

Data Science & Artificial Intelligence

Spyros Samothrakis

Deputy Director

Institute for Analytics and Data Science

University of Essex

September 20, 2019



BETTER SCIENCE THROUGH DATA

- ▶ Thousand years ago: empirical branch
 - ▶ You observed stuff and you wrote down about it
- ▶ Last few hundred years: theoretical branch
 - ▶ Equations of gravity, equations of electromagnetism
- ▶ Last few decades: computational branch
 - ▶ Modelling at the micro level, observing at the macro level
- ▶ Today: data exploration
 - ▶ Machines create models using vast amounts of **observational** data

Manyika, James, et al. "Jim Gray on eScience: a transformed scientific method." (2009).

<http://languagelog.ldc.upenn.edu/myl/JimGrayOnE-Science.pdf>

BETTER BUSINESS THROUGH DATA

- ▶ There was that report from McKinsey...
- ▶ Urges everyone to monetise “Big Data”
- ▶ Use the data provided within your organisation to gain insights
- ▶ Had a massive impact on businesses

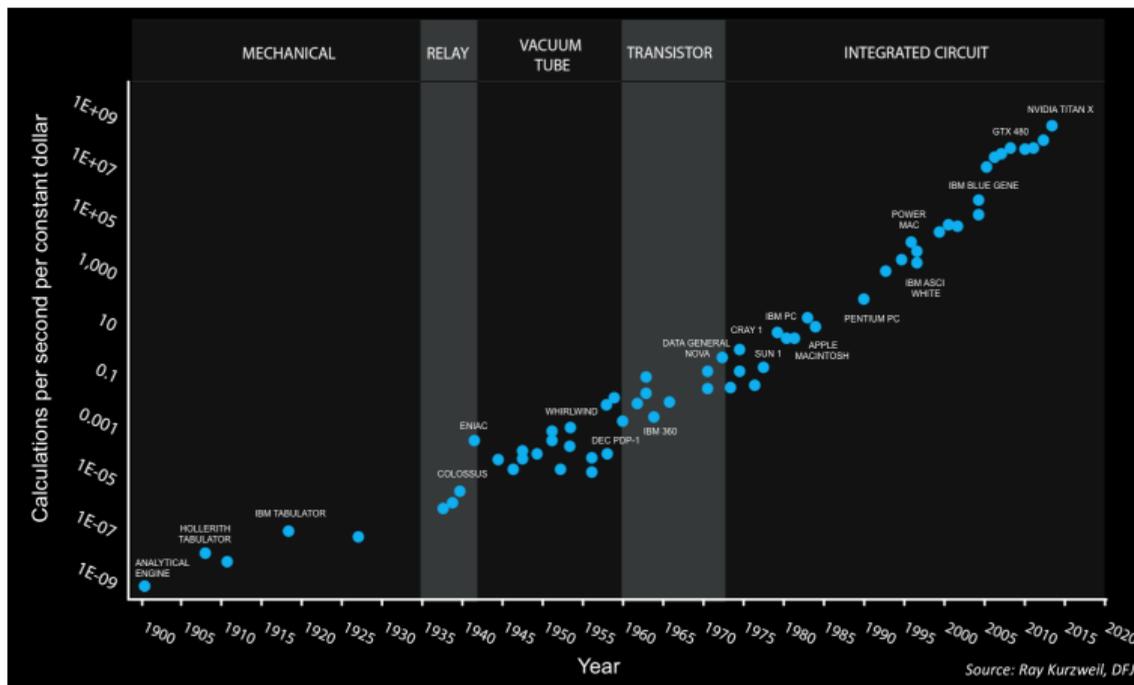


Manyika, James, et al. "Big data: The next frontier for innovation, competition, and productivity." (2011).

[http:](http://www.mckinsey.com/business-functions/digital-mckinsey/our-insights/big-data-the-next-frontier-for-innovation)

[//www.mckinsey.com/business-functions/digital-mckinsey/our-insights/big-data-the-next-frontier-for-innovation](http://www.mckinsey.com/business-functions/digital-mckinsey/our-insights/big-data-the-next-frontier-for-innovation)

MOORE'S LAW



120 years of Moore's Law.

<https://www.flickr.com/photos/jurvetson/31409423572>

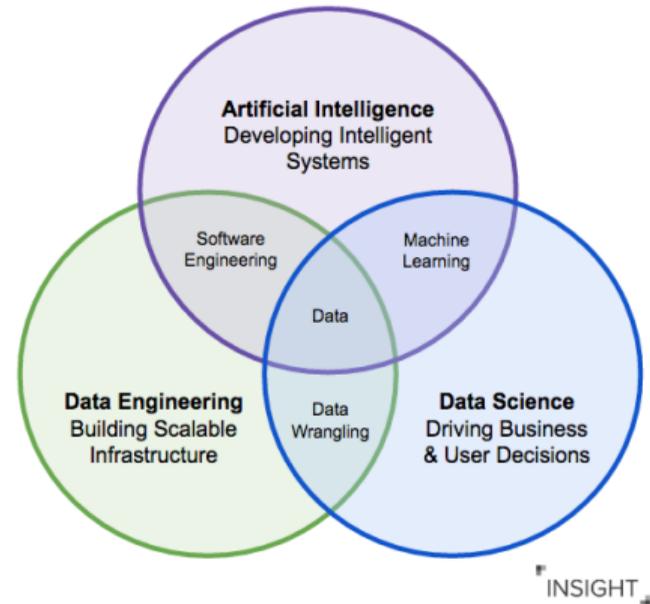
MORE IS DIFFERENT!

- ▶ The idea of emergence
 - ▶ You put people together, you go from psychology to sociology
- ▶ “Quantity changes into quality”
 - ▶ Think difference between a calculator and your laptop
- ▶ Key points: **more data, more computing power**
 - ▶ All of a sudden, certain methods (e.g. neural networks/deep learning) become viable

Anderson, Philip W. "More is different." *Science* 177.4047 (1972): 393-396.
https://www.tkm.kit.edu/downloads/TKM1_2011_more_is_different_PWA.pdf

ARTIFICIAL INTELLIGENCE & DATA SCIENCE

- ▶ An umbrella term where people fit whatever they think is roughly related to intelligence
 - ▶ Almost all of machine learning
 - ▶ Predictive modelling
 - ▶ Reinforcement learning
 - ▶ Network science
 - ▶ Classic AI (i.e. logic, ontologies)
 - ▶ Game theory



Karnowski, Jeremy. "How AI Careers Fit into the Data Landscape." (2015).
<https://blog.insightdatascience.com/how-emerging-ai-roles-fit-in-the-data-landscape-d4cd922c389b>

INSTITUTE FOR ANALYTICS AND DATA SCIENCE (IADS)

- ▶ A loose group of around 80 academics that have expertise in working with data
- ▶ Around 9 dedicated staff members and growing
- ▶ We have experts from all fields of AI/data processing
- ▶ Links to other data-related groups within the university (e.g. ISER, ESRC BLG, UKDA, CSEE)

NATURAL LANGUAGE PROCESSING

- ▶ Chatbots
 - ▶ Agents that can communicate via natural language with the user
 - ▶ Can greatly help with front line services
 - ▶ Amazon has a yearly prize
- ▶ Sentiments analysis
 - ▶ Great when launching new products/services (how did my audience respond)
- ▶ Tons of other uses

The Alexa Prize

<https://developer.amazon.com/alexaprize>

Signal Media

<https://signalmedia.co/>



CREATIVE ARTIFICIAL INTELLIGENCE

- ▶ New music
- ▶ New games
- ▶ New recipes
- ▶ New art
- ▶ New products

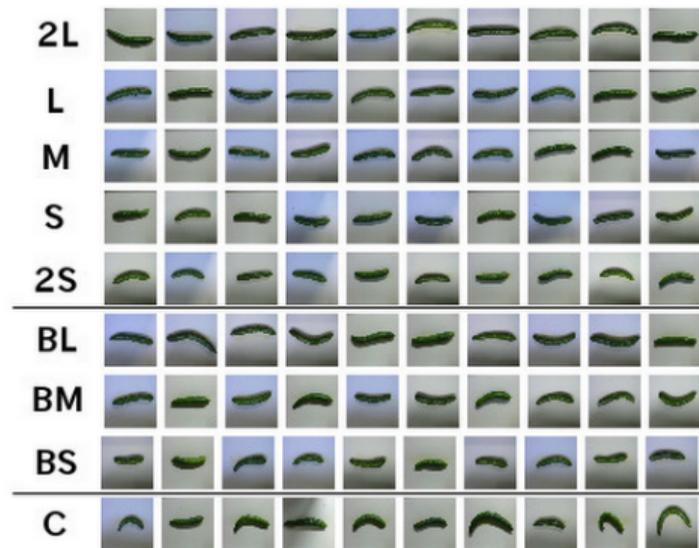


Supercharging Style Transfer

<https://research.googleblog.com/2016/10/supercharging-style-transfer.html>

COMPUTER VISION

- ▶ One of the original AI themes
- ▶ Very data-hungry
- ▶ Expertise in CSEE/IADS



How a Japanese cucumber farmer is using deep learning and TensorFlow
<https://cloud.google.com/blog/products/gcp/how-a-japanese-cucumber-farmer-is-using-deep-learning-and-tensorflow>

AI FOR POLICY MAKING

- ▶ A few groups within the university work on this
 - ▶ Causal inference
 - ▶ Reinforcement learning
- ▶ “Given the data that I hold, what is the best next action?”
- ▶ This can range from real time monitoring and acting loops
 - ▶ e.g. a fully automated greenhouse, where a vision system monitors everything, actuators/robots/humans take care of plants
 - ▶ “Viable system”
- ▶ ...to high level business planning

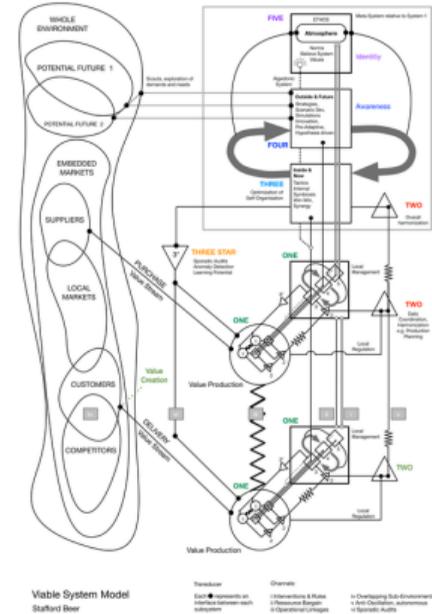
BIG DATA VS SMALL DATA

- ▶ Genuine “big data” problems are rare
- ▶ Most problems are *small data* problems
- ▶ Most businesses work with data sizes up to a GB
- ▶ If your data can fit a laptop (i.e. <1TB), you don't really have big data

statistics, data generation

MOVEMENT

- ▶ We are moving from...
 - ▶ *Reactive to predictive*
 - ▶ *Generic to personalised*
 - ▶ *Predictions to causal inference*
 - ▶ *Black boxes to greater transparency*
- ▶ A resurrection of ideas coming from cybernetics
- ▶ Thank you!



Viable system model

https://en.wikipedia.org/wiki/Viable_system_model