Intro
DQN In Theory
DQN In Practice
Beyond DQN
AI Revolution

Deep Reinforcement Learning

Deep Learning — Unit 10

Dr. Jon Krohn
jon@untapt.com

Slides available at jonkrohn.com/talks

April 7th, 2018

NYC DATA SCIENCE ACADEMY

Intro
DQN In Theory
DQN In Practice
Beyond DQN
AI Revolution

Outline

1 Introduction
2 Deep Q-Learning Theory
3 Cart Pole Deep Q-Learning
4 Beyond DQN
5 The AI Revolution

Intro
DQN In Theory
DQN In Practice
Beyond DQN
AI Revolution

Outline

1. Introduction
2. Deep Q-Learning Theory
3. Cart Pole Deep Q-Learning
4. Beyond DQN
5. The AI Revolution

Intro
DQN In Theory
DQN In Practice
Beyond DQN
AI Revolution

Outline

1. Introduction
2. Deep Q-Learning Theory
3. Cart Pole Deep Q-Learning
4. Beyond DQN
5. The AI Revolution

Intro
DQN In Theory
DQN In Practice
Beyond DQN
AI Revolution

Outline

1. Introduction
2. Deep Q-Learning Theory
3. Cart Pole Deep Q-Learning
4. Beyond DQN
5. The AI Revolution

Intro
DQN In Theory
DQN In Practice
Beyond DQN
AI Revolution

Outline

1. Introduction
2. Deep Q-Learning Theory
3. Cart Pole Deep Q-Learning
4. Beyond DQN
5. The AI Revolution

Intro
DQN In Theory
DQN In Practice
Beyond DQN
AI Revolution

Outline

1 Introduction
2 Deep Q-Learning Theory
3 Cart Pole Deep Q-Learning
4 Beyond DQN
5 The AI Revolution

NYC DATA SCIENCE ACADEMY

Categories of A.I.

1. ANI: *Narrow*
2. AGI: *General*
3. ASI: *Super*

Categories of A.I.

1. ANI: *Narrow*
2. AGI: *General*
3. ASI: *Super*

Categories of A.I.

1. ANI: *Narrow*
2. AGI: *General*
3. ASI: *Super*

Deep Learning Categories

1. supervised learning
2. unsupervised learning
3. reinforcement learning

Deep Learning Categories

1. supervised learning
2. unsupervised learning
3. reinforcement learning

Deep Learning Categories

1. supervised learning
2. unsupervised learning
3. reinforcement learning

Supervised Learning

- have x and y
- goal is to approximate function that predicts y with x
- e.g.: classification (MNIST digits, IMDB sentiment), regression (sales of a product, future value of an asset)

Supervised Learning

- have x and y
- goal is to approximate function that predicts y with x
- e.g.: classification (MNIST digits, IMDB sentiment), regression (sales of a product, future value of an asset)

Supervised Learning

- have x and y
- goal is to approximate function that predicts y with x
- e.g.: classification (MNIST digits, IMDB sentiment), regression (sales of a product, future value of an asset)

Unsupervised Learning

- have *x* without *y* labels
- goal is to learn data's underlying hidden structure
- e.g.: word vectors, GANs

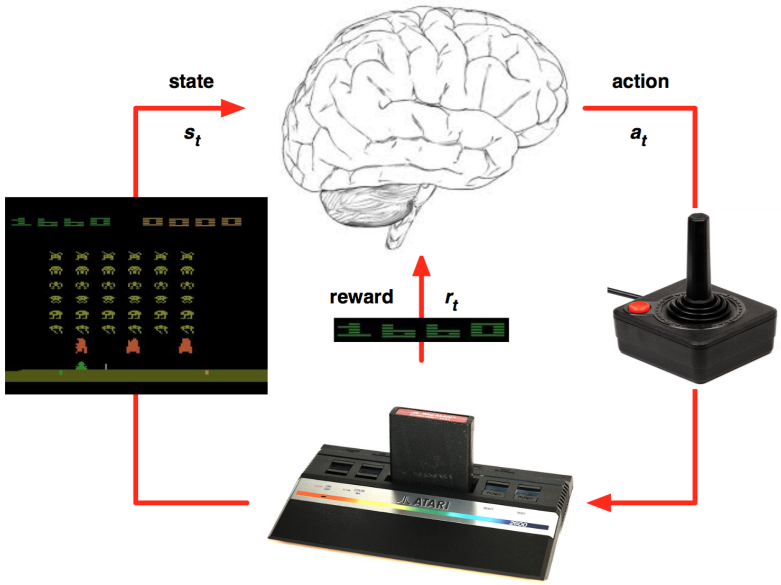
Unsupervised Learning

- have *x* without *y* labels
- goal is to learn data's underlying hidden structure
- e.g.: word vectors, GANs

Unsupervised Learning

- have *x* without *y* labels
- goal is to learn data's underlying hidden structure
- e.g.: word vectors, GANs

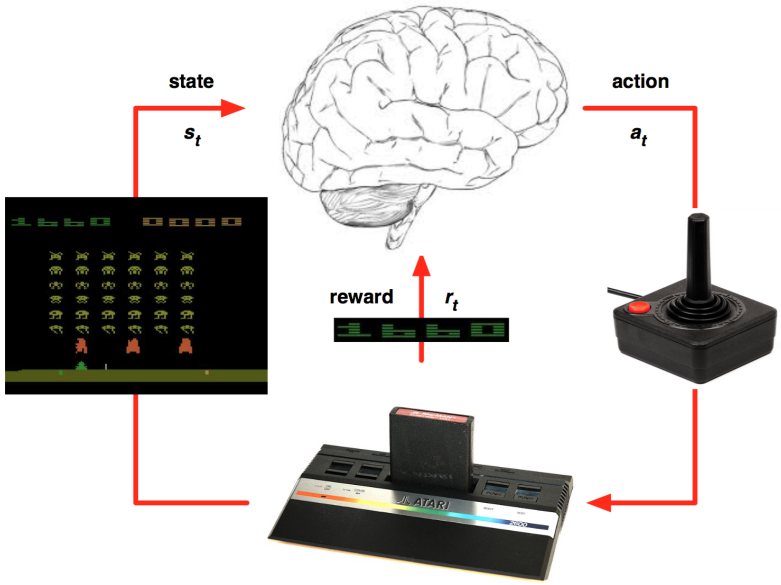
Reinforcement Learning



- an *agent* takes action in an *environment*
- environment returns *state*, i.e., *reward* at t, state at $t+1$
- repeat

NYC DATA SCIENCE ACADEMY

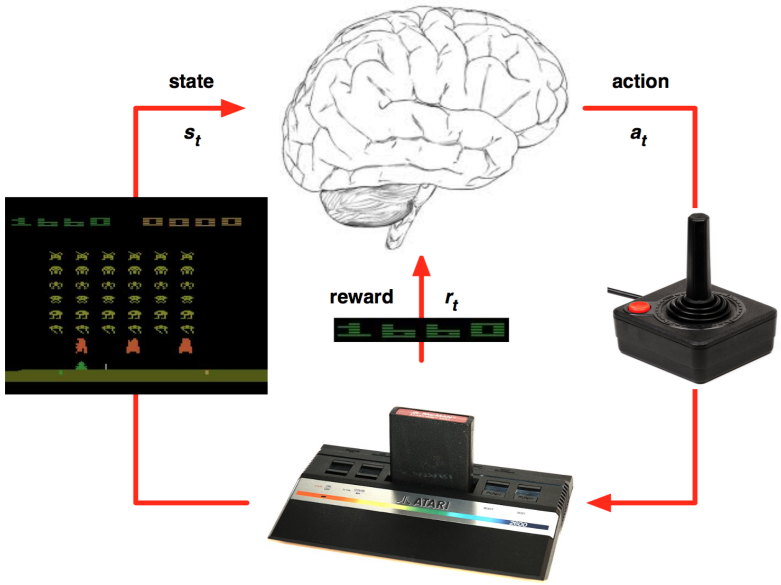
Reinforcement Learning



- an *agent* takes action in an *environment*
- environment returns *state*, i.e., *reward* at t, state at $t+1$
- repeat

NYC DATA SCIENCE ACADEMY

Reinforcement Learning



- an *agent* takes action in an *environment*
- environment returns *state*, i.e., *reward* at t, state at $t+1$
- repeat

Deep Q-Learning

Mnih et al. (2013)

- “Playing Atari with Deep RL”
- NIPS and arXiv paper
- played seven Atari 2600 games
- beat previous ML approaches on six
- beat human expert on three

Deep Q-Learning

Mnih et al. (2013)

- “Playing Atari with Deep RL”
- NIPS and arXiv paper
- played seven Atari 2600 games
- beat previous ML approaches on six
- beat human expert on three

Deep Q-Learning

Mnih et al. (2013)

- “Playing Atari with Deep RL”
- NIPS and arXiv paper
- played seven Atari 2600 games
- beat previous ML approaches on six
- beat human expert on three

Deep Q-Learning

Mnih et al. (2013)

- “Playing Atari with Deep RL”
- NIPS and arXiv paper
- played seven Atari 2600 games
- beat previous ML approaches on six
- beat human expert on three

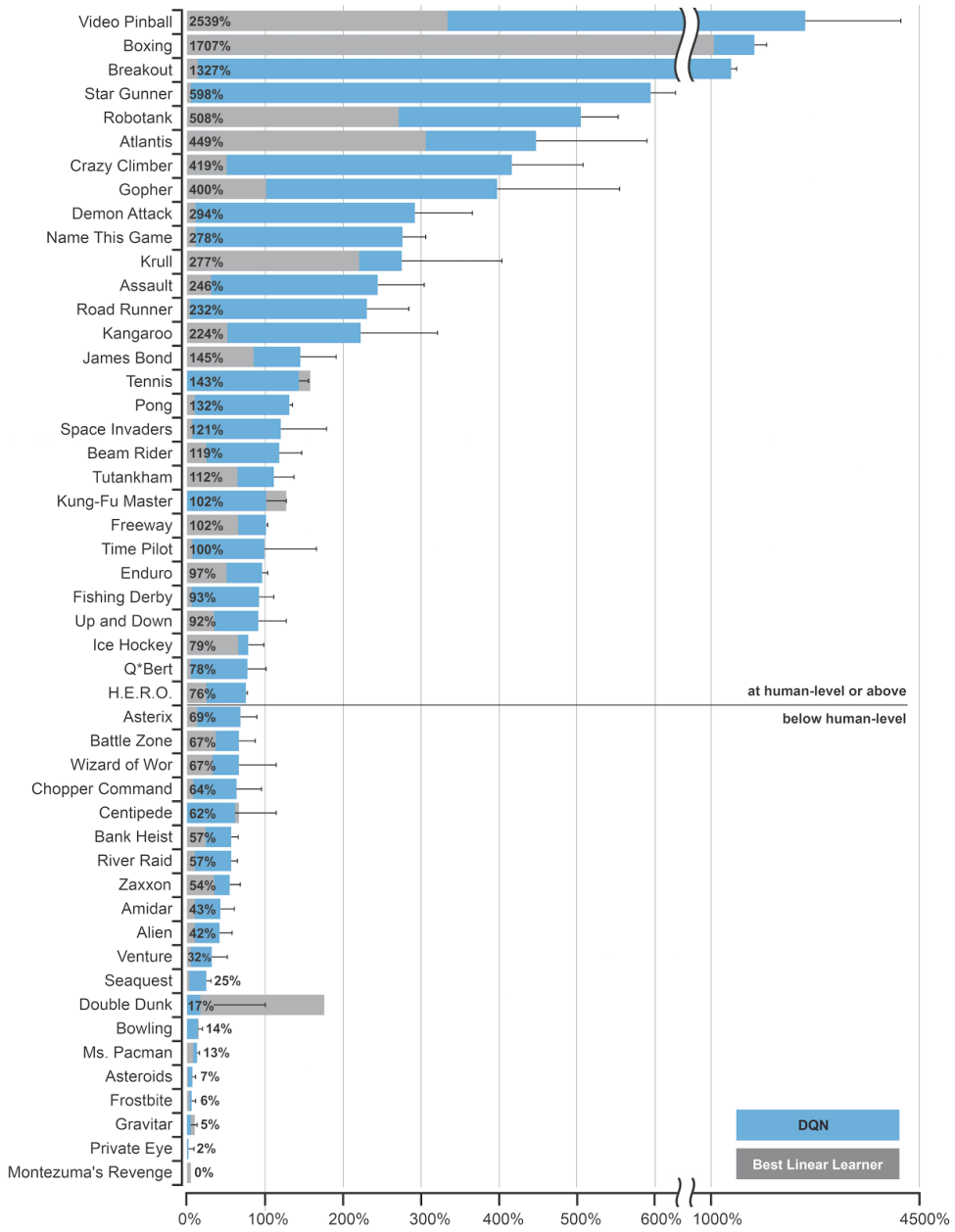
Deep Q-Learning

Mnih et al. (2013)

- “Playing Atari with Deep RL”
- NIPS and arXiv paper
- played seven Atari 2600 games
- beat previous ML approaches on six
- beat human expert on three

Deep Q-Learning

Mnih et al. (2015)

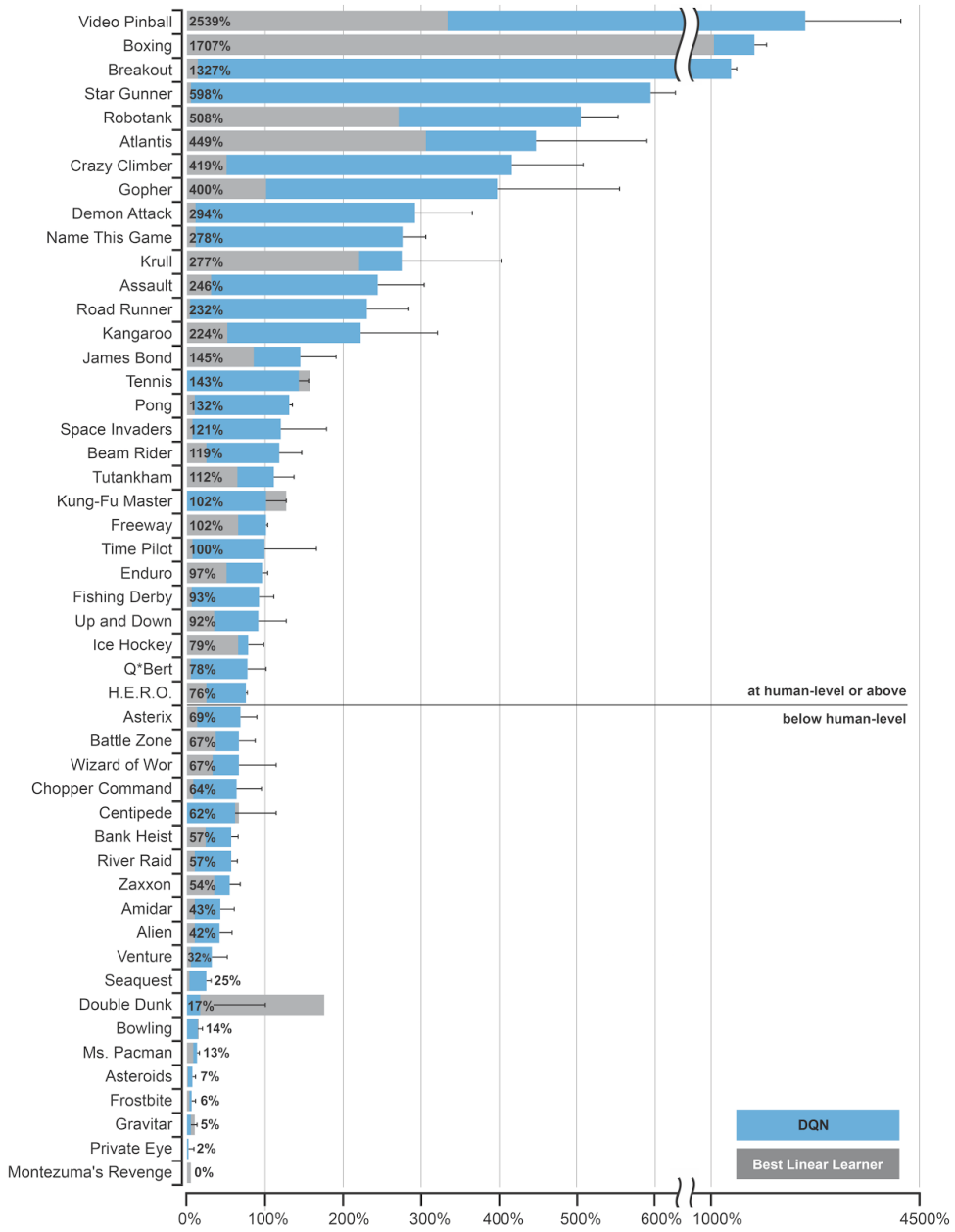


[Atari Games]

NYC DATA SCIENCE ACADEMY

Deep Q-Learning

Mnih et al. (2015)

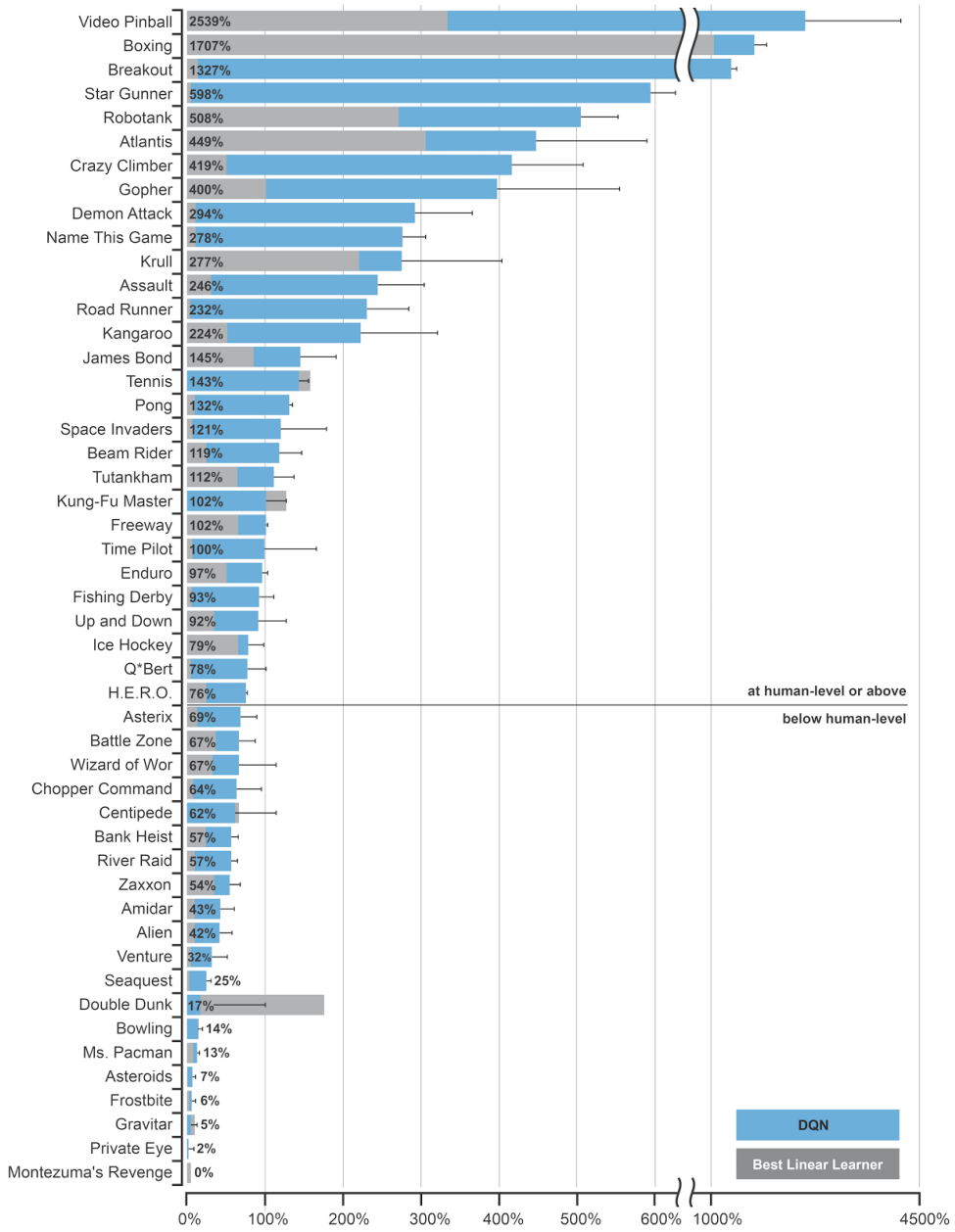


[Atari Games]

NYC DATA SCIENCE ACADEMY

Deep Q-Learning

Mnih et al. (2015)

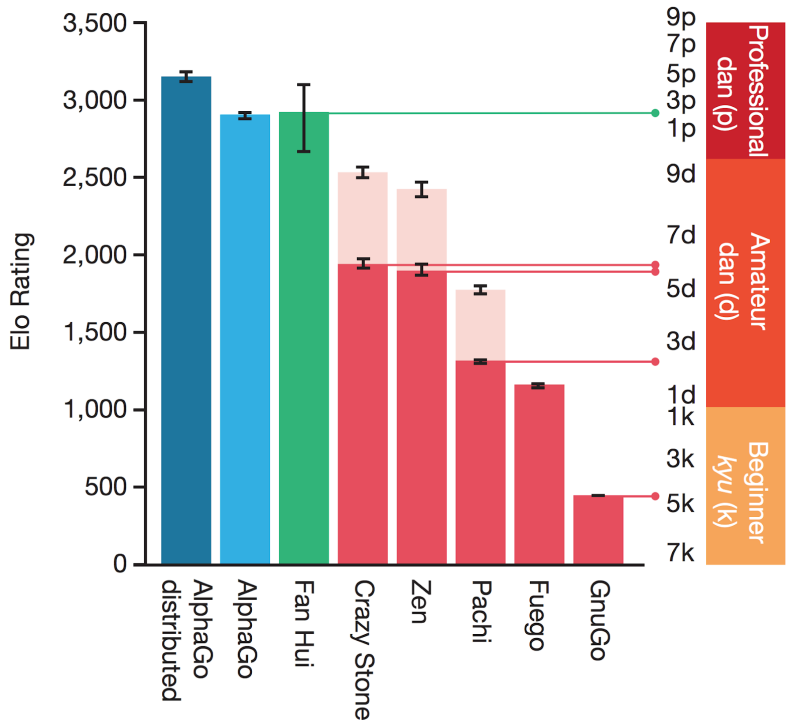


[Atari Games]

NYC DATA SCIENCE ACADEMY

AlphaGo

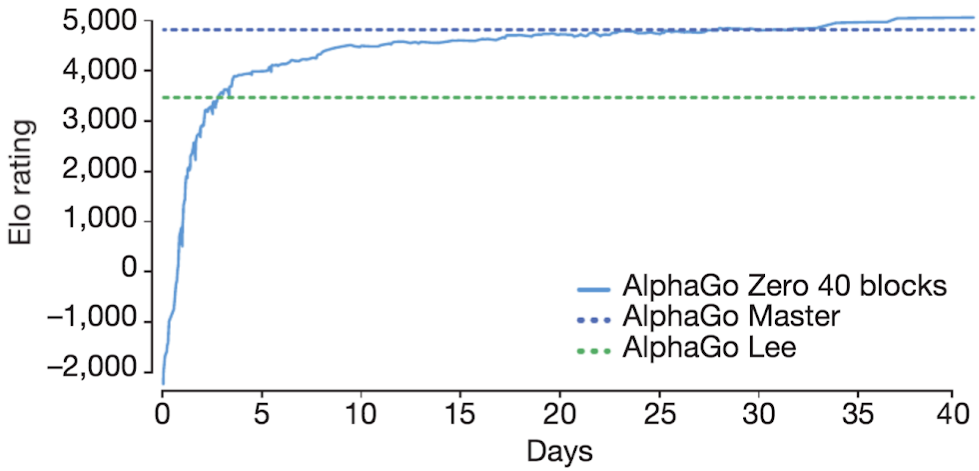
Silver et al. (2016)

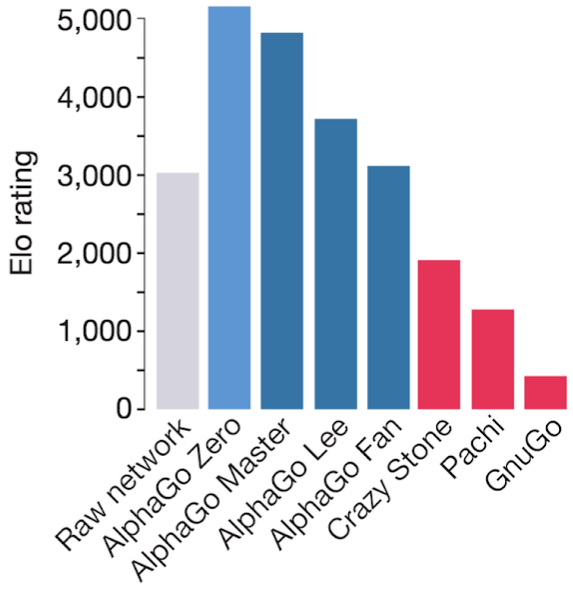


NYC DATA SCIENCE ACADEMY

AlphaGo

Silver et al. (2017a)

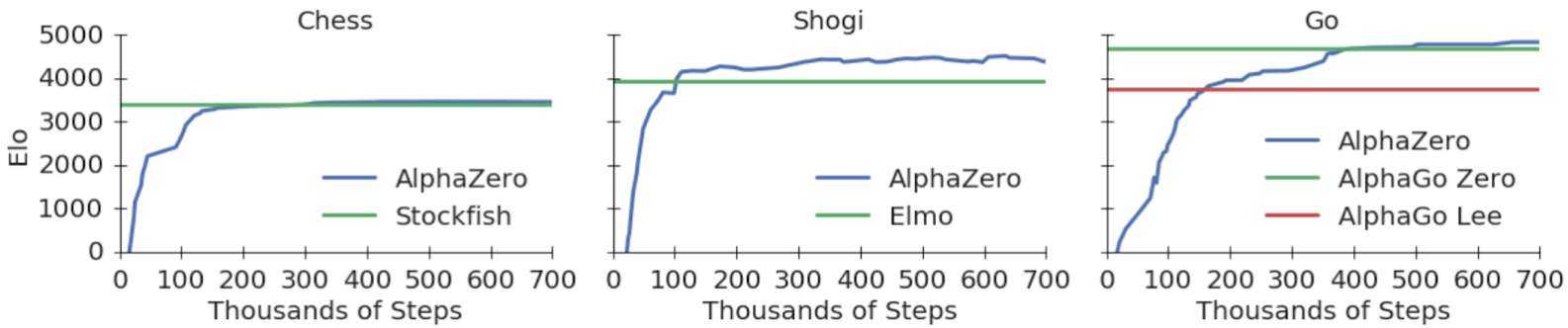




NYC DATA SCIENCE ACADEMY

AlphaZero

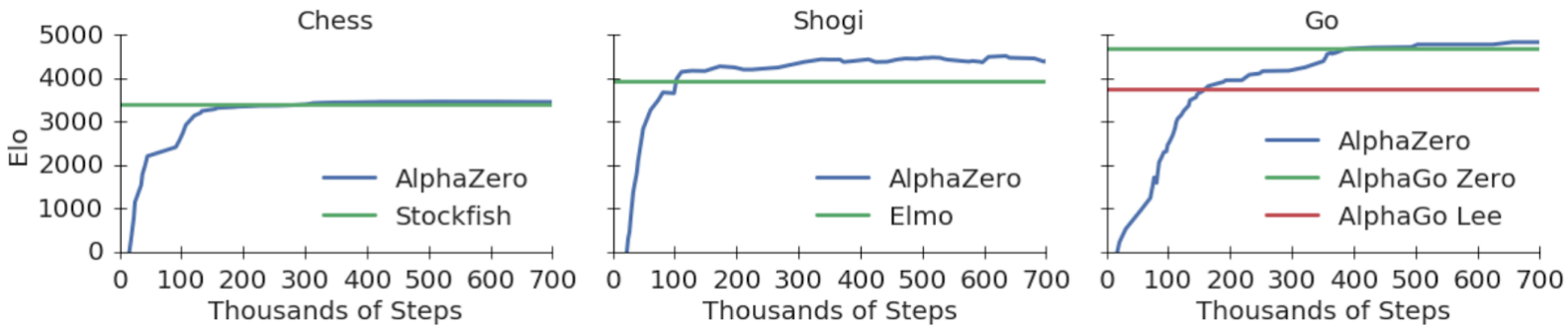
Silver et al. (2017b)



NYC DATA SCIENCE ACADEMY

AlphaZero

Silver et al. (2017b)



NYC DATA SCIENCE ACADEMY

Environments

1. OpenAI Gym
2. DeepMind Lab
3. Unity

Environments

1. OpenAI Gym
2. DeepMind Lab
3. Unity

Environments

1. OpenAI Gym
2. DeepMind Lab
3. Unity

Intro
DQN In Theory
DQN In Practice
Beyond DQN
AI Revolution

Outline

1 Introduction

2 Deep Q-Learning Theory

3 Cart Pole Deep Q-Learning

4 Beyond DQN

5 The AI Revolution

Essential Deep Q-Learning Theory

[whiteboarding]

Outline

1 Introduction

2 Deep Q-Learning Theory

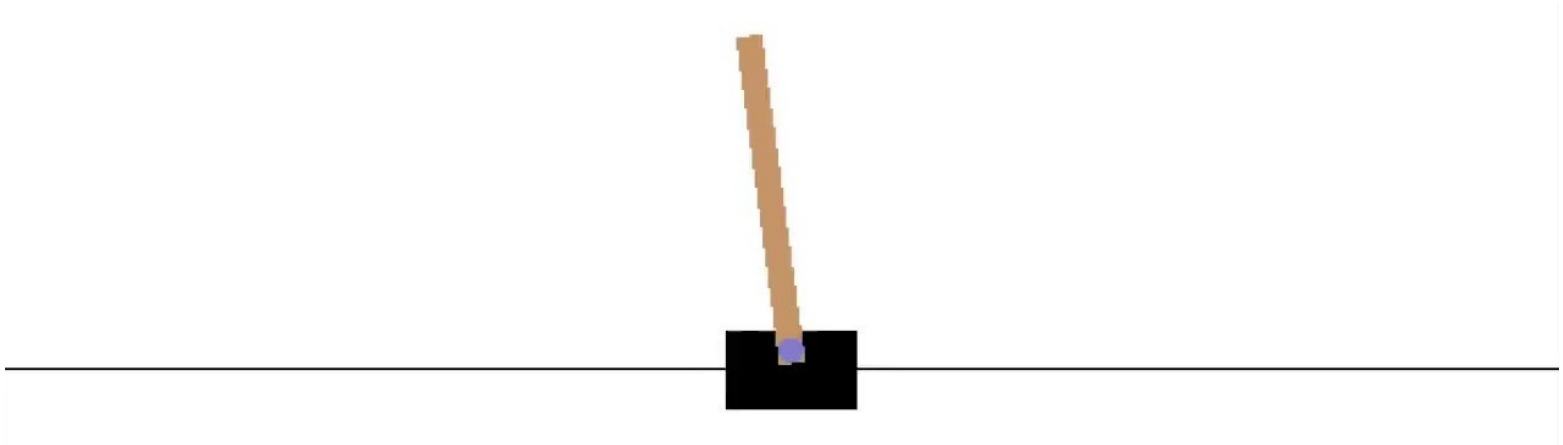
3 Cart Pole Deep Q-Learning

4 Beyond DQN

5 The AI Revolution

The Cartpole Game

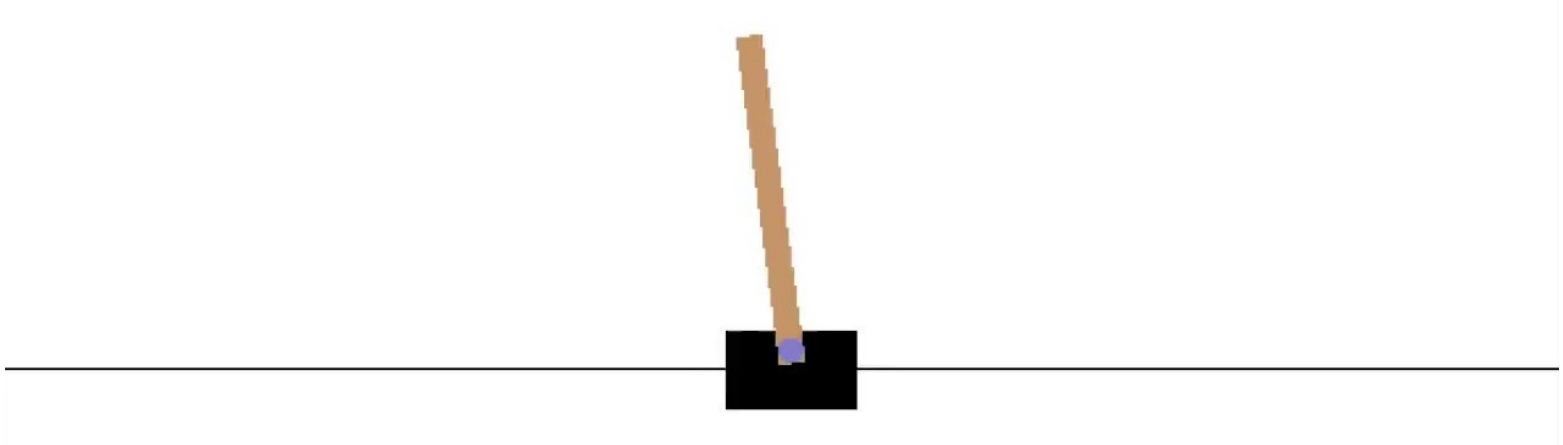
OpenAI



1. cart position
2. cart velocity
3. pole angle
4. pole angular velocity

The Cartpole Game

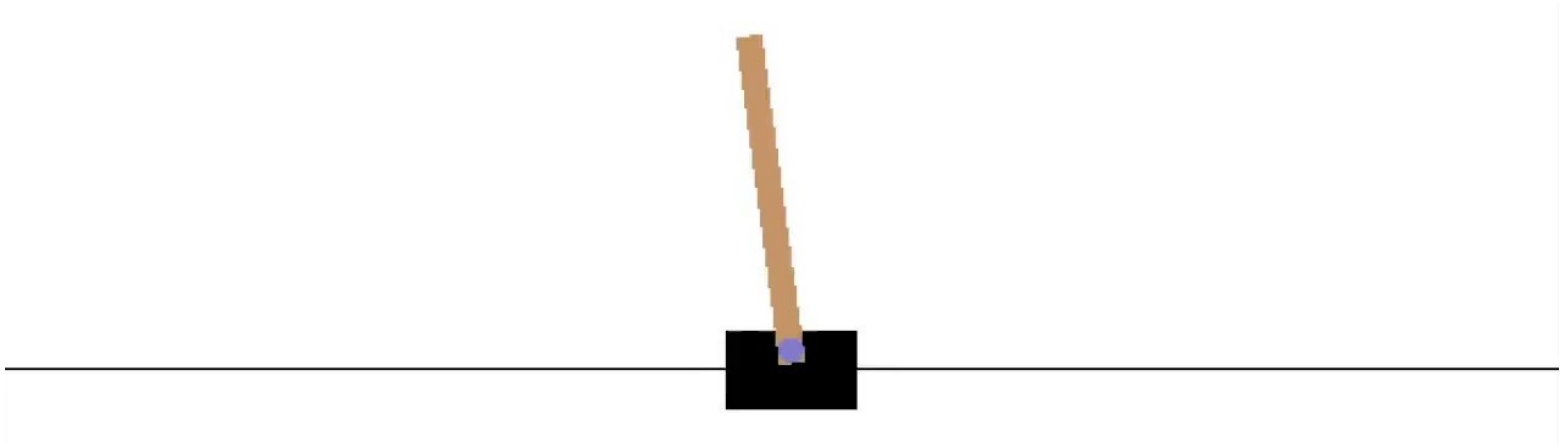
OpenAI



1. cart position
2. cart velocity
3. pole angle
4. pole angular velocity

The Cartpole Game

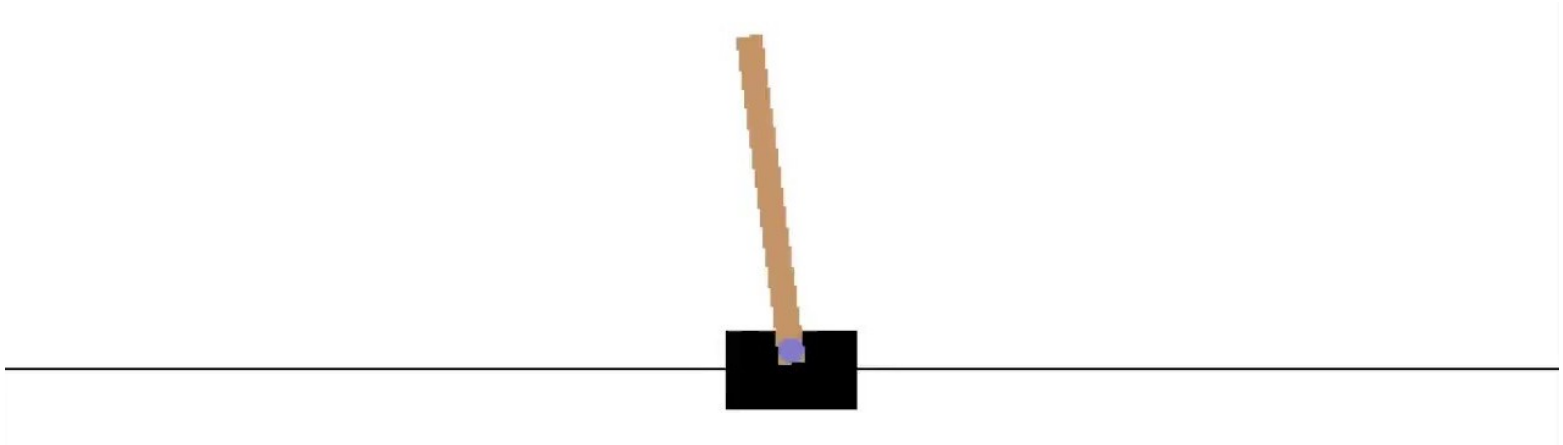
OpenAI



1. cart position
2. cart velocity
3. pole angle
4. pole angular velocity

The Cartpole Game

OpenAI



1. cart position
2. cart velocity
3. pole angle
4. pole angular velocity

Installation

(outside my container

1. https://github.com/openai/gym
2. anaconda
3. `(pip install gym[atari])`
4. `pip uninstall pyglet`
5. `pip install pyglet==1.2.4`
6. `jupyter notebook` in `nyc-ds-academy` directory

Installation

(outside my container

1. https://github.com/openai/gym
2. anaconda
3. `(pip install gym[atari])`
4. `pip uninstall pyglet`
5. `pip install pyglet==1.2.4`
6. `jupyter notebook` in `nyc-ds-academy` directory

Installation

(outside my container

1. https://github.com/openai/gym
2. anaconda
3. (`pip install gym[atari]`)
4. `pip uninstall pyglet`
5. `pip install pyglet==1.2.4`
6. `jupyter notebook` in `nyc-ds-academy` directory

Installation

(outside my container

1. https://github.com/openai/gym
2. anaconda
3. (`pip install gym[atari]`)
4. `pip uninstall pyglet`
5. `pip install pyglet==1.2.4`
6. `jupyter notebook` in `nyc-ds-academy` directory

Installation

(outside my container

1. https://github.com/openai/gym
2. anaconda
3. (`pip install gym[atari]`)
4. `pip uninstall pyglet`
5. `pip install pyglet==1.2.4`
6. `jupyter notebook` in `nyc-ds-academy` directory

Installation

(outside my container

1. https://github.com/openai/gym
2. anaconda
3. (`pip install gym[atari]`)
4. `pip uninstall pyglet`
5. `pip install pyglet==1.2.4`
6. `jupyter notebook` in `nyc-ds-academy` directory

[*cartpole* notebook]

Intro
DQN In Theory
DQN In Practice
Beyond DQN
AI Revolution

Outline

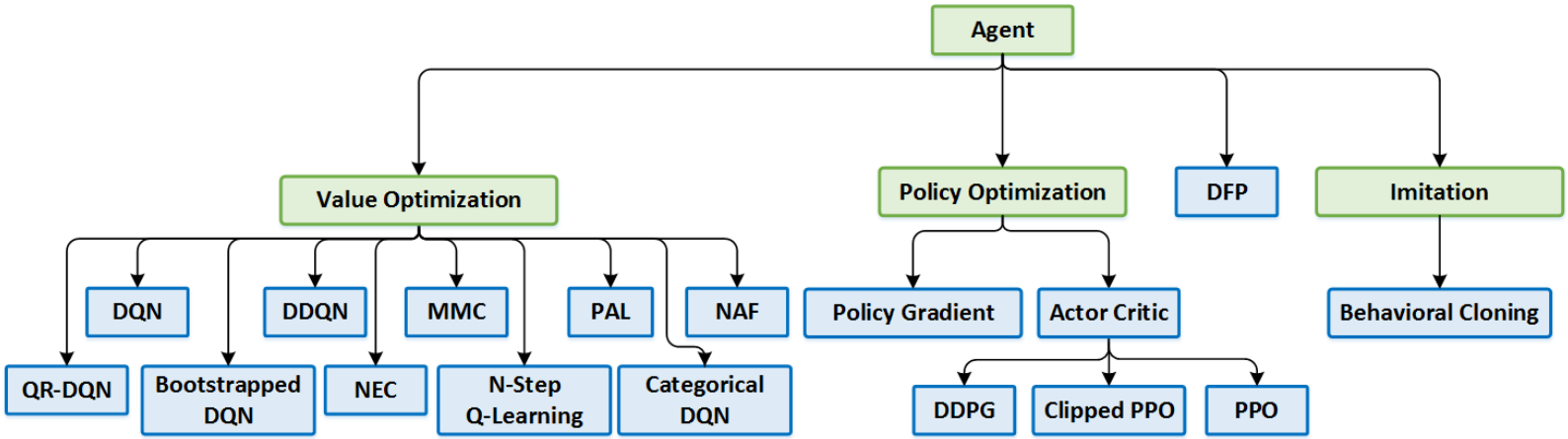
1 Introduction
2 Deep Q-Learning Theory
3 Cart Pole Deep Q-Learning
4 Beyond DQN
5 The AI Revolution

NYC DATA SCIENCE ACADEMY

Genetic Algorithms

[Mar/I/O]

Ready-Made Agents



[Coach]

NYC DATA SCIENCE ACADEMY

Automated Experimentation

[OpenAI Lab]

Human Imitation Learning

Warneken & Tomasello (2006)

[videos]

Machine Imitation Learning

Finn, Levine & Abbeel

[videos]

NYC DATA SCIENCE ACADEMY

Software 2.0 Pros

Karpathy (2017)

1. Computational Homogeneity
2. Constant Running Time
3. Constant Memory Use
4. Portable Across Devices
5. Easy
6. Superior

[Medium]

Software 2.0 Pros

Karpathy (2017)

1. Computational Homogeneity
2. Constant Running Time
3. Constant Memory Use
4. Portable Across Devices
5. Easy
6. Superior

[Medium]

Software 2.0 Pros

Karpathy (2017)

1. Computational Homogeneity
2. Constant Running Time
3. Constant Memory Use
4. Portable Across Devices
5. Easy
6. Superior

[Medium]

Software 2.0 Pros

Karpathy (2017)

1. Computational Homogeneity
2. Constant Running Time
3. Constant Memory Use
4. Portable Across Devices
5. Easy
6. Superior

[Medium]

Software 2.0 Pros

Karpathy (2017)

1. Computational Homogeneity
2. Constant Running Time
3. Constant Memory Use
4. Portable Across Devices
5. Easy
6. Superior

[Medium]

Software 2.0 Pros

Karpathy (2017)

1. Computational Homogeneity
2. Constant Running Time
3. Constant Memory Use
4. Portable Across Devices
5. Easy
6. Superior

[Medium]

Intro
DQN In Theory
DQN In Practice
Beyond DQN
AI Revolution

Outline

1 Introduction

2 Deep Q-Learning Theory

3 Cart Pole Deep Q-Learning

4 Beyond DQN

5 The AI Revolution

Jeanne Calment

Jeanne Calment

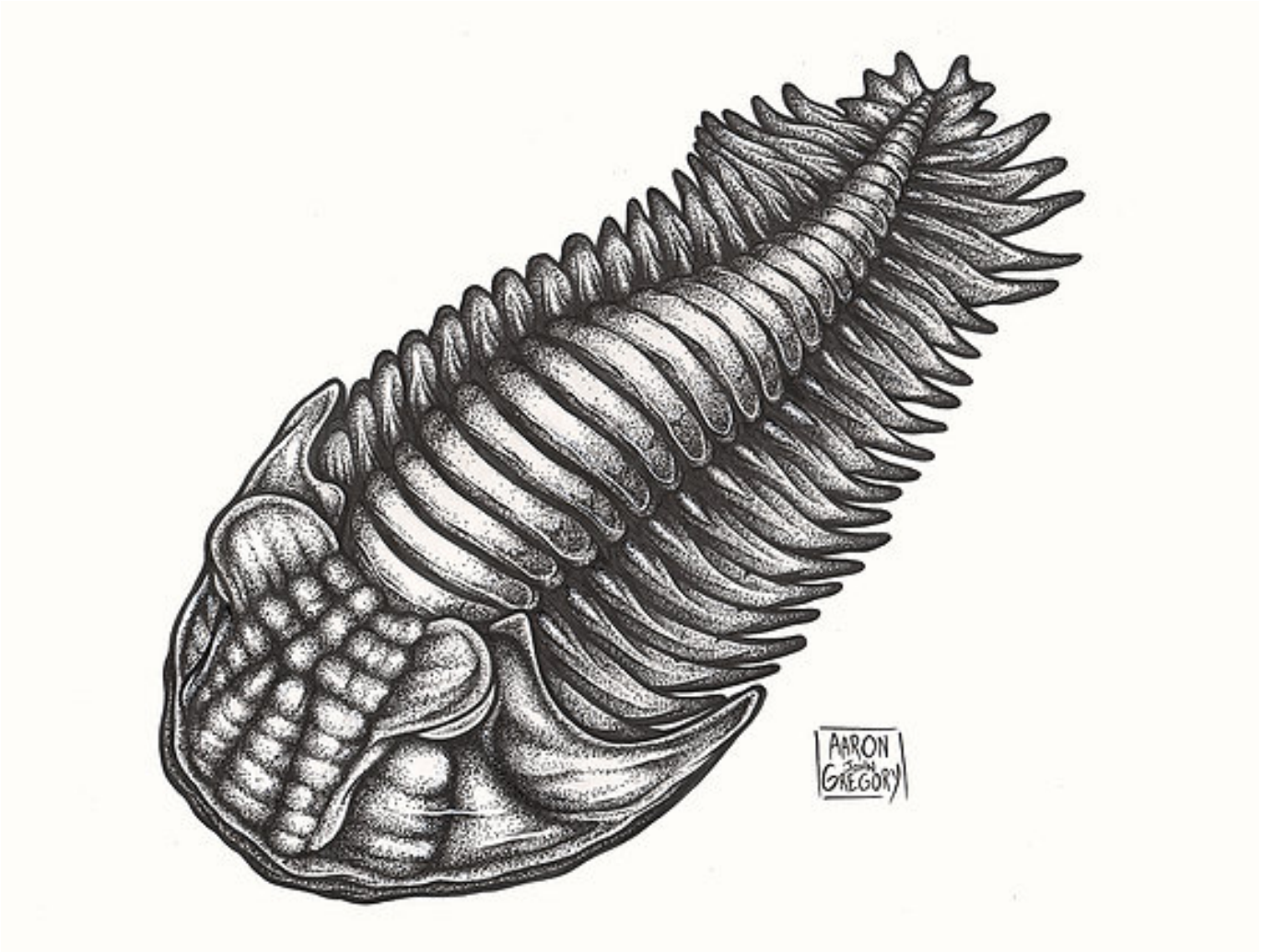
(1875-1997 — i.e., 122 years)

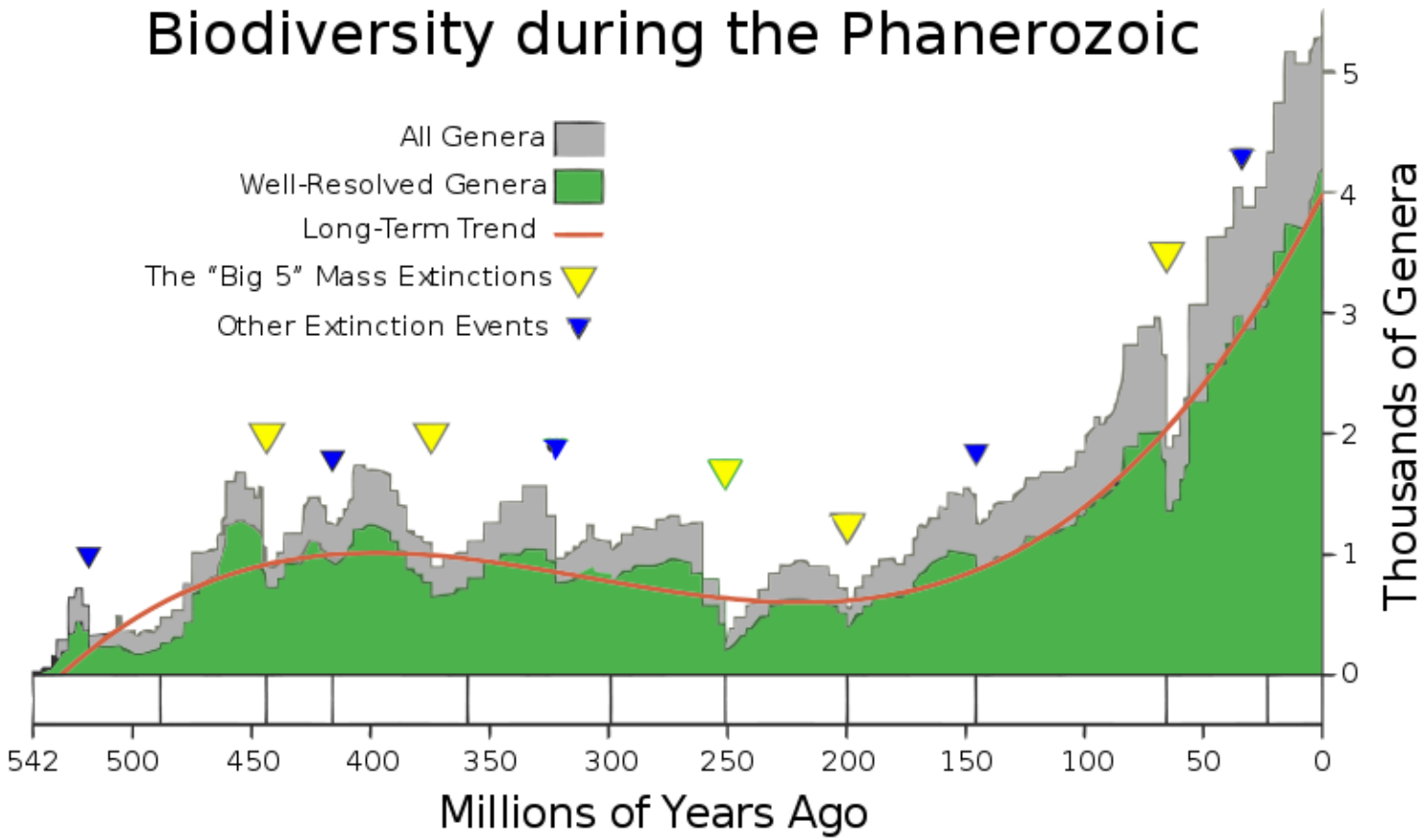
NYC DATA SCIENCE ACADEMY

Life in the Year 2138

The AI Revolution

Hasn't Even Begun


AARON
GREGORY



NYC DATA SCIENCE ACADEMY

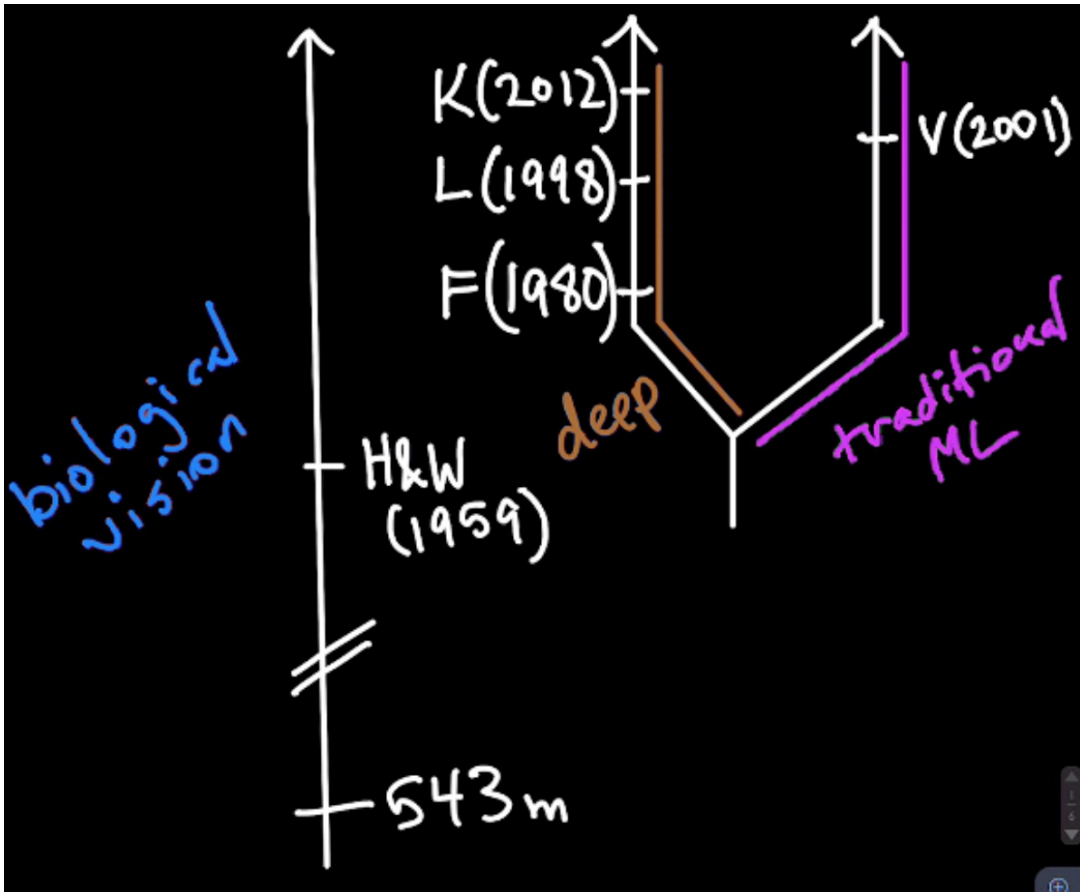
Intro

DQN In Theory

DQN In Practice

Beyond DQN

AI Revolution



NYC DATA SCIENCE ACADEMY

Factors Driving A.I.

1. Data, incl. Sensors Everywhere
2. Computing Power
3. Algorithms
4. Infrastructure

[Wait But Why]

Factors Driving A.I.

1. Data, incl. Sensors Everywhere
2. Computing Power
3. Algorithms
4. Infrastructure

[Wait But Why]

Factors Driving A.I.

1. Data, incl. Sensors Everywhere
2. Computing Power
3. Algorithms
4. Infrastructure

[Wait But Why]

Factors Driving A.I.

1. Data, incl. Sensors Everywhere
2. Computing Power
3. Algorithms
4. Infrastructure

[Wait But Why]

Categories of A.I.

1. ANI: *Narrow*
2. AGI: *General*
3. ASI: *Super*

Categories of A.I.

1. ANI: *Narrow*
2. AGI: *General*
3. ASI: *Super*

Categories of A.I.

1. ANI: *Narrow*
2. AGI: *General*
3. ASI: *Super*

AGI

Deep Learning Barriers

1. Requires Many, Many Samples
2. Black Box
3. Doesn't Leverage Knowledge of the World
4. Correlation = Causation
5. Unintuitive / Embarrassing / Adversarial Failures

AGI

Deep Learning Barriers

1. Requires Many, Many Samples
2. Black Box
3. Doesn't Leverage Knowledge of the World
4. Correlation = Causation
5. Unintuitive / Embarrassing / Adversarial Failures

AGI

Deep Learning Barriers

1. Requires Many, Many Samples
2. Black Box
3. Doesn't Leverage Knowledge of the World
4. Correlation = Causation
5. Unintuitive / Embarrassing / Adversarial Failures

AGI

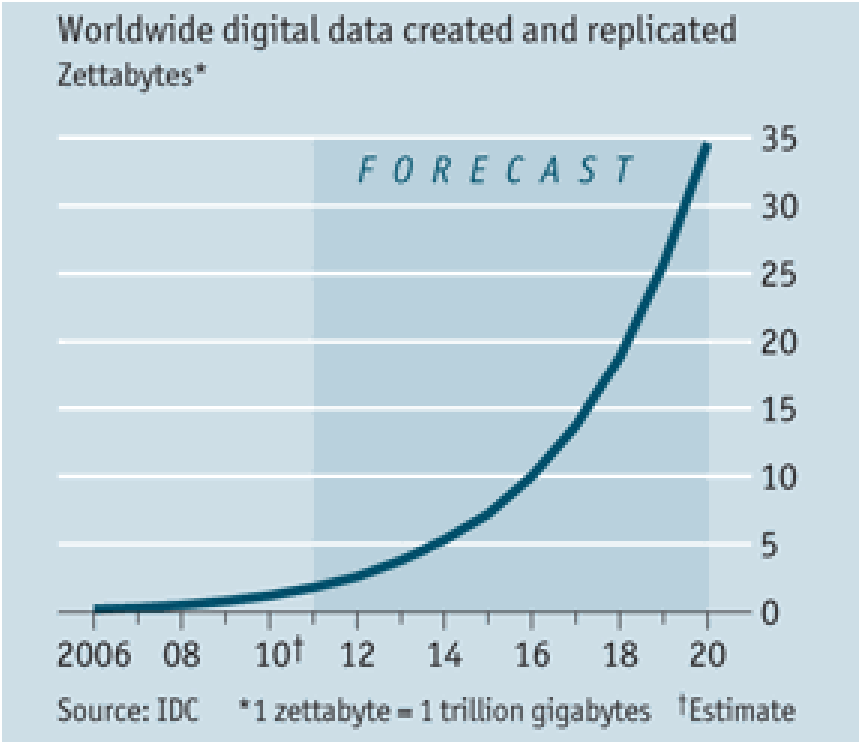
Deep Learning Barriers

1. Requires Many, Many Samples
2. Black Box
3. Doesn't Leverage Knowledge of the World
4. Correlation = Causation
5. Unintuitive / Embarrassing / Adversarial Failures

AGI

Deep Learning Barriers

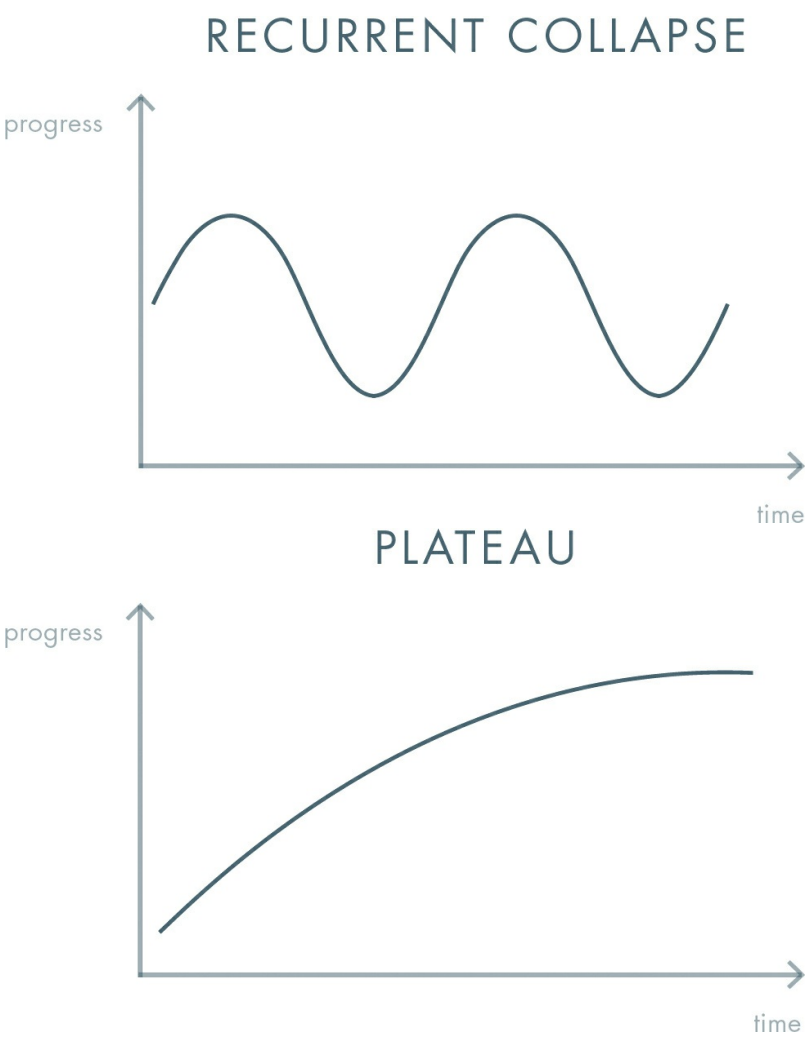
1. Requires Many, Many Samples
2. Black Box
3. Doesn't Leverage Knowledge of the World
4. Correlation = Causation
5. Unintuitive / Embarrassing / Adversarial Failures



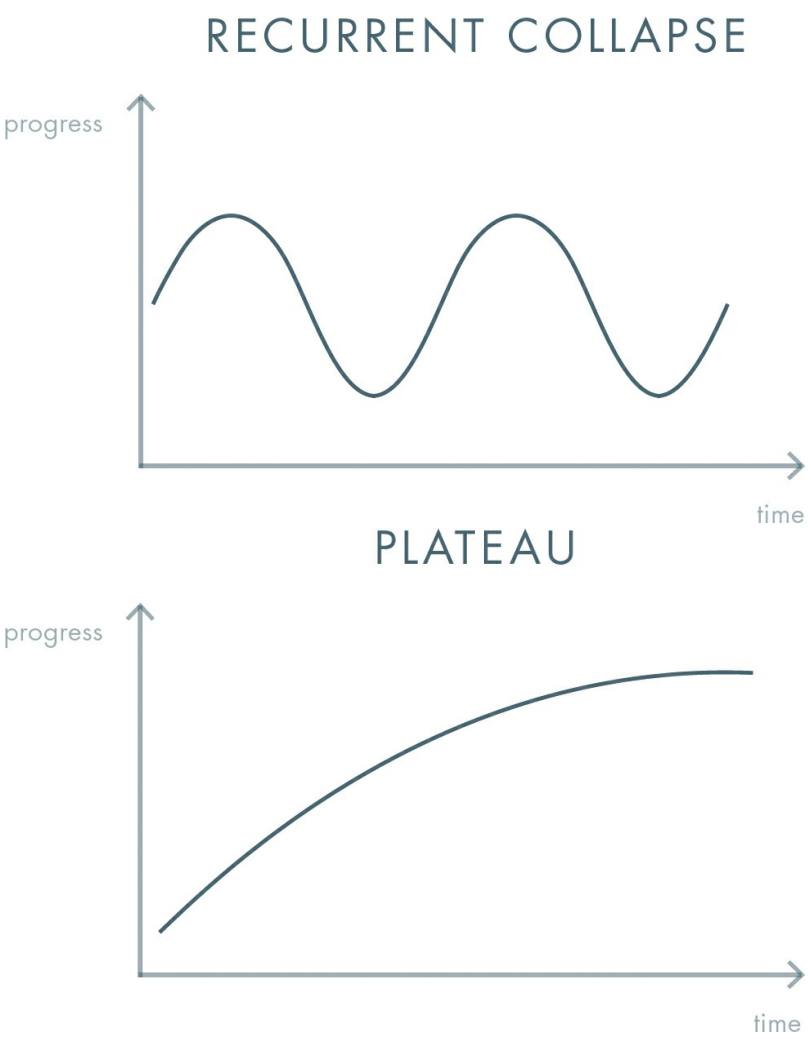
NYC DATA SCIENCE ACADEMY

Jeanne Calment

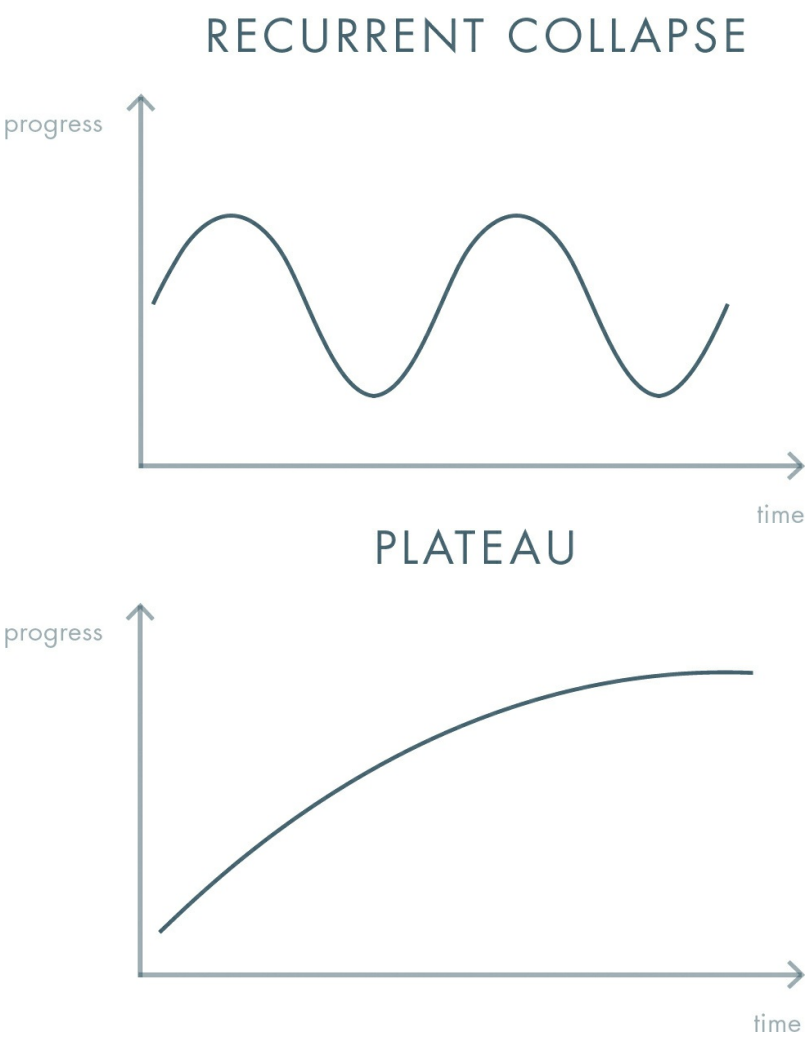
Thiel & Masters (2014)

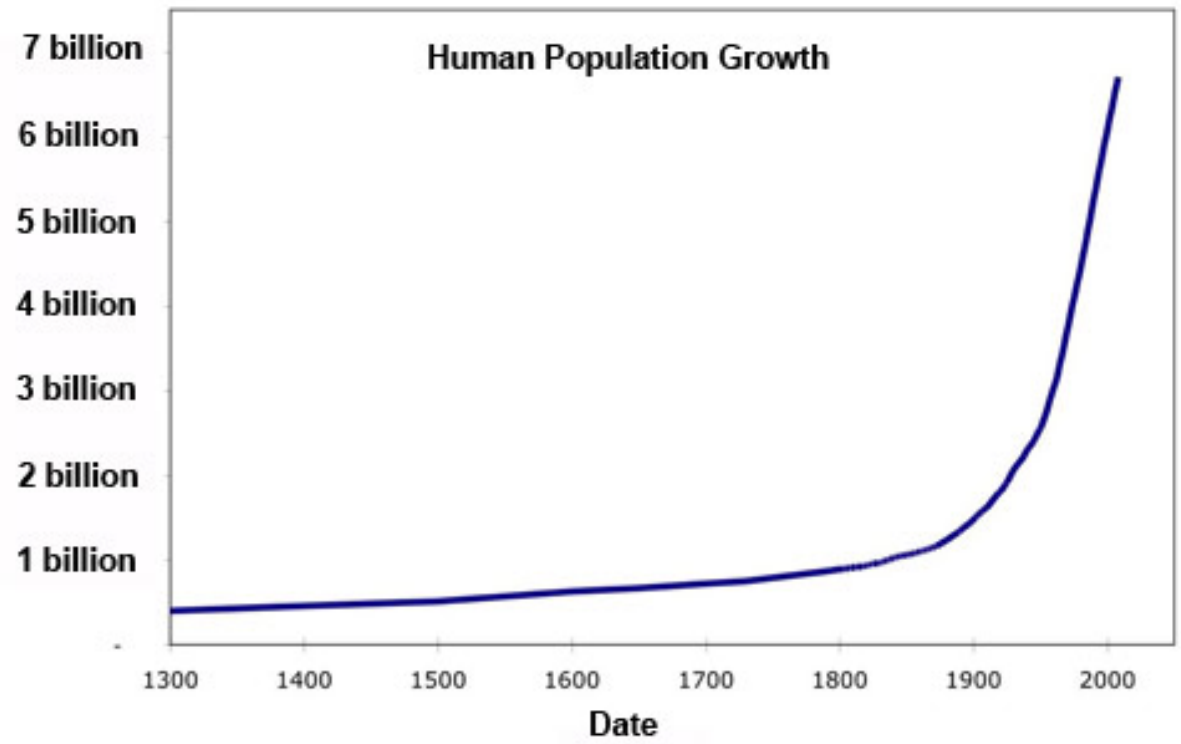


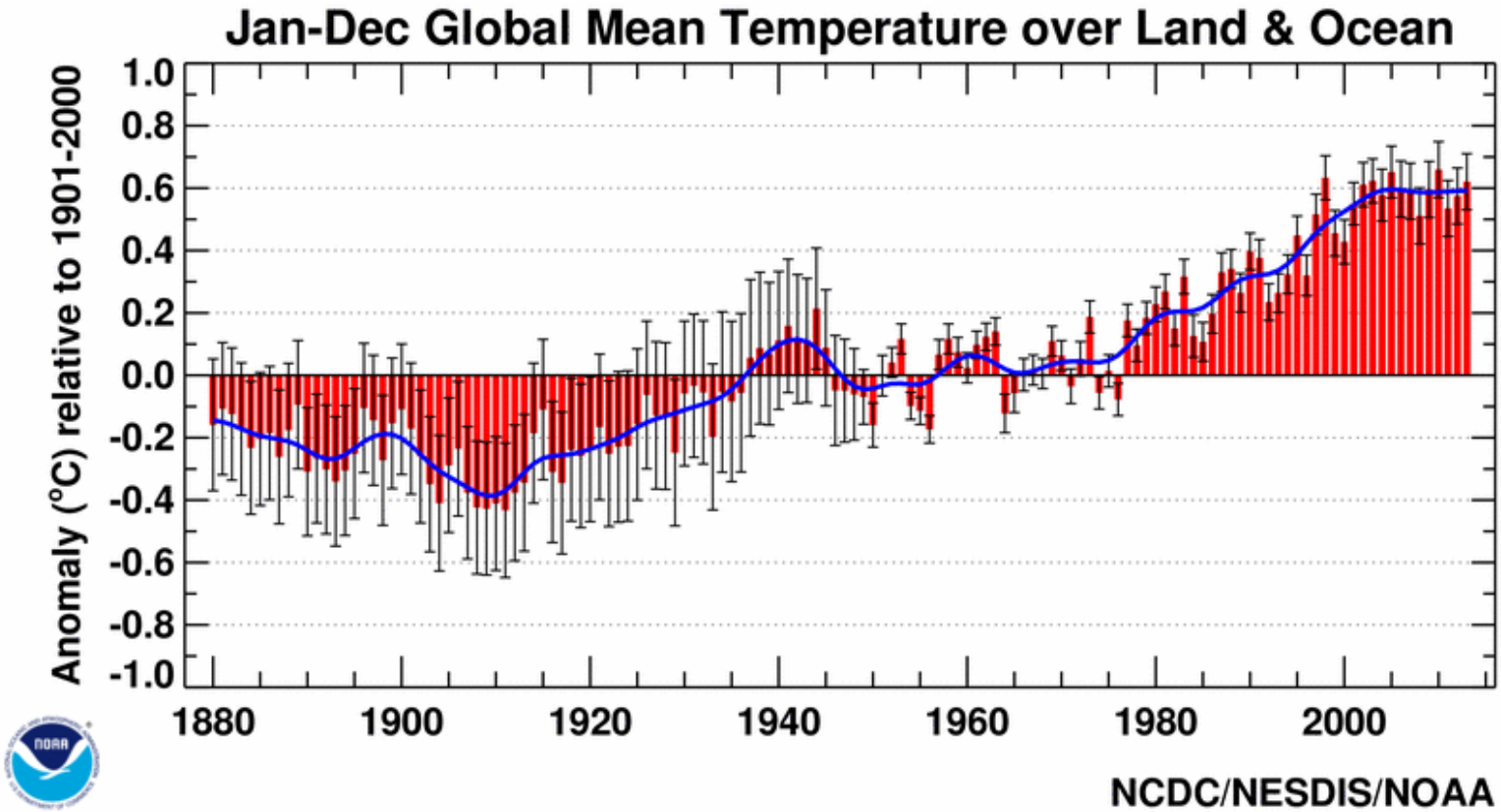
Thiel & Masters (2014)

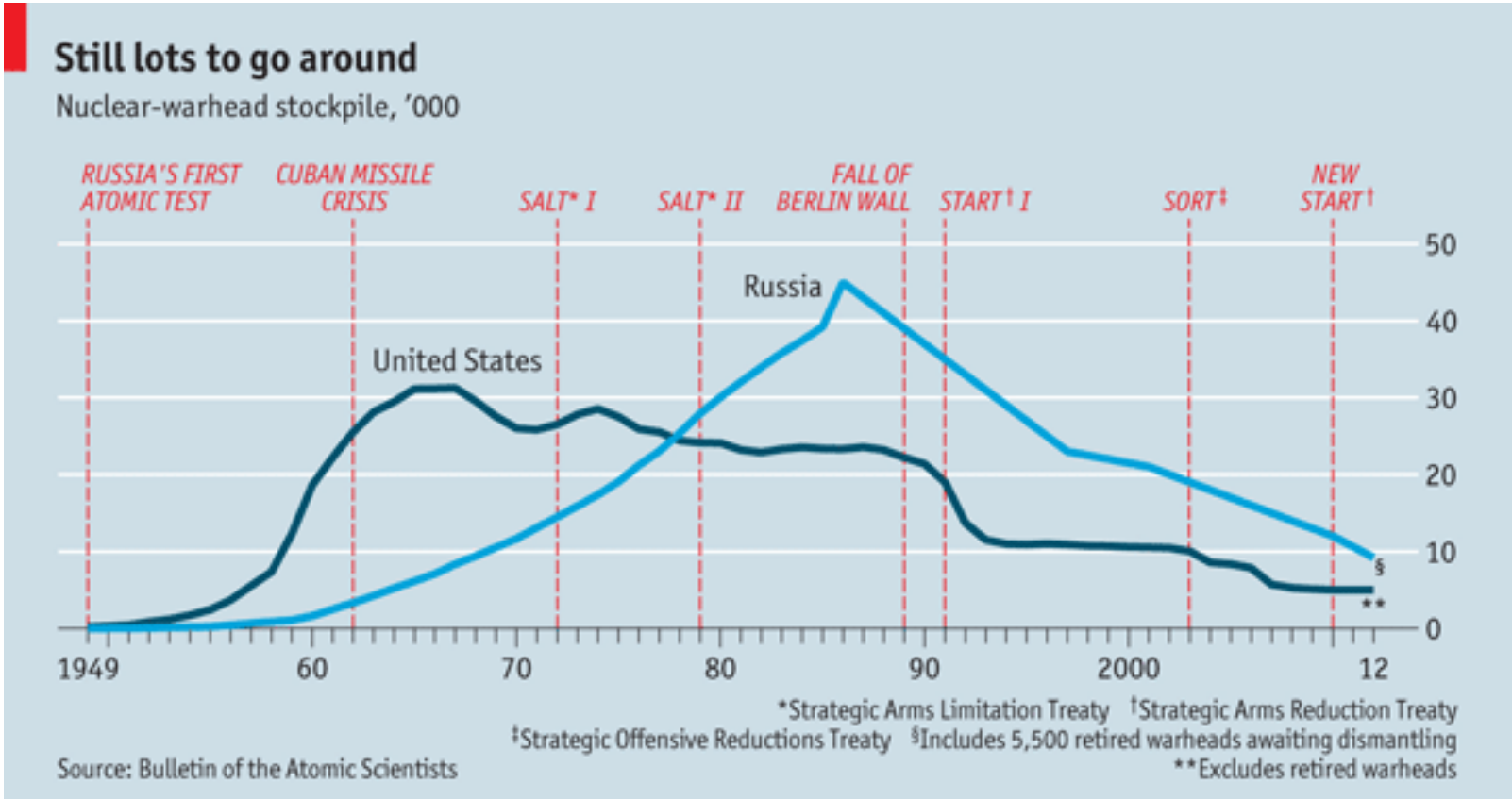


Thiel & Masters (2014)



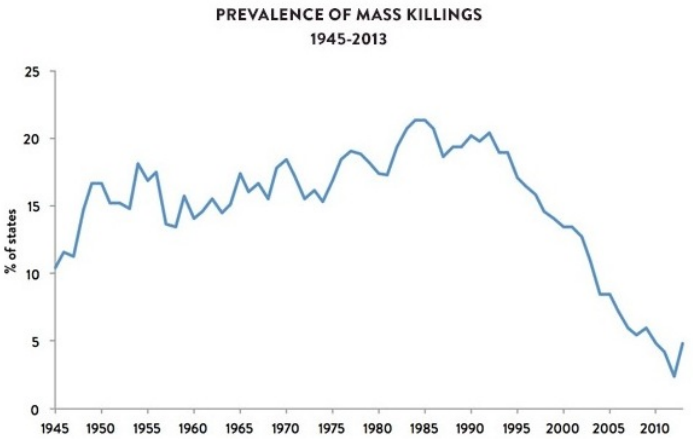

Human Population Growth
7 billion
6 billion
5 billion
4 billion
3 billion
2 billion
1 billion
1300
1400
1500
1600
1700
1800
1900
2000
Date


Jan-Dec Global Mean Temperature over Land & Ocean
Anomaly (°C) relative to 1901-2000
1.0
0.8
0.6
0.4
0.2
0.0
-0.2
-0.4
-0.6
-0.8
-1.0
1880
1900
1920
1940
1960
1980
2000
NCDC/NESDIS/NOAA

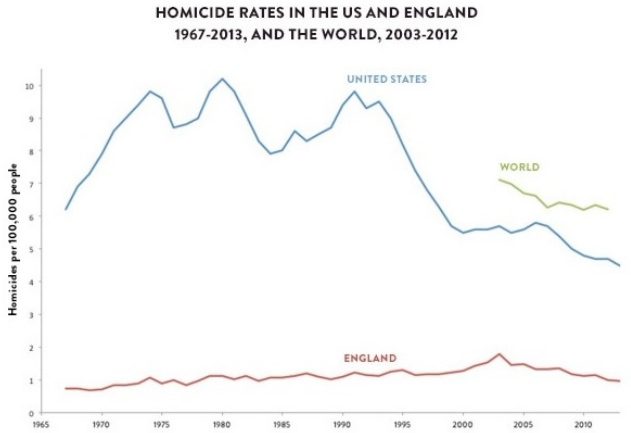


NYC DATA SCIENCE ACADEMY

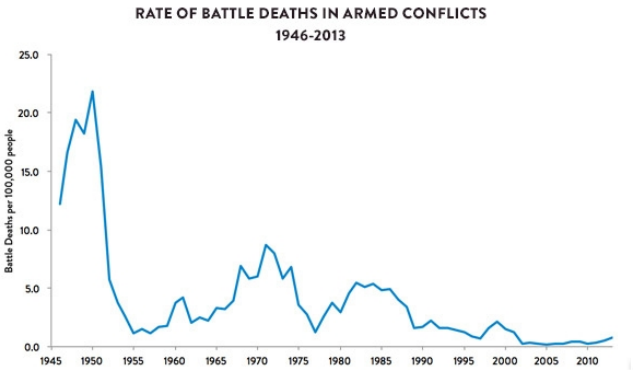
Pinker & Mack (2014)



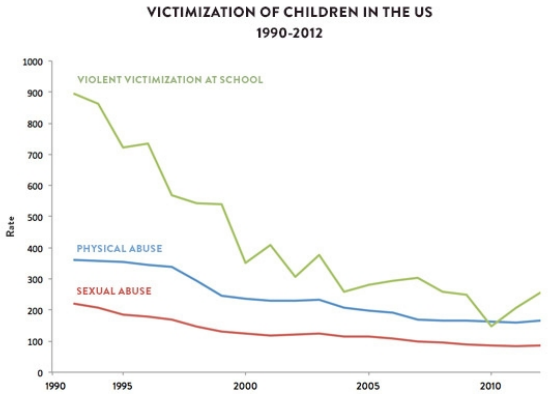
Source: Early Warning Project, https://cpgearlywarning.wordpress.com/; Ulfelder, Jay and Valentino, Benjamin, "Assessing Risks of State-Sponsored Mass Killing" (Feb. 1, 2008). Available at SSRN: http://ssrn.com/abstract=1703426 or http://dx.doi.org/10.2139/ssrn.1703426



Sources: U.S., FBI Uniform Crime Reports. England (including Wales): U.K. Office for National Statistics. World: U.N. Office on Drugs and Crime, reported in U.N. Economic and Social Council's "World crime trends and emerging issues and responses in the field of crime prevention and social justice," Feb. 12, 2014, Figure 1. The percentages were converted to homicide rates by setting the 2012 rate at 6.2, the figure reported in the UNODC Global Study on Homicide 2013, Page 12.



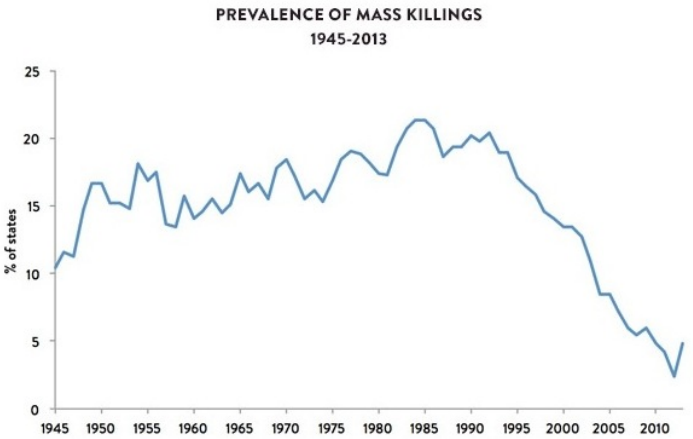
Source: Uppsala Conflict Data Program Battle-Related Deaths Dataset v.5-2014, 1989-2013, "Best Estimates." Missing estimate for 2013 for the Syrian civil war calculated from a revised UCDP "Low Estimate" and a "High Estimate" from the Peace Research Institute Oslo (both obtained in consultation with Erik Melander of UCDP). The "High Estimate" is for May 2013–April 2014, rather than calendar year 2013, and comes from Page 9 of M. Price, A. Gohdes, & P. Ball, "Updated Statistical Analysis of Documentation of Killings in the Syrian Arab Republic," Human Rights Data Analysis Group, http://www.ohchr.org/Documents/Countries/SY/HRDAGUpdatedReportAug2014.pdf. World population figures from U.S. Census Bureau, http://www.census.gov/population/international/data/worldpop/table_population.php



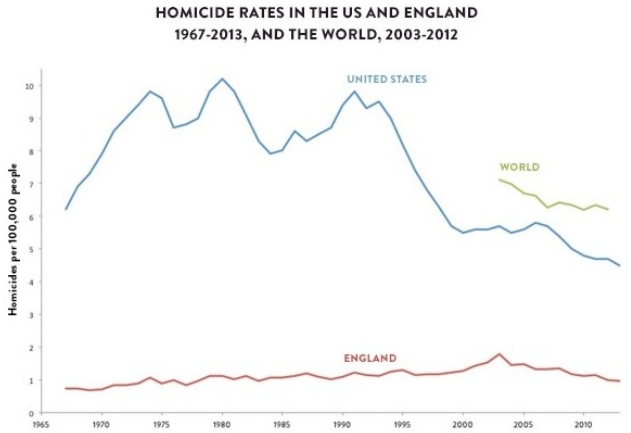
Rates for physical and sexual abuse are per 100,000 children younger than 18. Rates for violent victimization at school are per 10,000 children age 12–17.

Sources: Physical and sexual abuse, National Child Abuse and Neglect Data System, analyzed by David Finkelhor, 2014, in "Trends in Child Welfare," Presentation at the Carsey Institute Policy Series, March 31, 2014. Victimization at school, Bureau of Justice Statistics (bjs.gov), using the National Crime Victimization Survey Victimization Analysis Tool.

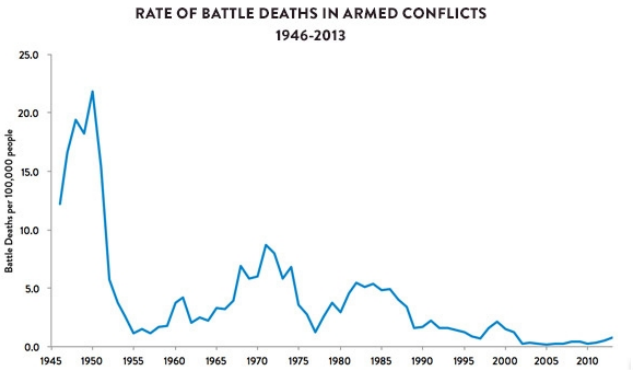
Pinker & Mack (2014)



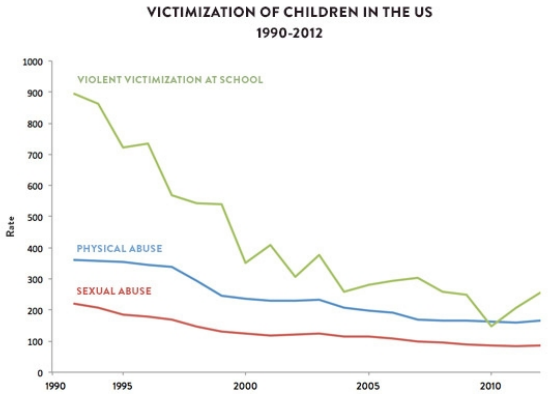
Source: Early Warning Project, https://cpgearlywarning.wordpress.com/; Ulfelder, Jay and Valentino, Benjamin, "Assessing Risks of State-Sponsored Mass Killing" (Feb. 1, 2008). Available at SSRN: http://ssrn.com/abstract=1703426 or http://dx.doi.org/10.2139/ssrn.1703426



Sources: U.S., FBI Uniform Crime Reports. England (including Wales): U.K. Office for National Statistics. World: U.N. Office on Drugs and Crime, reported in U.N. Economic and Social Council's "World crime trends and emerging issues and responses in the field of crime prevention and social justice," Feb. 12, 2014, Figure 1. The percentages were converted to homicide rates by setting the 2012 rate at 6.2, the figure reported in the UNODC Global Study on Homicide 2013, Page 12.



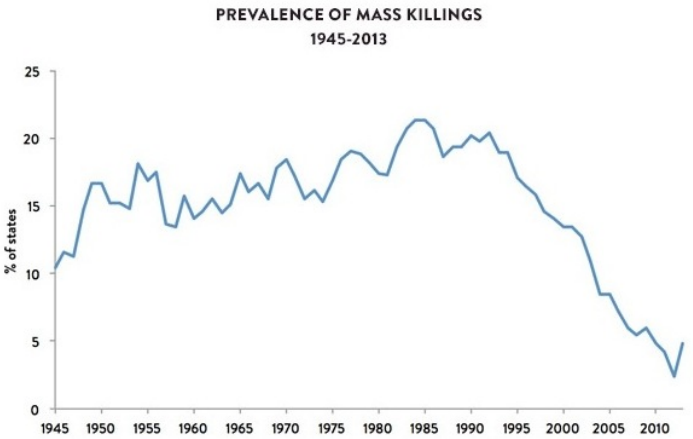
Source: Uppsala Conflict Data Program Battle-Related Deaths Dataset v.5-2014, 1989-2013, "Best Estimates." Missing estimate for 2013 for the Syrian civil war calculated from a revised UCDP "Low Estimate" and a "High Estimate" from the Peace Research Institute Oslo (both obtained in consultation with Erik Melander of UCDP). The "High Estimate" is for May 2013–April 2014, rather than calendar year 2013, and comes from Page 9 of M. Price, A. Gohdes, & P. Ball, "Updated Statistical Analysis of Documentation of Killings in the Syrian Arab Republic," Human Rights Data Analysis Group, http://www.ohchr.org/Documents/Countries/SY/HRDAGUpdatedReportAug2014.pdf. World population figures from U.S. Census Bureau, http://www.census.gov/population/international/data/worldpop/table_population.php



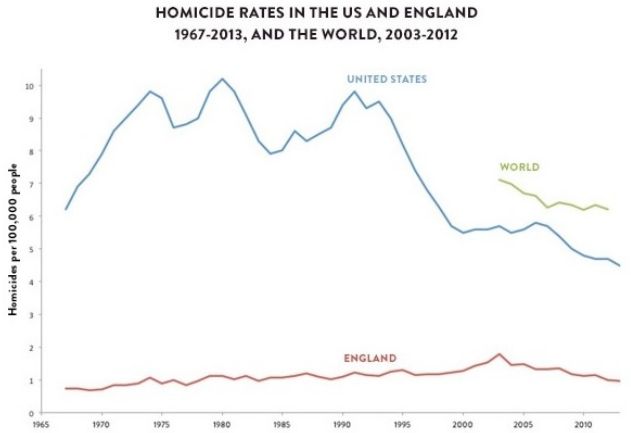
Rates for physical and sexual abuse are per 100,000 children younger than 18. Rates for violent victimization at school are per 10,000 children age 12–17.

Sources: Physical and sexual abuse, National Child Abuse and Neglect Data System, analyzed by David Finkelhor, 2014, in "Trends in Child Welfare," Presentation at the Carsey Institute Policy Series, March 31, 2014. Victimization at school, Bureau of Justice Statistics (bjs.gov), using the National Crime Victimization Survey Victimization Analysis Tool.

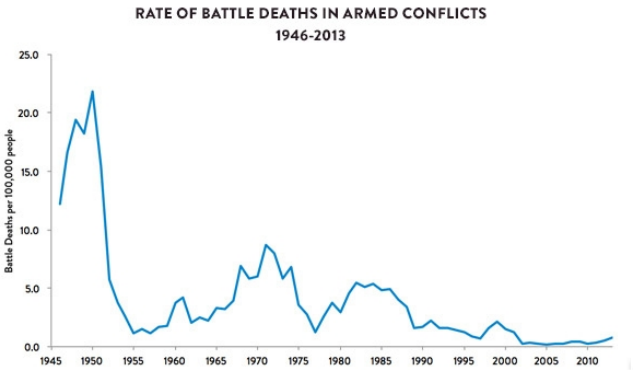
Pinker & Mack (2014)



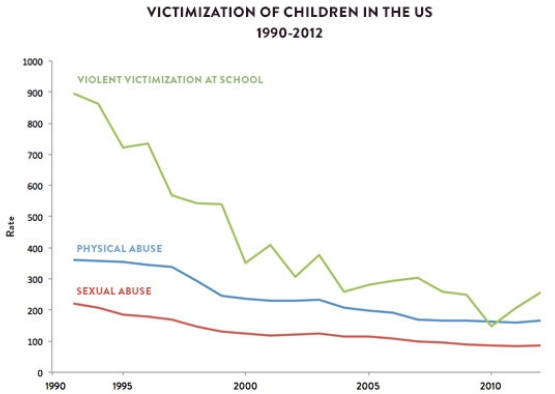
Source: Early Warning Project, https://cpgearlywarning.wordpress.com/; Ulfelder, Jay and Valentino, Benjamin, "Assessing Risks of State-Sponsored Mass Killing" (Feb. 1, 2008). Available at SSRN: http://ssrn.com/abstract=1703426 or http://dx.doi.org/10.2139/ssrn.1703426



Sources: U.S., FBI Uniform Crime Reports. England (including Wales): U.K. Office for National Statistics. World: U.N. Office on Drugs and Crime, reported in U.N. Economic and Social Council's "World crime trends and emerging issues and responses in the field of crime prevention and social justice," Feb. 12, 2014, Figure 1. The percentages were converted to homicide rates by setting the 2012 rate at 6.2, the figure reported in the UNODC Global Study on Homicide 2013, Page 12.



Source: Uppsala Conflict Data Program Battle-Related Deaths Dataset v.5-2014, 1989-2013, "Best Estimates." Missing estimate for 2013 for the Syrian civil war calculated from a revised UCDP "Low Estimate" and a "High Estimate" from the Peace Research Institute Oslo (both obtained in consultation with Erik Melander of UCDP). The "High Estimate" is for May 2013–April 2014, rather than calendar year 2013, and comes from Page 9 of M. Price, A. Gohdes, & P. Ball, "Updated Statistical Analysis of Documentation of Killings in the Syrian Arab Republic," Human Rights Data Analysis Group, http://www.ohchr.org/Documents/Countries/SY/HRDAGUpdatedReportAug2014.pdf. World population figures from U.S. Census Bureau, http://www.census.gov/population/international/data/worldpop/table_population.php

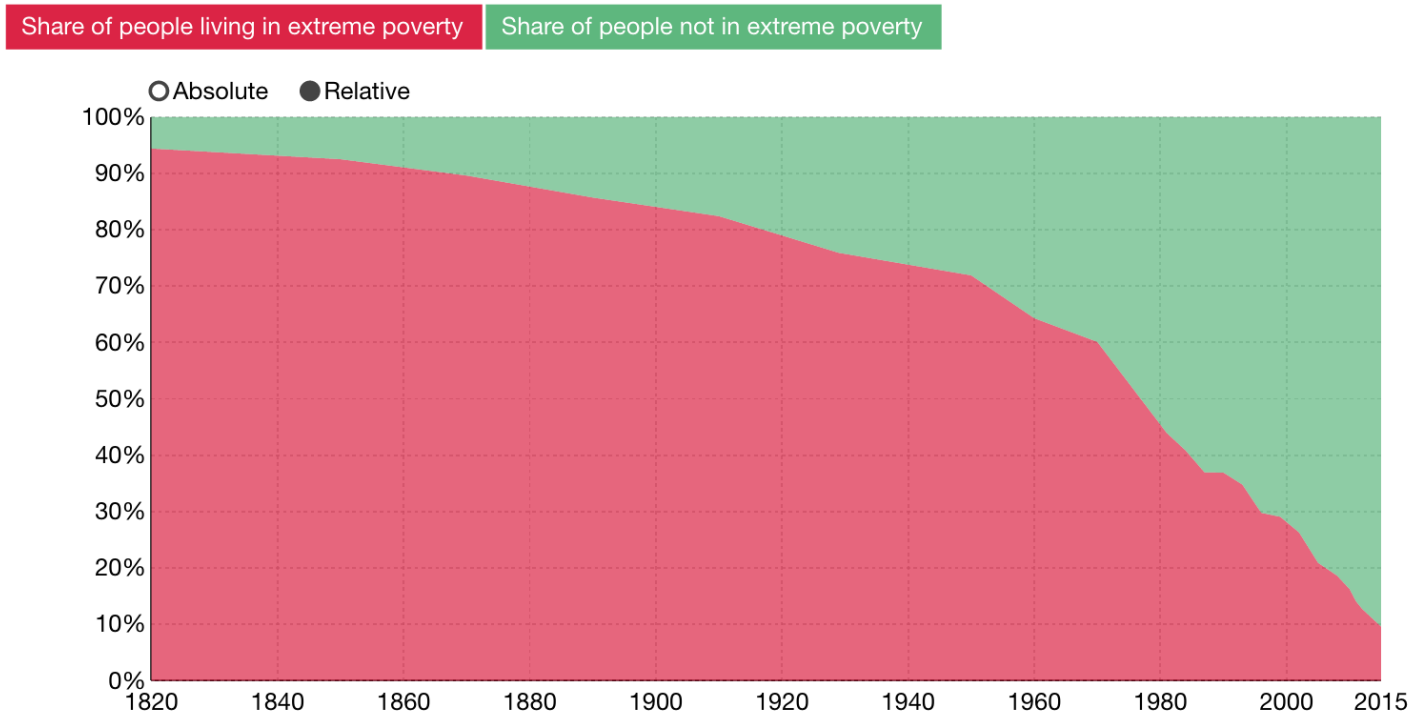


Rates for physical and sexual abuse are per 100,000 children younger than 18. Rates for violent victimization at school are per 10,000 children age 12–17.

Sources: Physical and sexual abuse, National Child Abuse and Neglect Data System, analyzed by David Finkelhor, 2014, in "Trends in Child Welfare," Presentation at the Carsey Institute Policy Series, March 31, 2014. Victimization at school, Bureau of Justice Statistics (bjs.gov), using the National Crime Victimization Survey Victimization Analysis Tool.

World population living in extreme poverty, 1820 to 2015

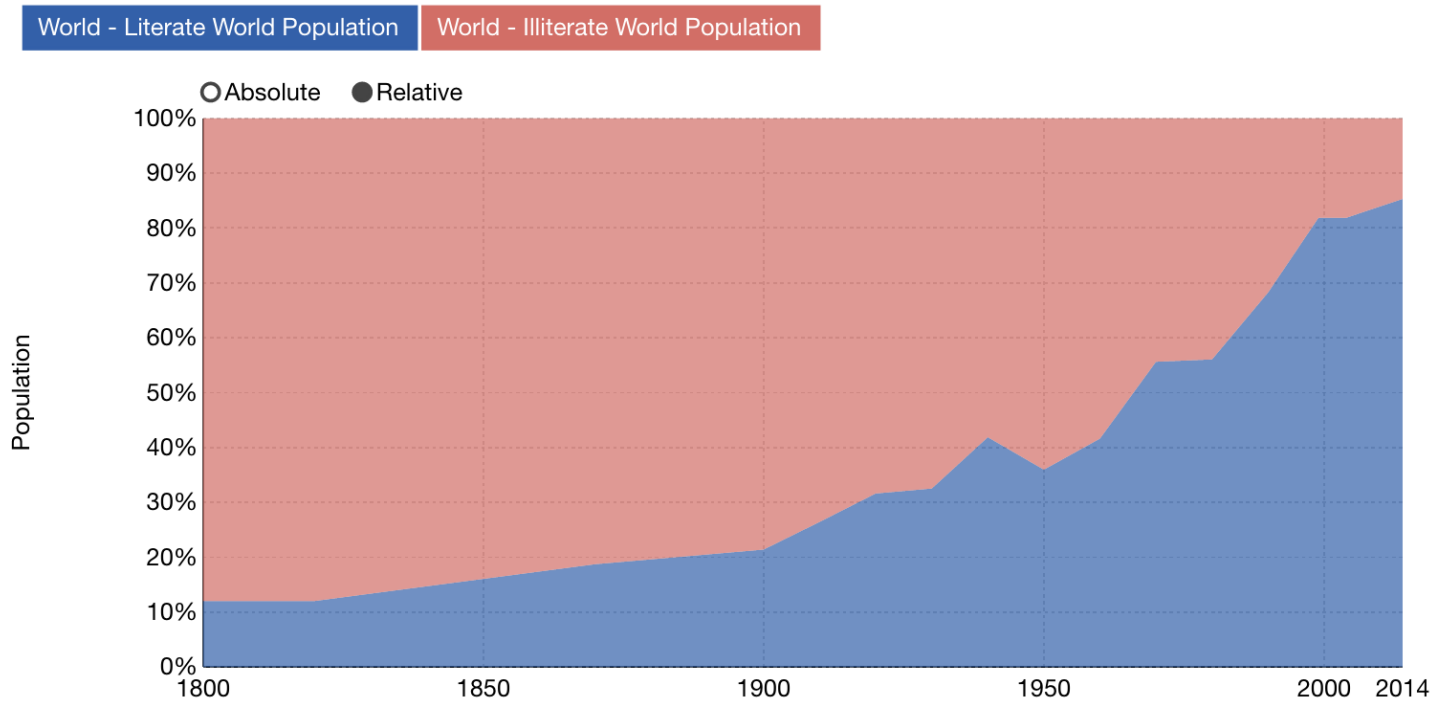
Our World in Data



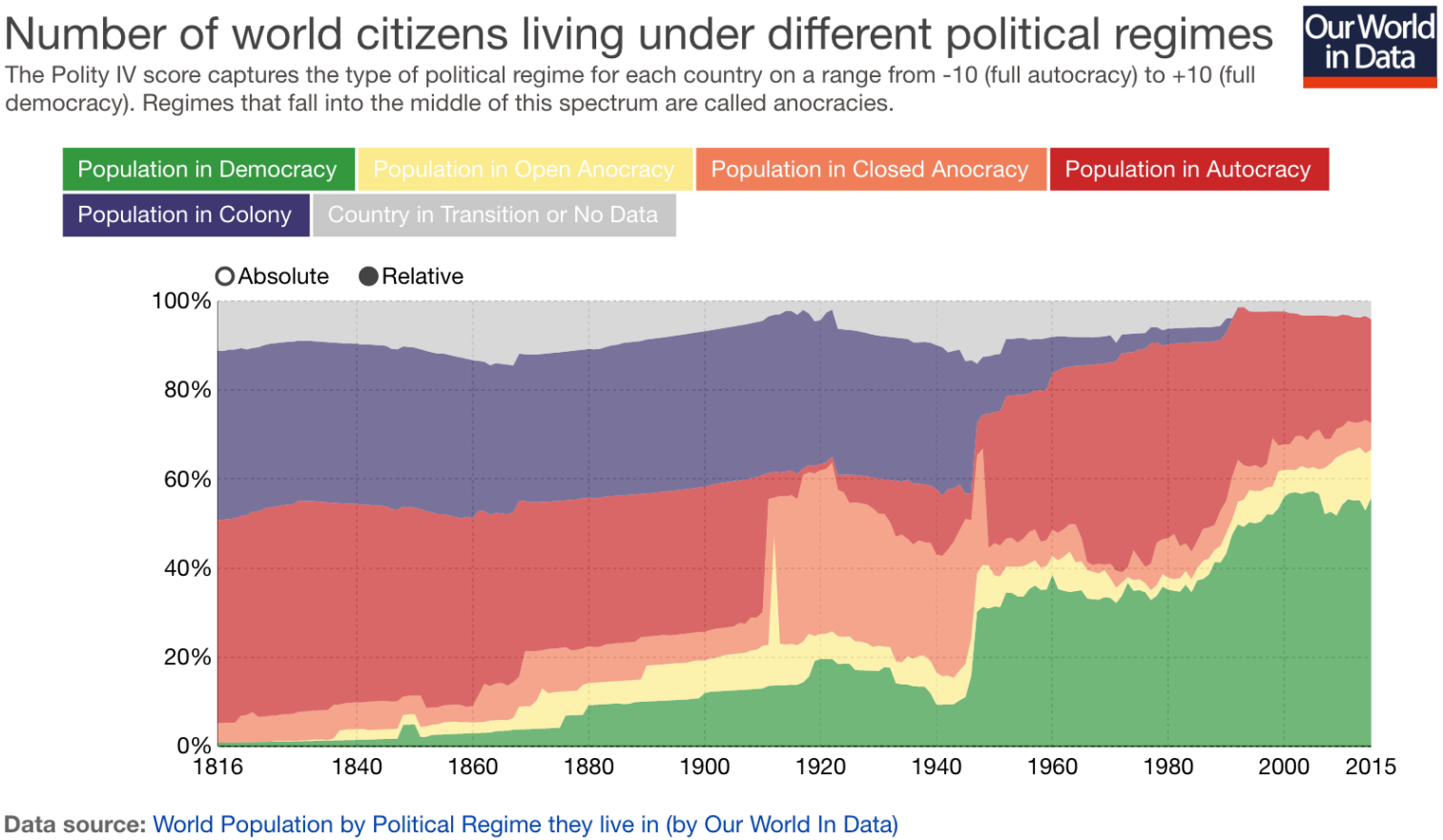
Data source: World Poverty in absolute numbers (Max Roser based on World Bank and Bourguignon and Morrisson (2002))

Literate and illiterate world population, 1800 to 2014

Our World in Data



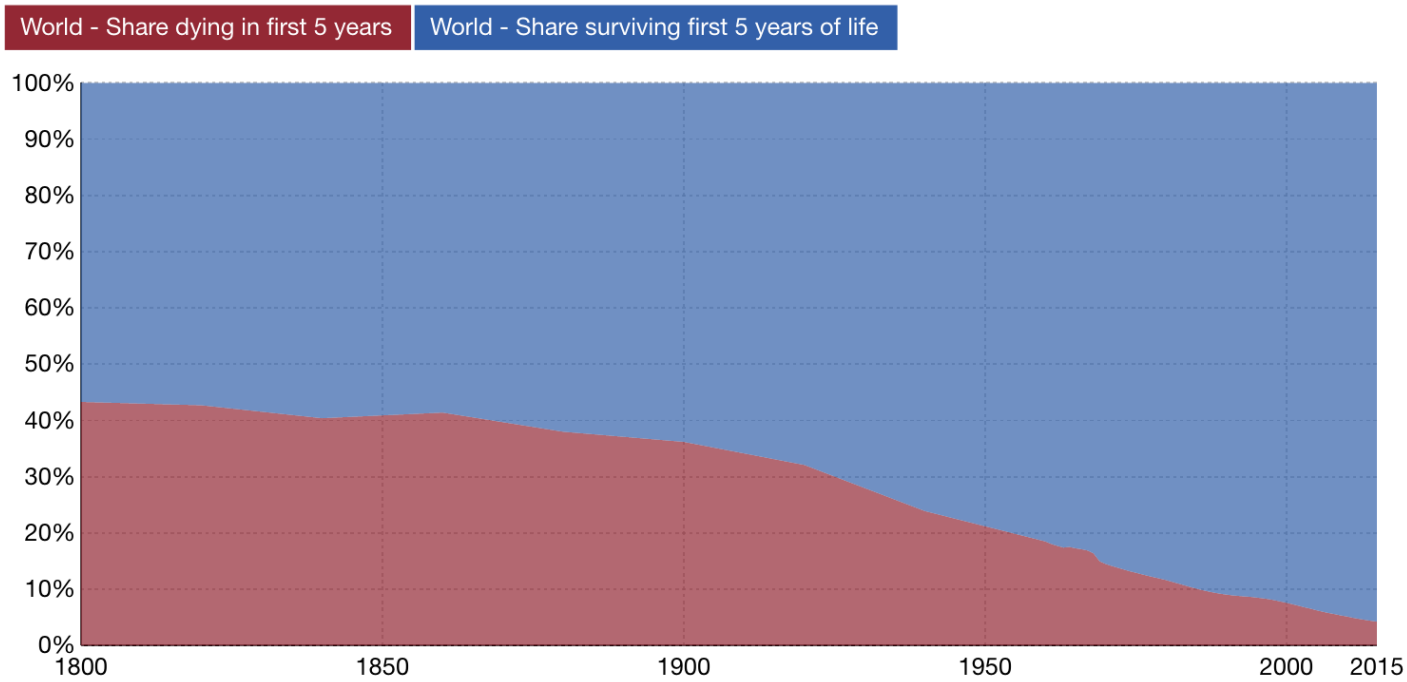
Data source: Literate World Population (Our World In Data based on OECD and UNESCO)


Number of world citizens living under different political regimes
The Polity IV score captures the type of political regime for each country on a range from -10 (full autocracy) to +10 (full democracy). Regimes that fall into the middle of this spectrum are called anocracies.
Our World in Data
Population in Democracy
Population in Open Anocracy
Population in Closed Anocracy
Population in Autocracy
Population in Colony
Country in Transition or No Data
Absolute
Relative
100%
80%
60%
40%
20%
0%
1816
1840
1860
1880
1900
1920
1940
1960
1980
2000
2015
Data source: World Population by Political Regime they live in (by Our World In Data)

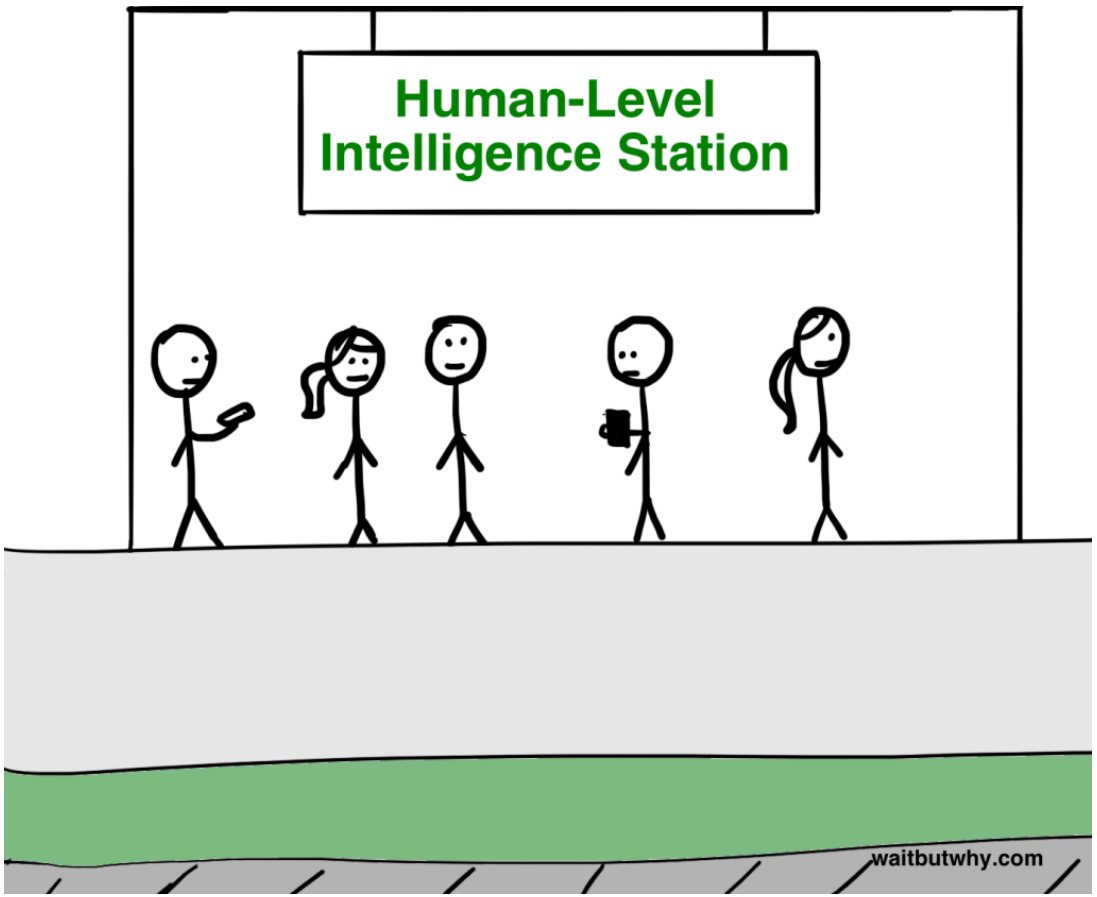
Global child mortality, 1800 to 2015

Share of the world population dying and surviving the first 5 years of life.

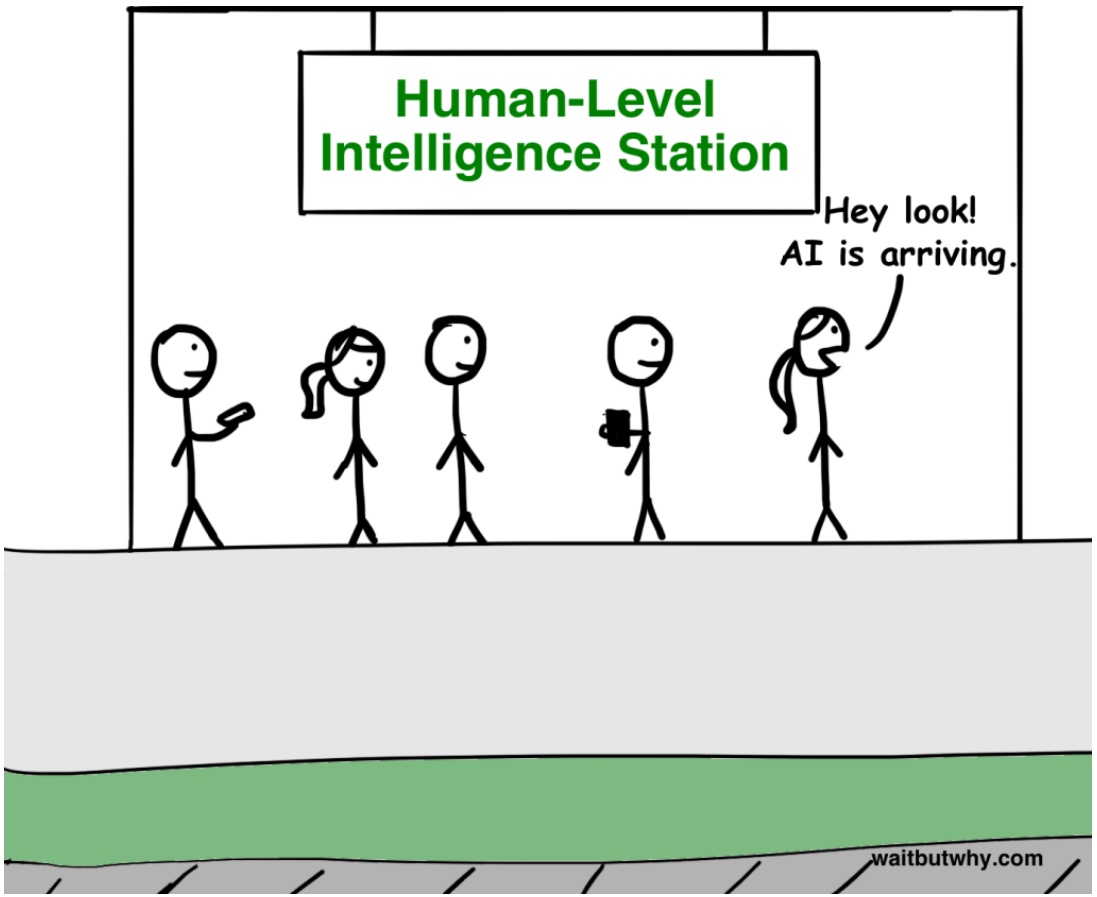
Our World in Data



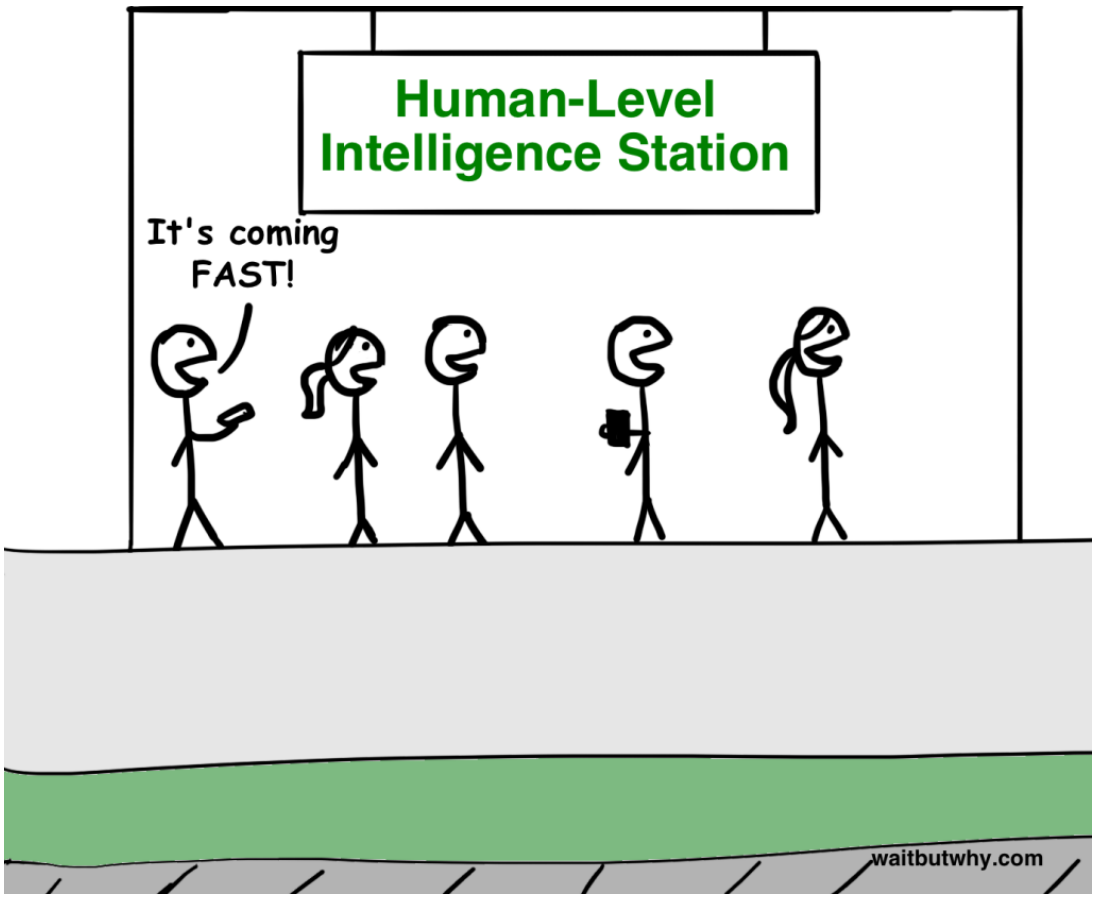
Data source: Global child mortality (since 1800) based on Gapminder and World Bank

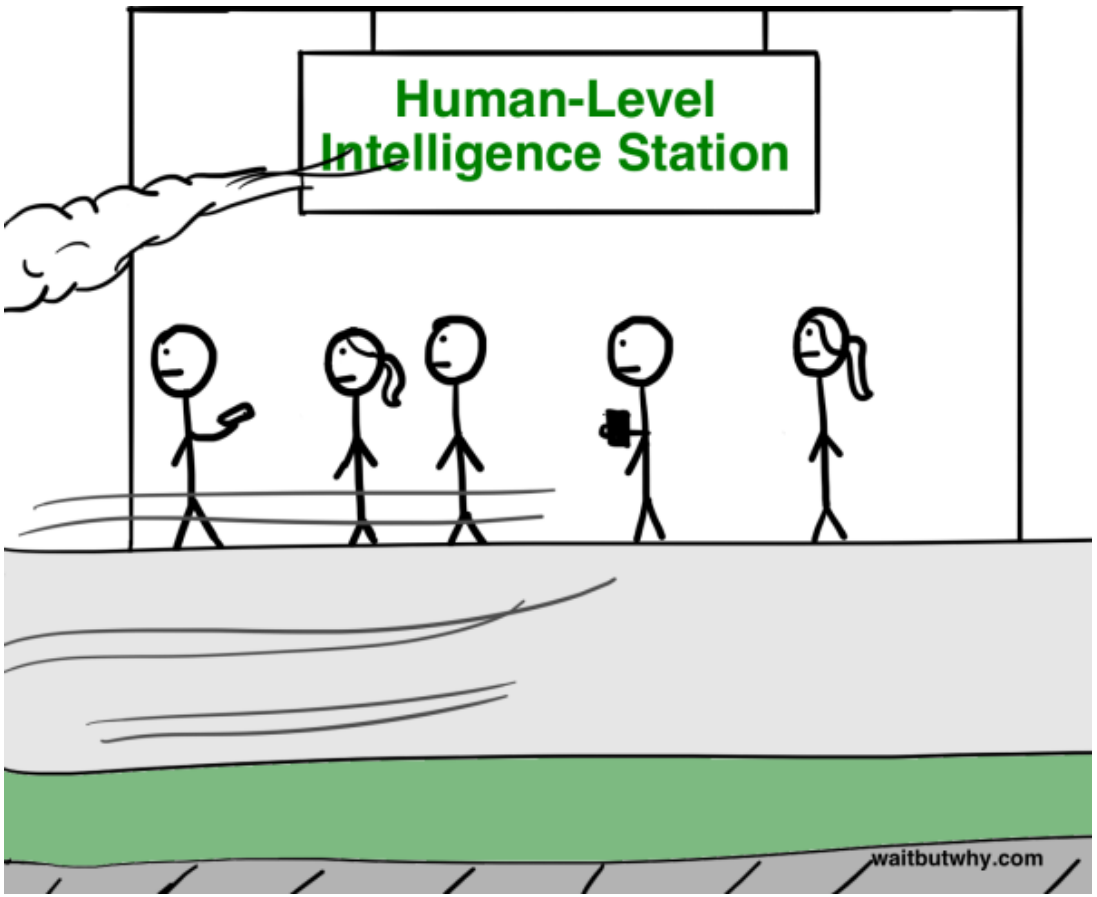

Human-Level Intelligence Station
waitbutwhy.com

Intro
DQN In Theory
DQN In Practice
Beyond DQN
AI Revolution



NYC DATA SCIENCE ACADEMY


Human-Level Intelligence Station
It's coming FAST!
waitbutwhy.com


Human-Level Intelligence Station
waitbutwhy.com

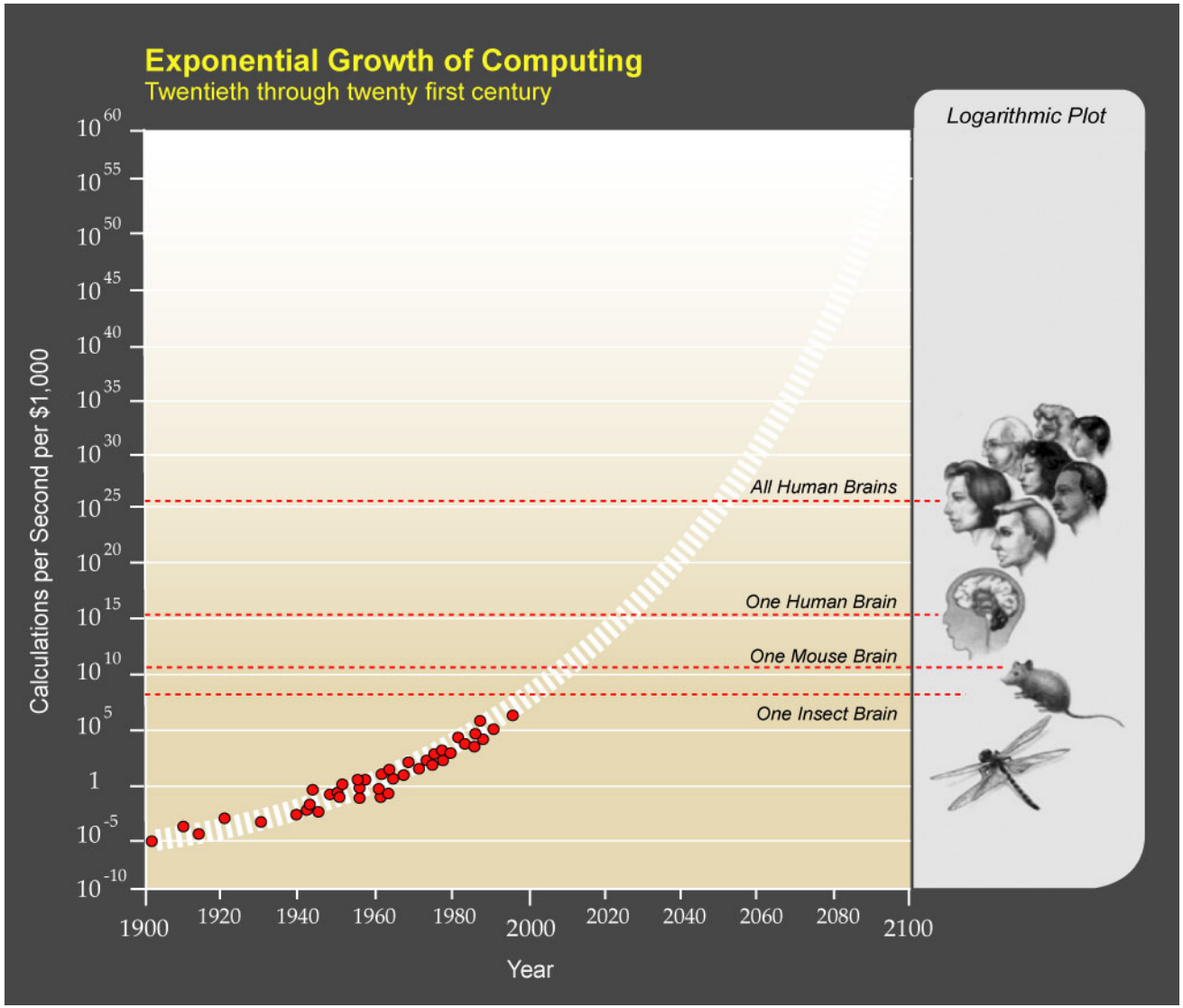
Intro

DQN In Theory

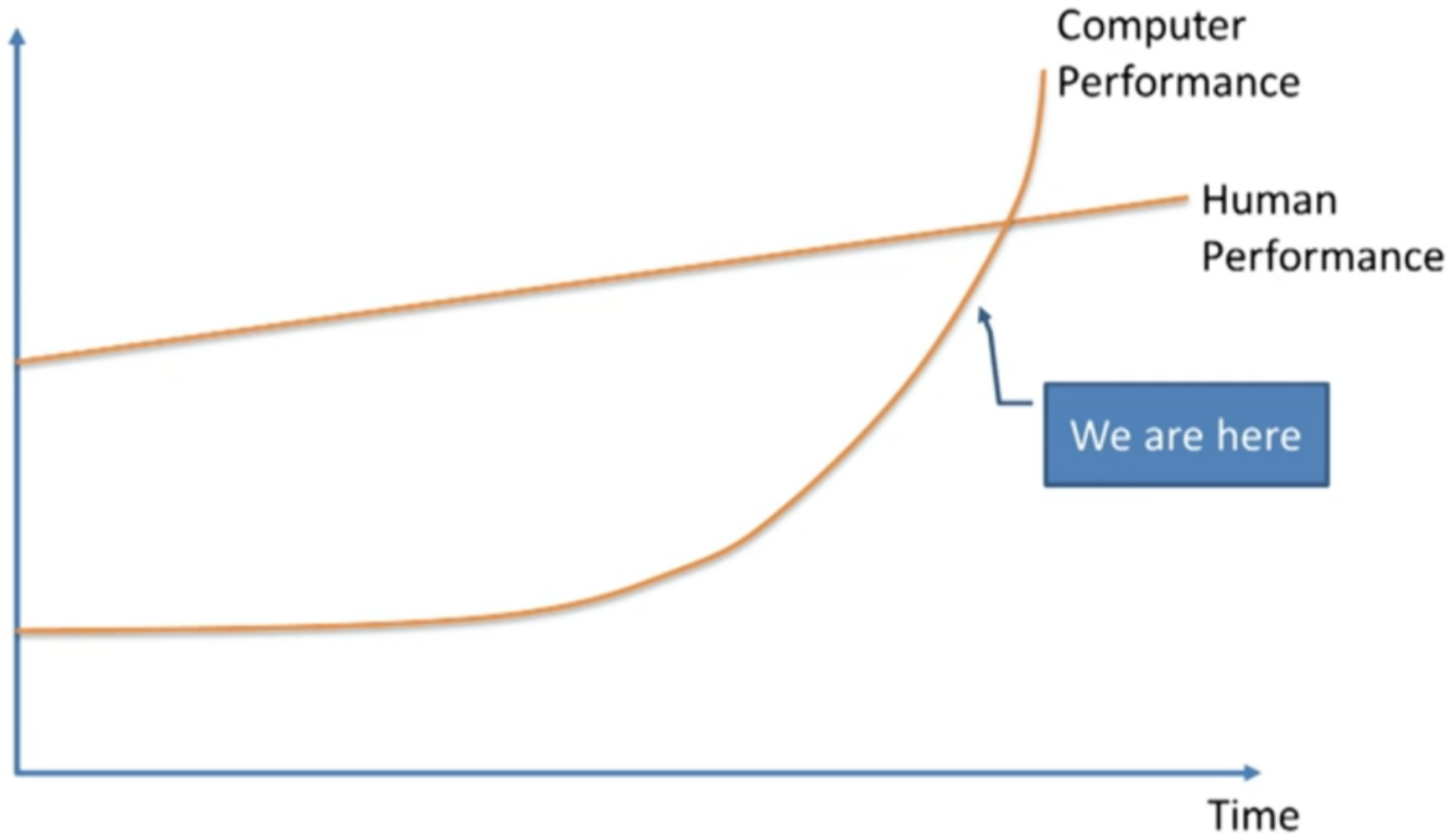
DQN In Practice

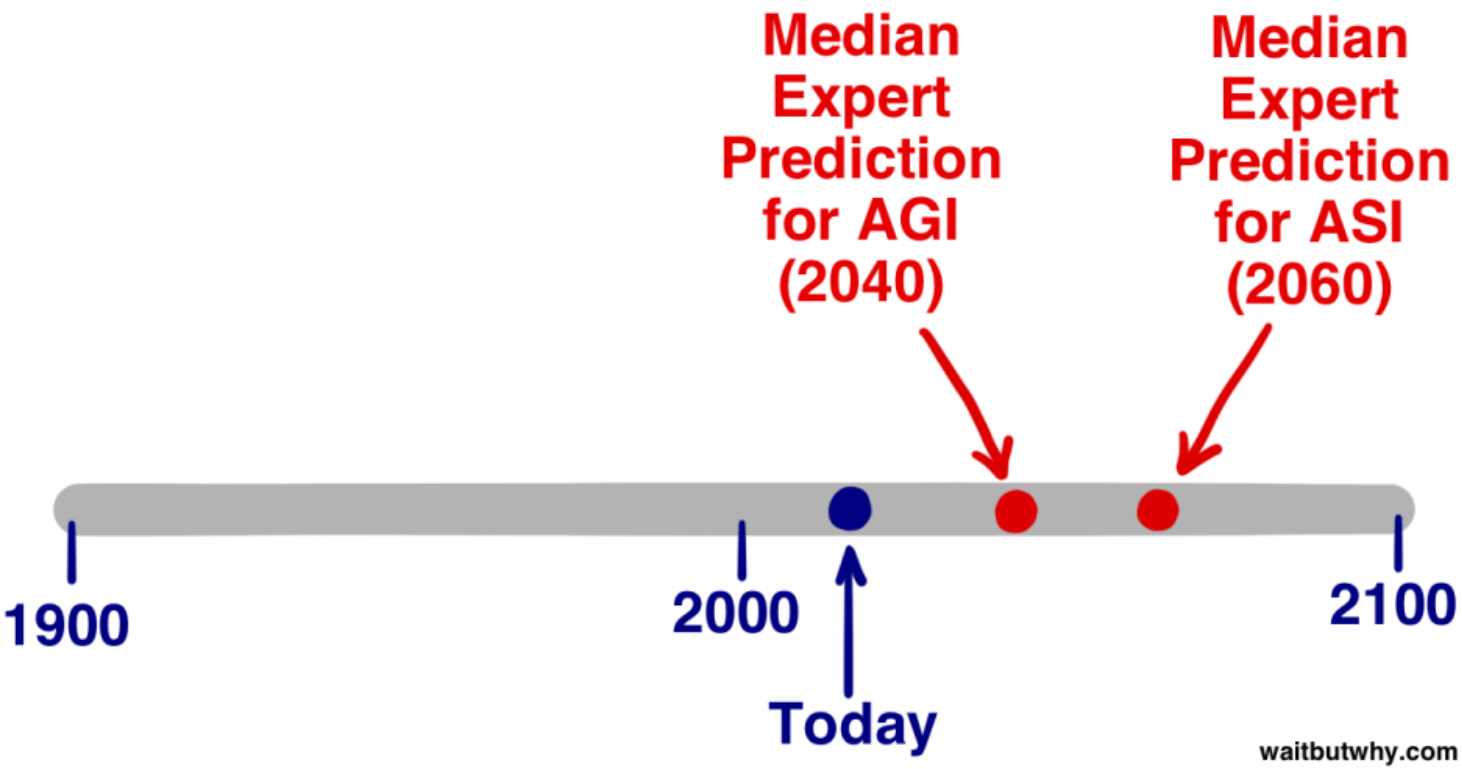
Beyond DQN

AI Revolution

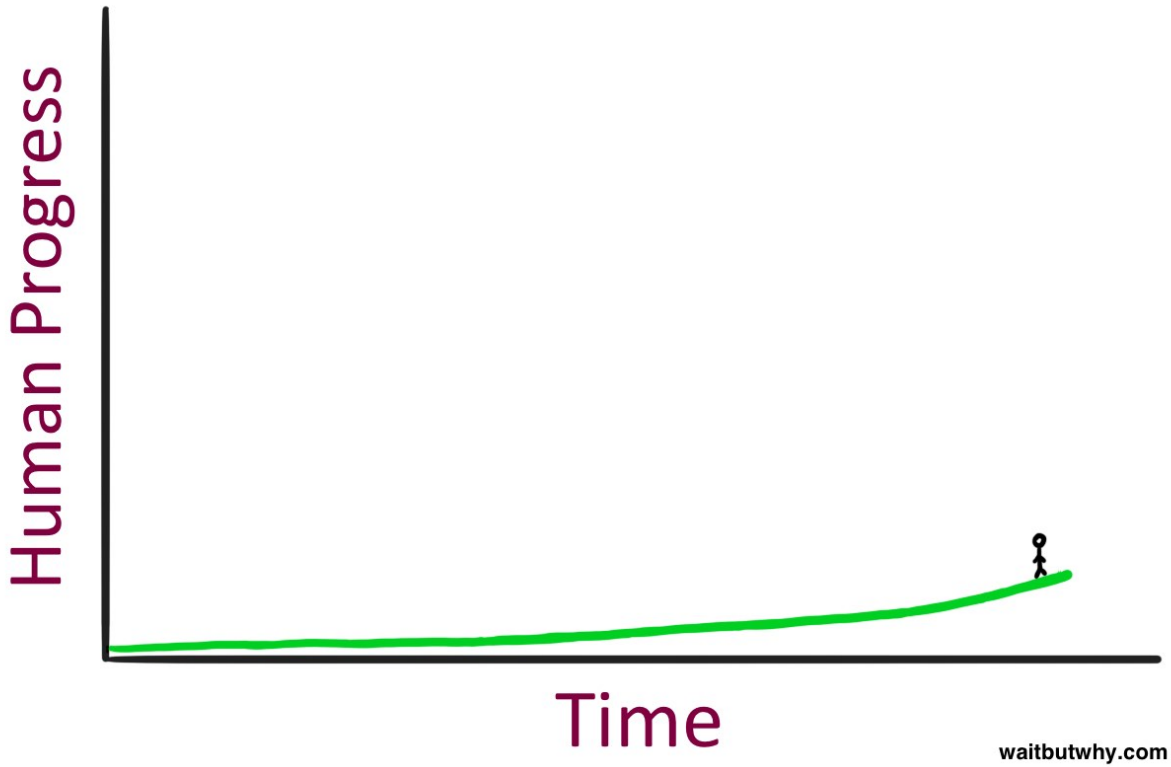


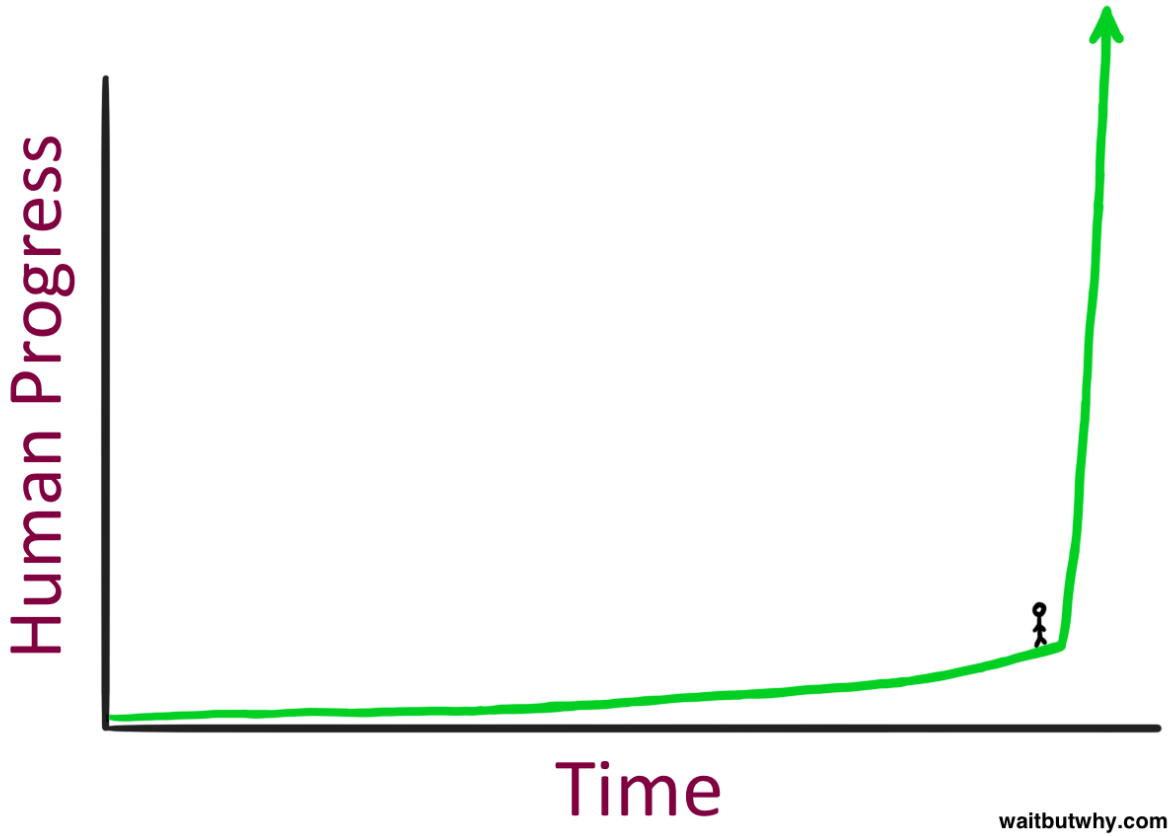
NYC DATA SCIENCE ACADEMY


Computer Performance
Human Performance
We are here
Time



NYC DATA SCIENCE ACADEMY


Human Progress
Time
waitbutwhy.com


Human Progress
Time
waitbutwhy.com