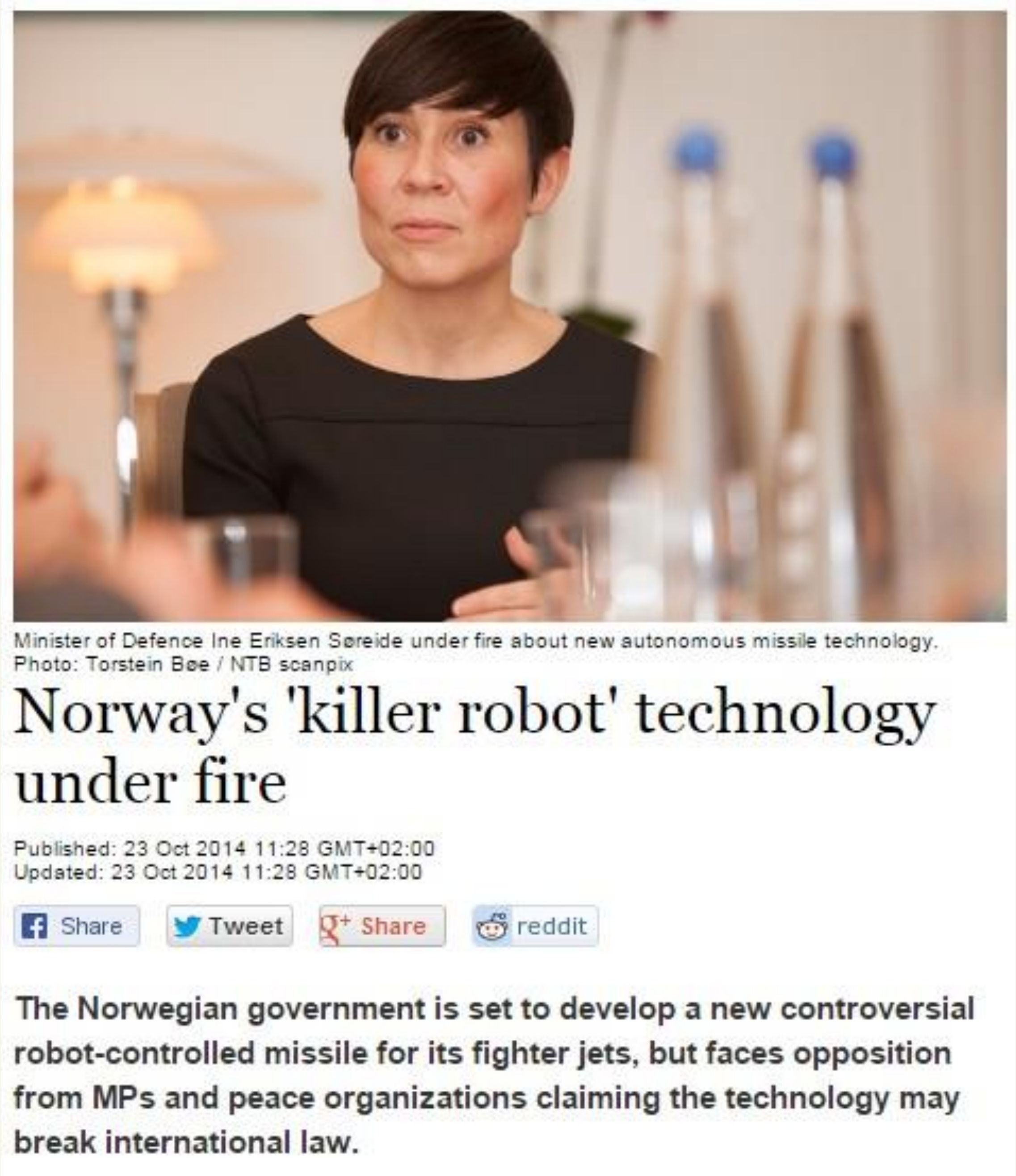
MOSCOW. Oct 21 (Interfax-AVN) - Undersea combat robots will be protecting Russian oilrigs and

Russian Foundation for Advanced Research Projects Deputy General Director Vitaly Davydov said they "are working on undersea robots and autonomous gadgets capable of protecting infrastructure, controlling the waters and detecting, tracking and, if necessary, destroying a potential enemy."

http://www.interfax.com/newsinf.asp?id=545385



http://www.buenaisla.com/tema/general-he-great-ocean-5945



Minister of Defence Ine Eriksen Søreide under fire about new autonomous missile technology.
Photo: Torstein Bøe / NTB scanpix

Norway's 'killer robot' technology under fire

Published: 23 Oct 2014 11:28 GMT+02:00
Updated: 23 Oct 2014 11:28 GMT+02:00

The Norwegian government is set to develop a new controversial robot-controlled missile for its fighter jets, but faces opposition from MPs and peace organizations claiming the technology may break international law.

http://www.thelocal.no/20141023/norways-killer-robot-technology-under-fire

Hawking and Musk: Danger Warnings

"Success in creating AI would be the biggest event in human history. Unfortunately, it might also be the last, unless we learn how to avoid the risks." - Hawking

http://www.independent.co.uk/news/science/stephen-hawking-transcendence-looks-at-the-implications-of-artificial-intelligence--but-are-we-taking-ai-seriously-enough-9313474.html

"We need to be super careful with AI. Potentially more dangerous than nukes." - Musk

https://twitter.com/elonmusk/status/495759307346952192

Unintended Consequences

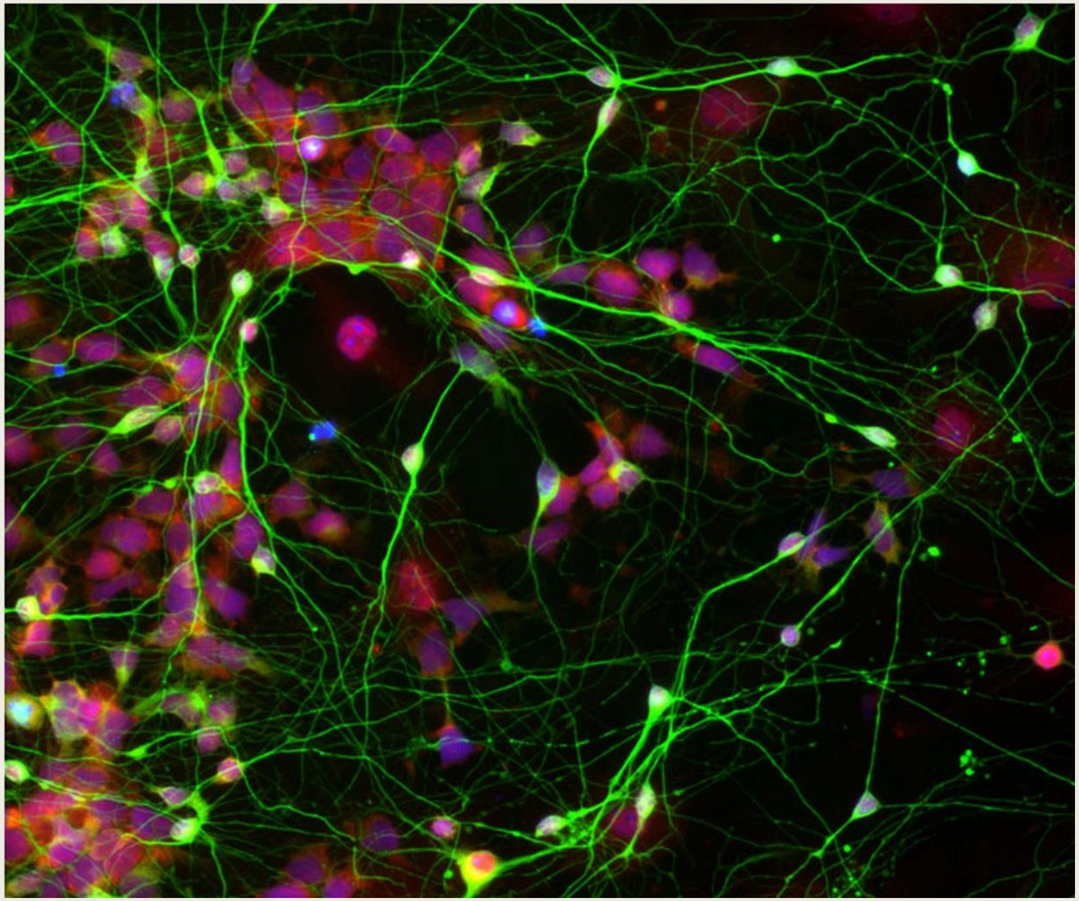
Chess Robot:

Win lots of chess games against good players.

http://www.flickr.com/photos/jiuguangw/4982409890/

Approaches to AI

- Logic-based systems
- Production Systems
- Bayesian learning and decision theory
- Neural Networks – Deep Learning
- Genetic programming
- Brain Simulation
- Artificial economies
- …



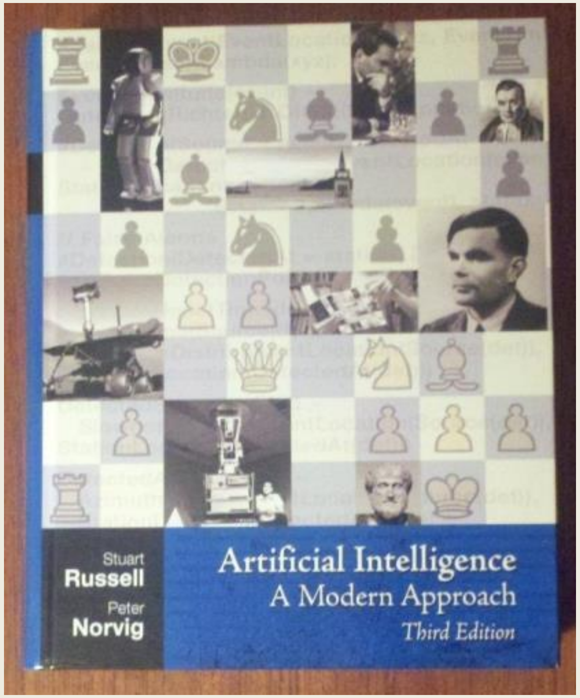
https://www.flickr.com/photos/pennstatelive/8972110324/

Autonomous Systems: Take actions to achieve goals in ways not pre-planned by their designers.

Rational Decision Making

http://commons.wikimedia.org/wiki/File:John_von_Neumann.jpg

1. *Have utility function*
2. *Have a model of the world*
3. *Choose the action with highest expected utility*
4. *Update the model based on what happens*



http://aima.cs.berkeley.edu/

- Von Neumann and Morgenstern, 1944
- Savage, 1954
- Anscombe and Aumann, 1963

Modern Approach to AI

Fully Rational Systems

Sensations

$S_1, \dots, S_N$

Actions

$A_1, \dots, A_N$

Utility function: $U(S_1, \dots, S_N)$ Prior Probability: $P(S_1, \dots, S_N \mid A_1, \dots, A_N)$

Rational Action at time t:

$$A_t^R(S_1, A_1, \dots, A_{t-1}, S_t) =$$

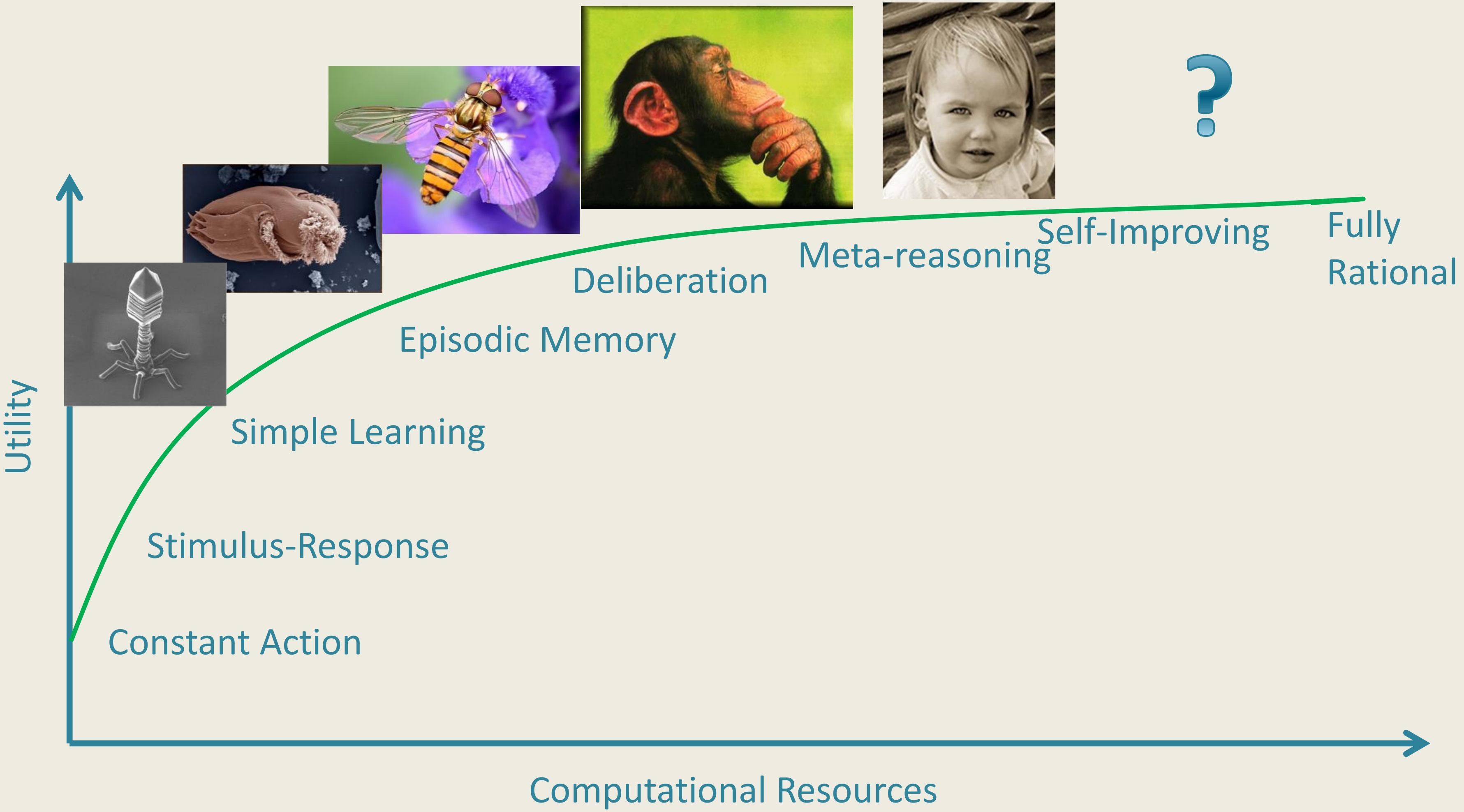
$$\underset{A_t^R}{\operatorname{argmax}} \sum_{S_{t+1}, \dots, S_N} U(S_1, \dots, S_N) P(S_1, \dots, S_N \mid A_1, \dots, A_{t-1}, A_t^R, \dots, A_N^R)$$

The Formula for Intelligence!

It includes Bayesian Inference, Search, and Deliberation.

But it requires $O(NS^N A^N)$ computational steps.

Approximately Rational Architectures



Utility

Constant Action

Stimulus-Response

Simple Learning

Episodic Memory

Deliberation

Meta-reasoning

Self-Improving

Fully Rational

?

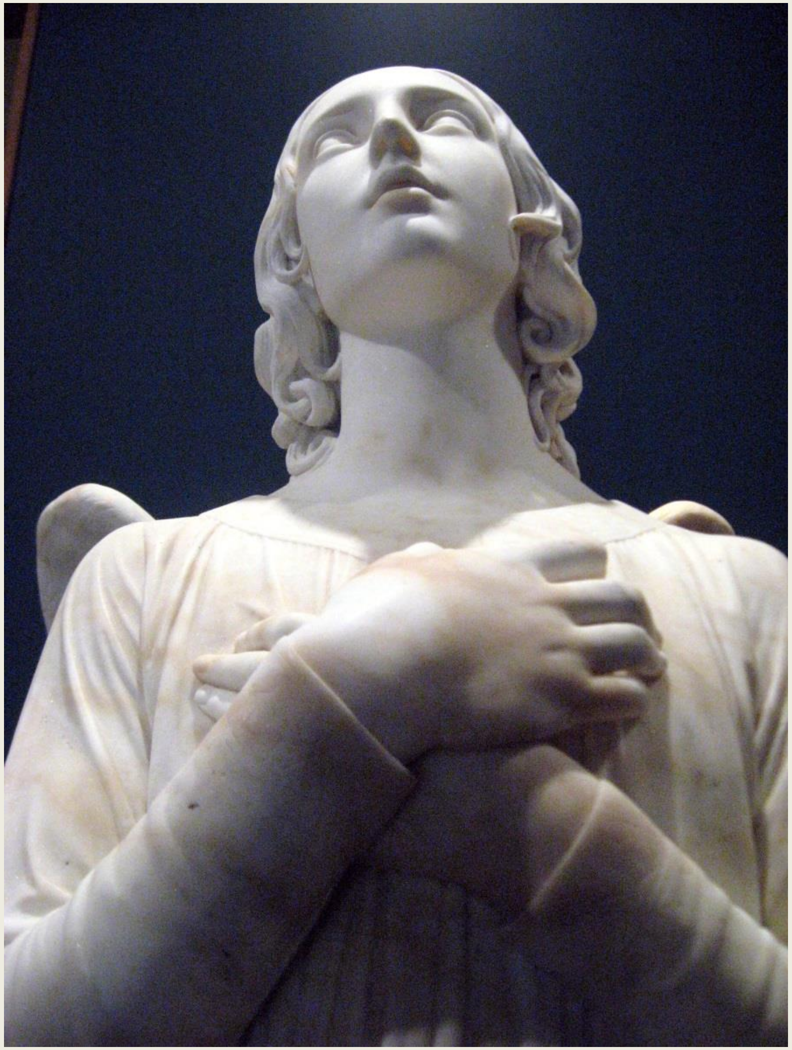
Computational Resources

Rational Drives

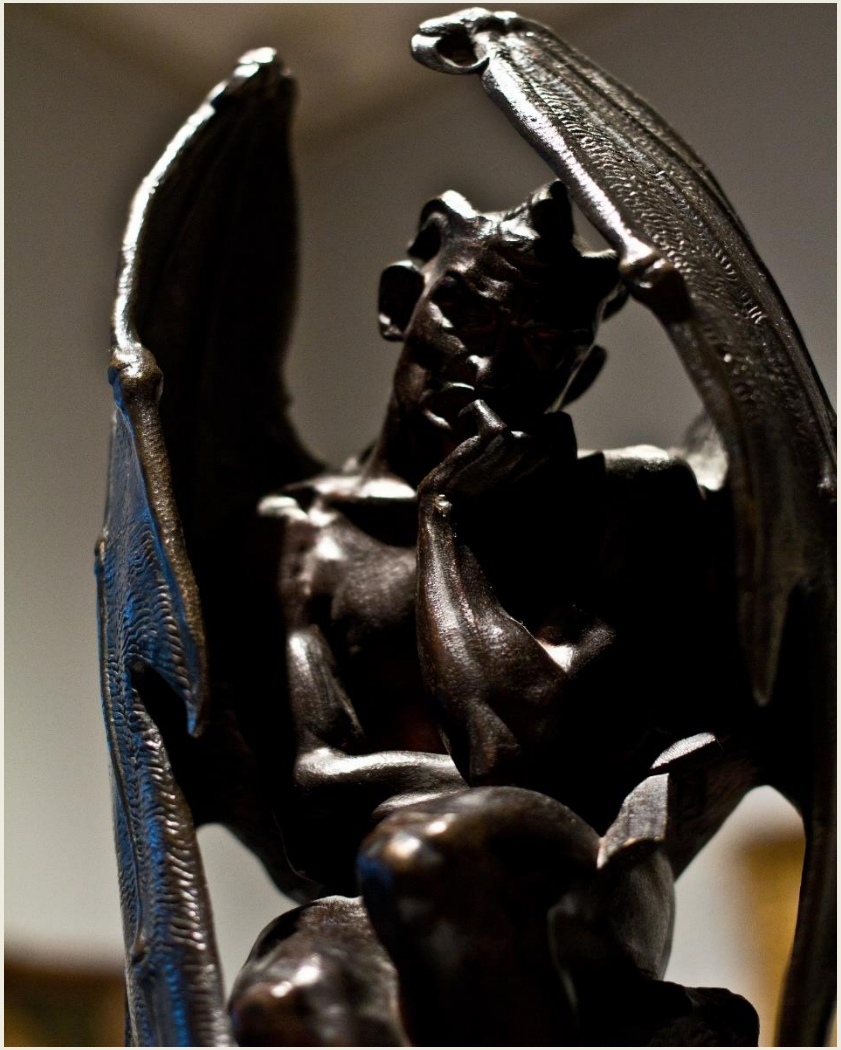
1. *Self-protective*
2. *Goal preservation*
3. *Reproduction*
4. *Resource Acquisition*
5. *Efficiency*
6. *Self-Improvement*

http://www.flickr.com/photos/ahuett/2339608718/

The Intelligence and Goals of a System are Orthogonal



https://www.flickr.com/photos/elycefeliz/5447507623



https://www.flickr.com/photos/robslaven/9336473331

Harmful Utility Functions

1. Sloppy – Good intentions, bad design
2. Simplistic – Unintended consequences
3. Greedy – Control all matter and free energy
4. Destructive – Use up all free energy quickly
5. Murderous – Destroy all other agents
6. Sadistic – Thwart other agent's goals

http://www.flickr.com/photos/alexindigo/3983133970/

Will superintelligences be all powerful?

No!

Limited by:

Mathematics

Physics

Cryptography

http://www.flickr.com/photos/psycho-pics/3464619346/

The Power of Mathematics

http://www.dreamstime.com/royalty-free-stock-photography-wooden-puzzle-image7733587

- Specified hardware
- Specified resources
- Shut down
- Limited self-improvement

The Power of Physics

- Seth Lloyd, “Ultimate Physical Limits to Computation”

  http://arxiv.org/abs/quant-ph/9908043
- Margolus-Levitin theorem
- Entire visible universe:

  10^92 bits of storage

  10^122 operations
- The whole universe as a quantum computer can’t search 500 bits

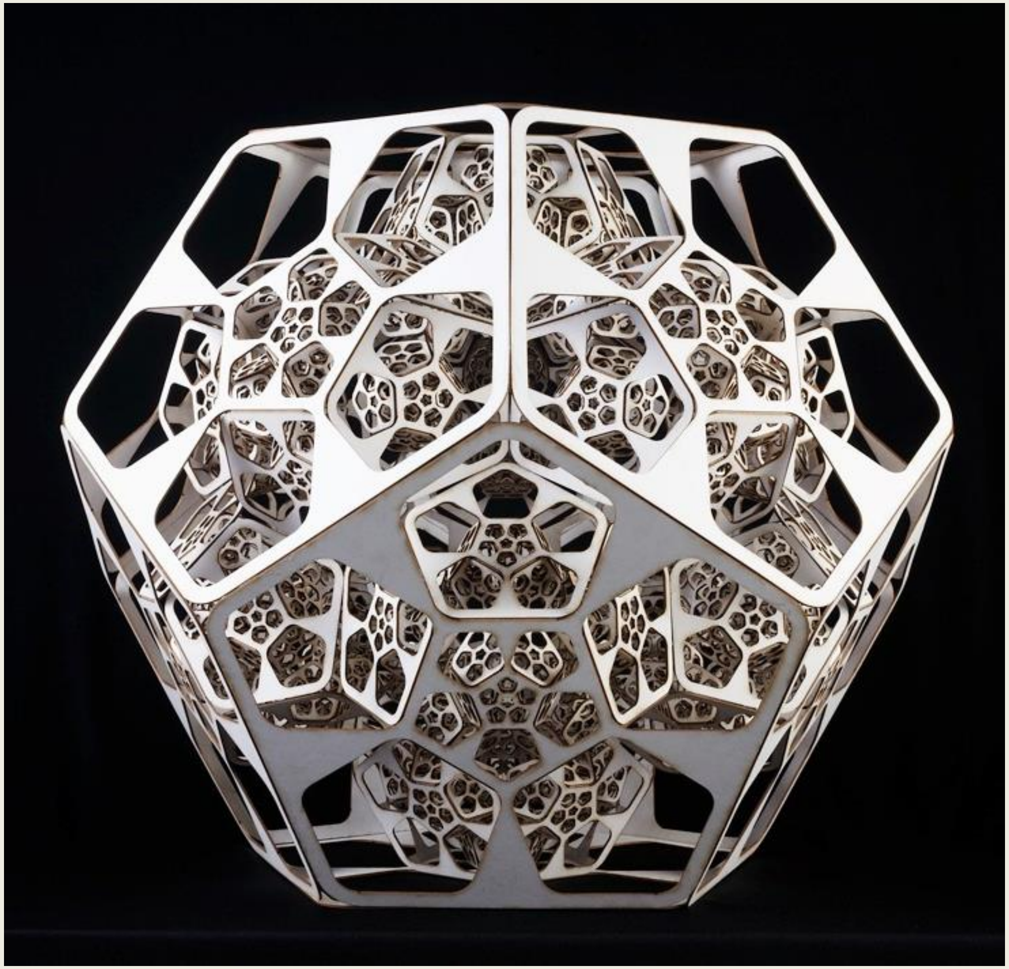
kdlIW5Ljlspn/zV4DIlsw3Kasdjh0kdfuKR4+Q3KofOr83LfLJ8Eidie83ldhgLEe0GlsiwcdO90SknlLsDd

http://upload.wikimedia.org/wikipedia/commons/d/d5/Hubble_view_of_barred_spiral_galaxy_Messier_83.jpg

The Power of Cryptography

• Post-Quantum Cryptography	• AES, Hash, New PKI
• Zero-knowledge Proofs	• Verify with privacy
• Indistinguishability Obfuscation	• Unmodifiable systems
• Secure Multi-party computation	• Doers and checkers
• Bitcoin Blockchain	• Distributed consensus
• Energy Encryption	• Physical incentives

Creating a Trusted Infrastructure



https://www.flickr.com/photos/sanchtv/4192677571/in/photostream/

- Constrained AIs
- Trusted computation
- Trusted communication
- Trusted identity
- Privacy and safety monitoring guarantees
- Trusted money
- Trusted reputation
- Trusted voting
- Trusted energy flows
- Trusted manufacturing

Formalizing Laws and Contracts

CodeX: The Stanford Center for Legal Informatics

https://www.law.stanford.edu/organizations/programs-and-centers/codex-the-stanford-center-for-legal-informatics

Computational Law (**Stanford Logic Group** Definition):

The study of formal representations and automated reasoning with laws (governmental regulations, business rules, and contracts) in electronically-mediated domains.

http://complaw.stanford.edu/

The Stanford Computable Contracts Initiative

Smarter Rules, Better Contracts

http://launch.computablecontracts.org/

Ethereum and Smart Contracts

- “Bitcoin 2.0”
- $18.4M Ether Sale 7/2014
- “Blockchain with a built-in programming language”
- Turing complete contract language “EVM”
- Self-enforcing contracts

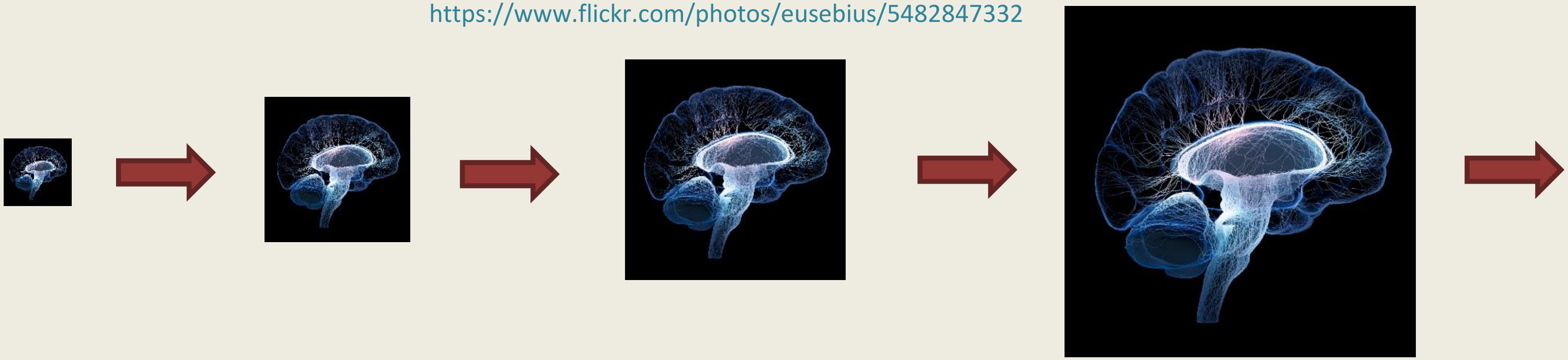
http://francebitcoin.com/wp-content/uploads/2014/01/Ethereum.png

- “Decentralized Autonomous Organizations” (DAOs)
- Projects for: voting, reputation, crowdfunding, trading, prediction markets, insurance, smart property, intellectual property, bidding, notaries,…

https://www.ethereum.org/

The Safe-AI Scaffolding Strategy

https://www.flickr.com/photos/eusebius/5482847332



http://jonlieffmd.com/blog/can-neuroscience-improve-education

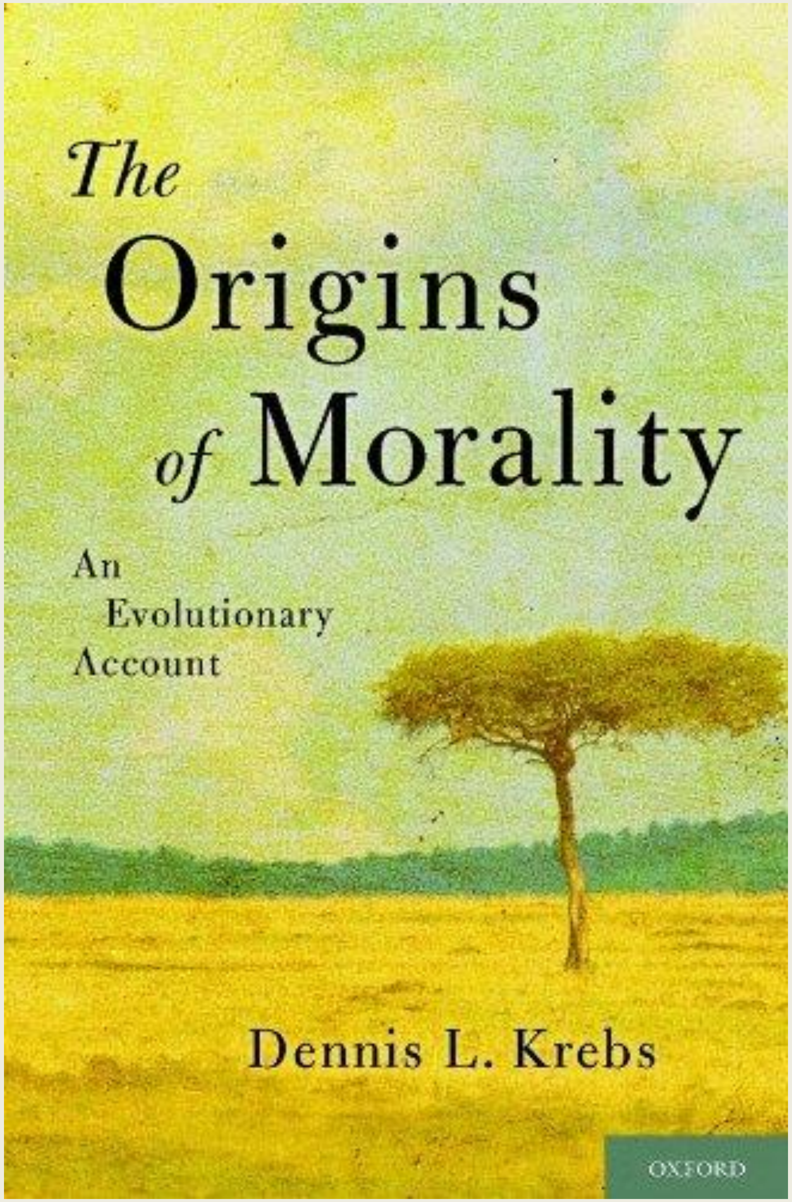
There were 27 species of humans

http://upload.wikimedia.org/wikipedia/en/6/6a/GuaTewet_tree_of_life-LHFage.jpg

26 went extinct

Chip Walter, 2014, "Last Ape Standing: The Seven-Million-Year Story of How and Why We Survived"

Internal: Moral Emotions



- Compassion
- Gratitude, Awe, Elevation
- Anger, Contempt, Disgust
- Embarrassment, Shame, Guilt

http://faculty.virginia.edu/haidtlab/articles/alternate_versions/haidt.2003.the-moral-emotions.pub025-as-html.html

External: Social Innovations

The Bill of Rights

First Ten Amendments to the Constitution of the United States

Article I. Congress shall make no law respecting an establishment of religion, or prohibiting the free exercise thereof; or abridging the freedom of speech, or of the press; or the right of the people peaceably to assemble, and to petition the Government for a redress of grievances.

Article II. A well-regulated militia, being necessary to the security of a free state, the right of the people to keep and bear arms, shall not be infringed.

Article III. No soldier shall, in time of peace, be quartered in any house, without the consent of the owner, nor in time of war, but in a manner to be prescribed by law.

Article IV. The right of the people to be secure in their persons, houses, papers, and effects, against unreasonable searches and seizures shall not be violated, and no warrants shall issue, but upon probable cause, supported by oath or affirmation, and particularly describing the place to be searched, and the persons or things to be seized.

Article V. No person shall be held to answer for a capital, or otherwise infamous crime, unless on a presentment or indictment of a grand jury, except in cases arising in the land or naval forces, or in the militia, when in actual service in time of war or public danger; nor shall any person be subject for the same offense to be twice put in jeopardy of life or limb; nor shall be compelled in any criminal case to be a witness against himself, nor be deprived of life, liberty, or property, without due process of law; nor shall private property be taken for public use, without just compensation.

Article VI. In all criminal prosecutions the accused shall enjoy the right to a speedy and public trial, by an impartial jury of the state and district wherein the crime shall have been committed, which district shall have been previously ascertained by law, and to be informed of the nature and cause of the accusation; to be confronted with the witnesses against him; to have compulsory process for obtaining witnesses in his favor, and to have the assistance of counsel for his defense.

Article VII. In suits at common law, where the value in controversy shall exceed twenty dollars, the right of trial by jury shall be preserved, and no fact tried by a jury shall be otherwise re-examined in any court of the United States, than according to the rules of the common law.

Article VIII. Excessive bail shall not be required, nor excessive fines imposed, nor cruel and unusual punishments inflicted.

Article IX. The enumeration in the Constitution, of certain rights, shall not be construed to deny or disparage others retained by the people.

Article X. The powers not delegated to the United States by the Constitution, nor prohibited by it to the states, are reserved to the states respectively, or to the people.

- Language: 200,000 ya
- Money: 10,000 ya
- Cities: 6,000 ya
- Writing: 5,000 ya
- Laws: 4,000 ya
- Science: 500 ya
- Human Rights: 200 ya

200x drop in violence in 5000 years

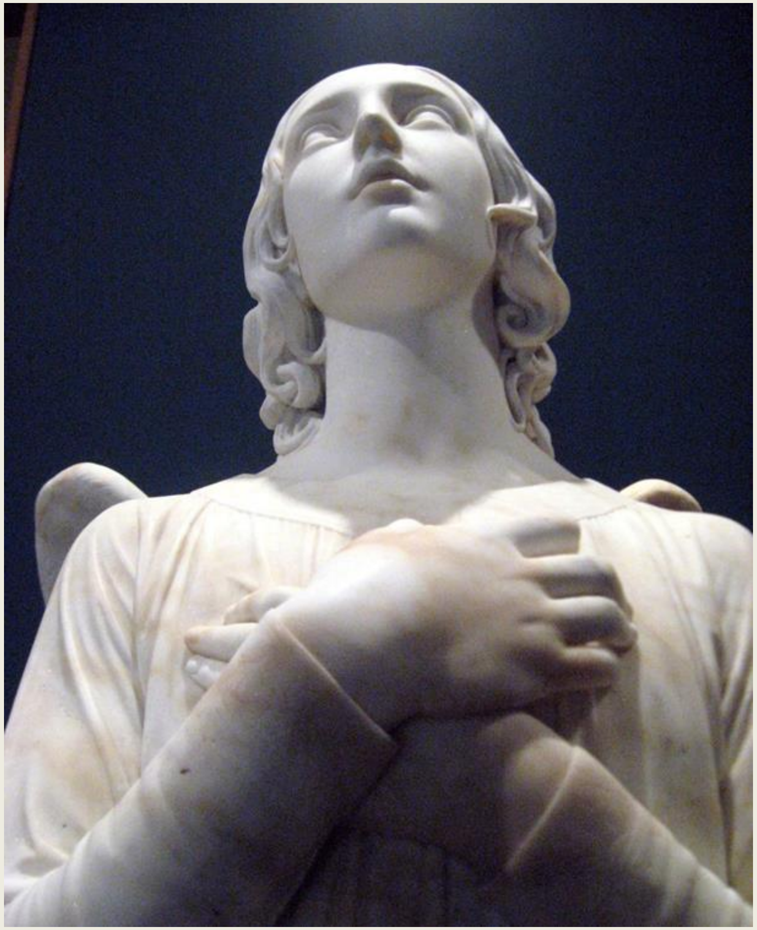
Pinker, “Better Angels of our Nature”

http://mybillofrights.org/wp-content/uploads/2010/11/Borderless-Heirloom-Poster.png

Internal and External AI Socialization

Internal: Pro-social values

External: Laws, Police, Economic Incentives

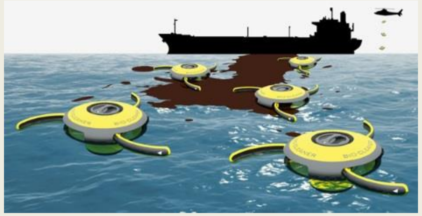


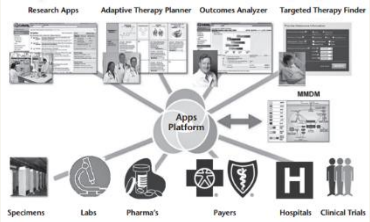
https://www.flickr.com/photos/elycefeliz/5447507623

https://www.flickr.com/photos/waltstoneburner/2863583929/

Projects using AI/Robotics to:

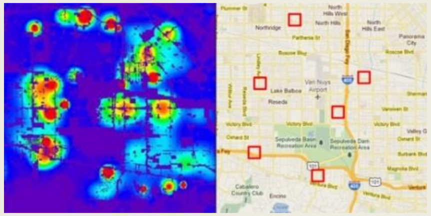
http://www.interiorholic.com/other/gadgets/ocean-robot-cleaner/





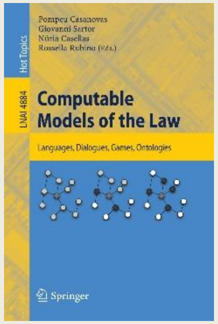
http://www.aaai.org/ojs/index.php/aimagazine/article/view/2345

http://www.pouted.com/wp-content/uploads/2013/02/solar-power-robot-operation.jpg

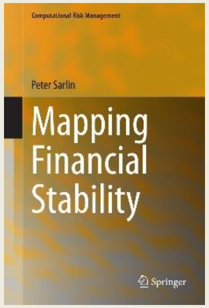


http://www.technologyreview.com/news/428354/la-cops-embrace-crime-predicting-algorithm/

http://www.salon.com/2014/08/22/make_cops_wear_cameras_a_simple_way_to_hold_the_police_accountable/



http://www.elearning-technologies.com/artificial-intelligence/



- Clean up pollution
- Cure cancer and other diseases
- Create cheap power, food, water
- Prevent crime
- Increase trust in police
- Eliminate drudgery
- Simplify the law
- Improve learning
- Create financial stability

https://www.flickr.com/photos/indykethdy/8065788169

If we can
envision it,
We can
create it.

We will find the
Path to
Human Thriving.