The need for Data Analysis

Madrid, Octubre 2021

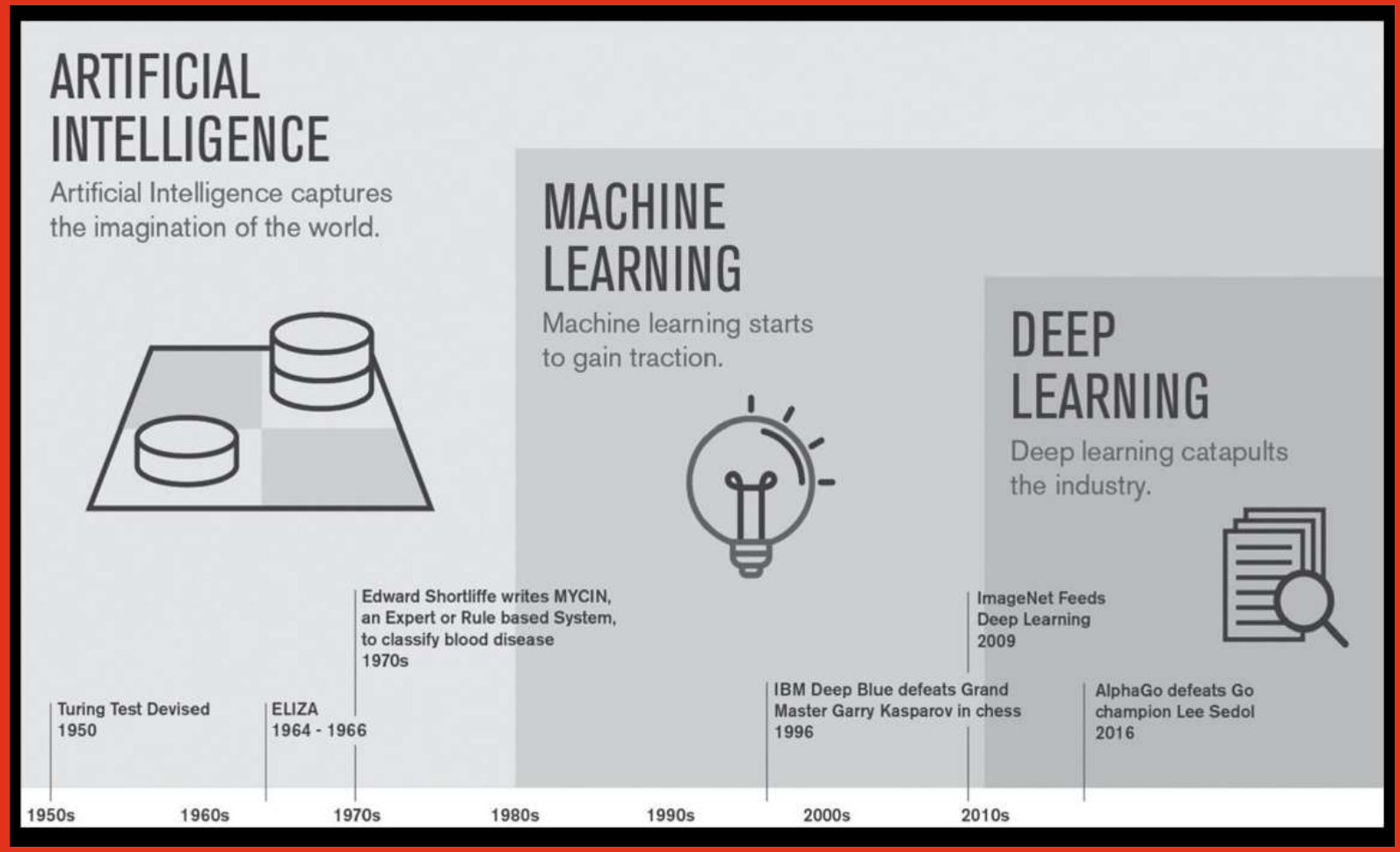
Data Science area and profiles

Intro to the current picture of the bursting of data

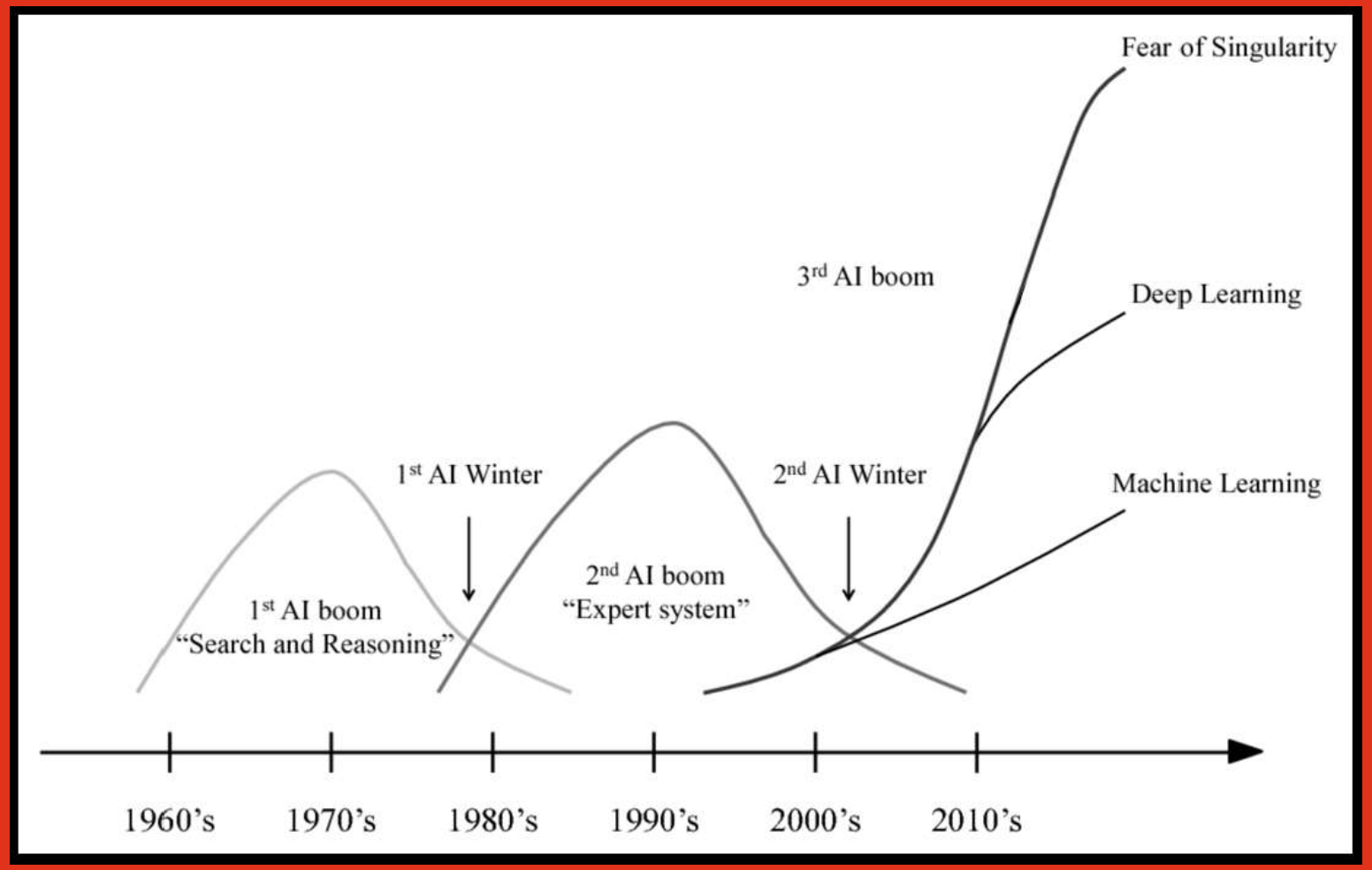
Digital profiles in the industry

How far data analysis may bring us

AI – ML – DL



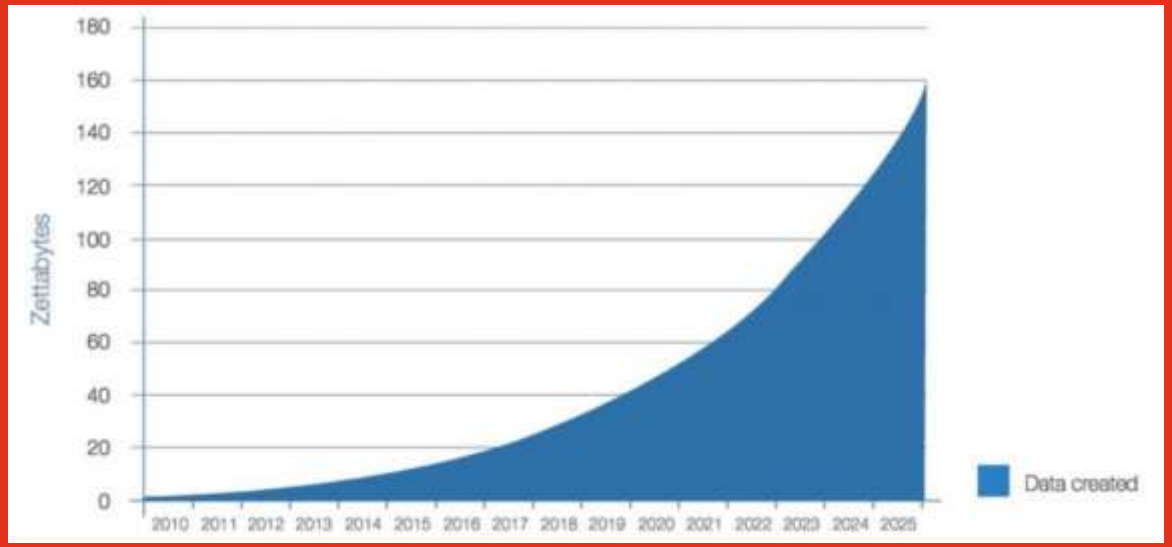
Intro to the current picture of the bursting of data



Previously on This is data… What happened in 2007?

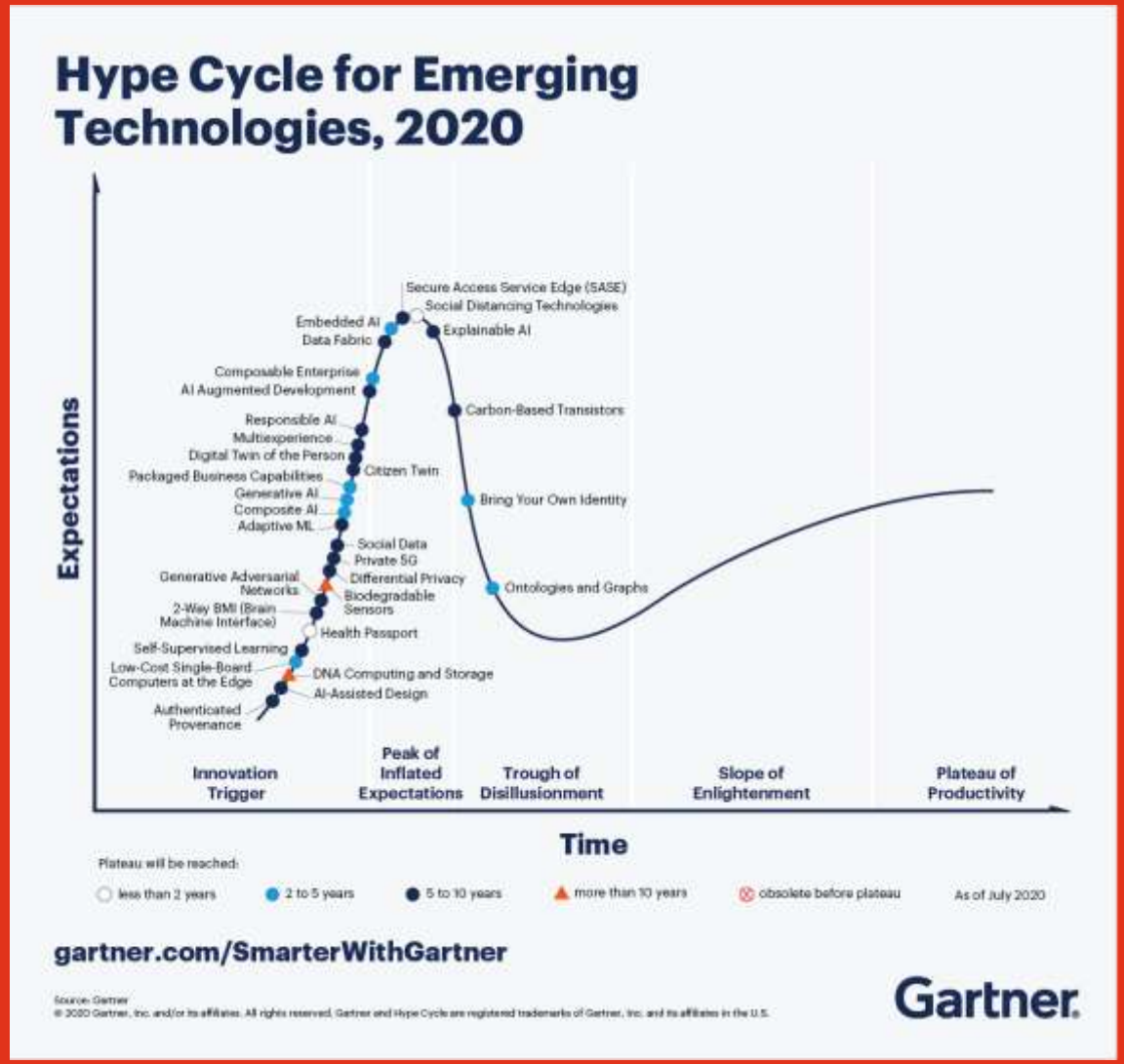
1. Apple presents the 1st Iphone to the world
2. Hadoop makes Big Data possible (from a Google paper)
3. GitHub appears (to let technical teams work together in the same project with different versions)
4. Change.org appears as well
5. Twitter begins to scale up!
6. Facebook begins to scale up!
7. Google presents Android
8. AirBnB posted its first house for renting

And from 2007 forwards...



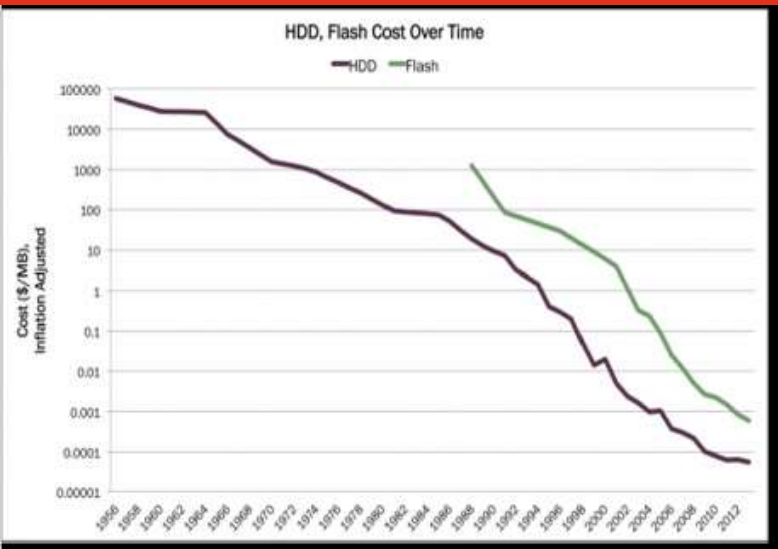
1 kilobyte	1,000
1 megabyte	1,000,000
1 gigabyte	1,000,000,000
1 terabyte	1,000,000,000,000
1 petabyte	1,000,000,000,000,000
1 exabyte	1,000,000,000,000,000,000
1 zettabyte	1,000,000,000,000,000,000,000

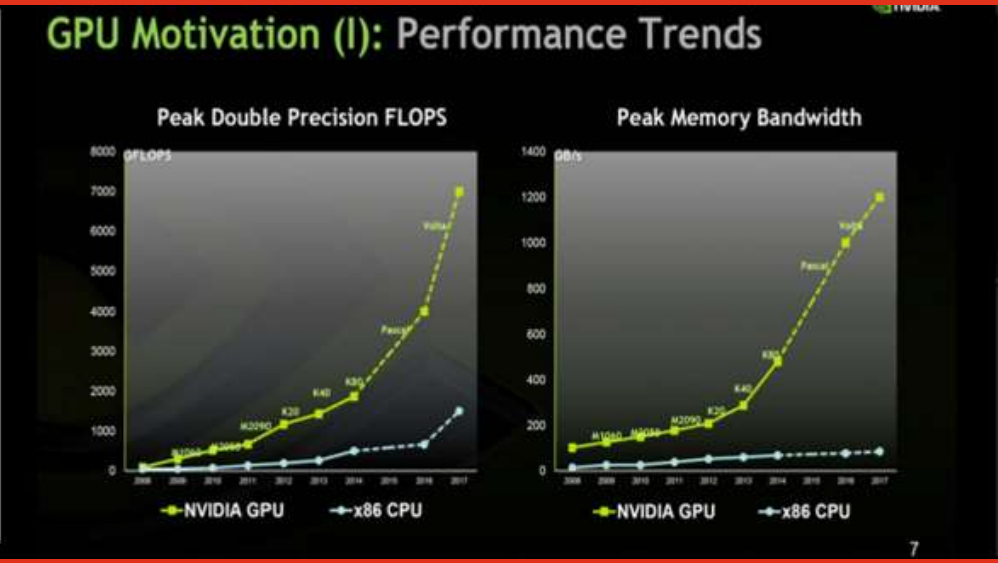
Current trends

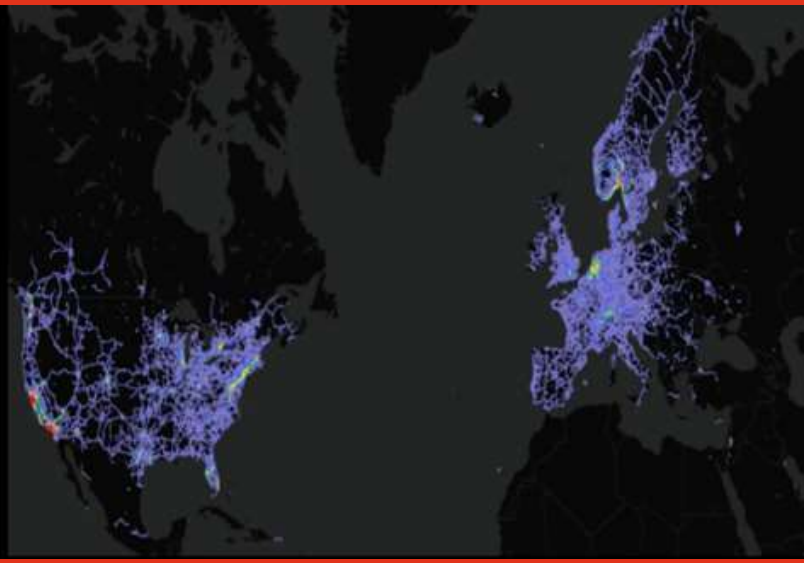


Loading data is cheaper and cheaper

Silicon Valley is named after the silicon, which comes from the sand. There is sand everywhere, it is what the chips are made of and it is extremely cheap!



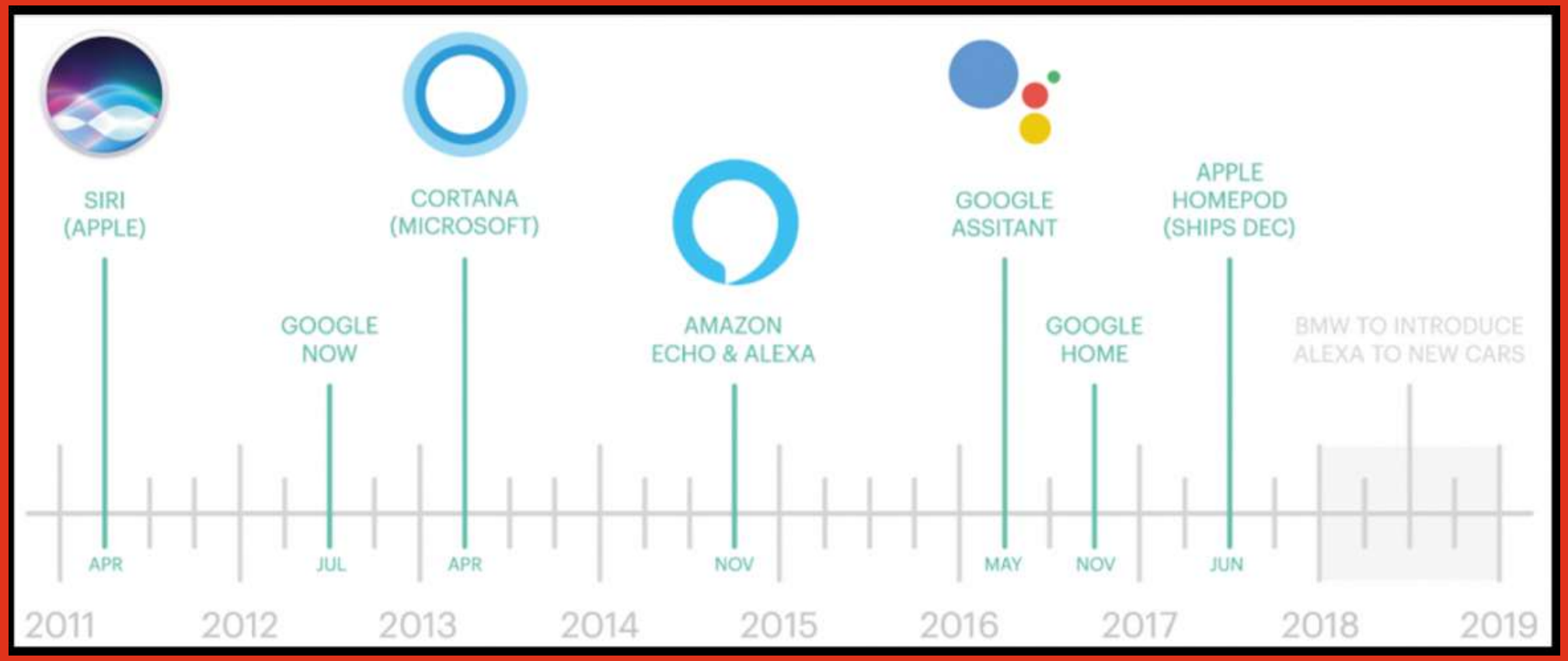




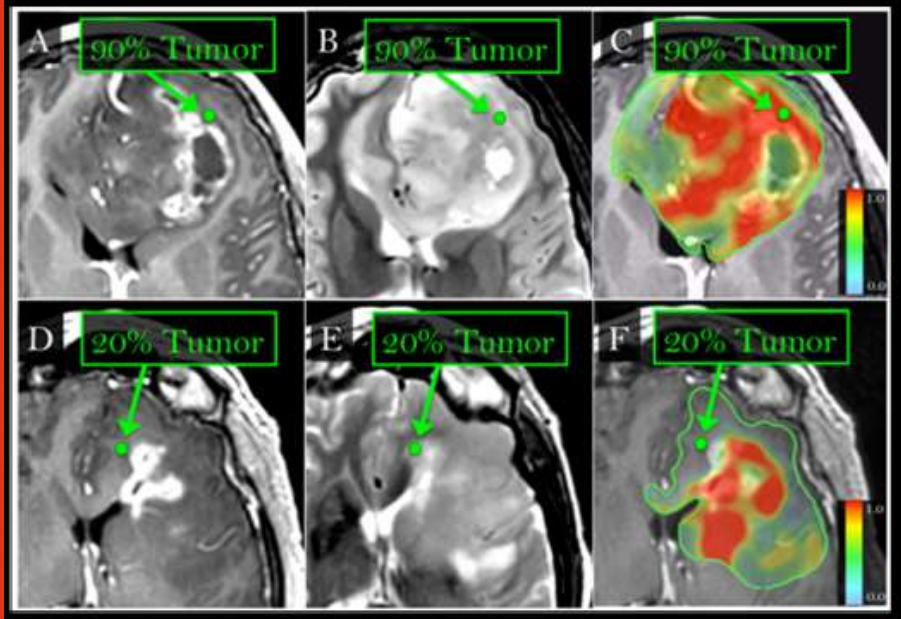
I do not work for free…or do I?

You tought ML algorithms to recognise objects. Thank you!

AI is coming home!



When AI can work better than humans



13 MAY 2019 NEWS

AI surpasses humans in predicting heart attack and death

By **Chloe Kent**

SHARE

Through repetition and adjustment, machine learning can exploit large amounts of data and identify complex patterns that may not be evident to humans. Credit: Shutterstock

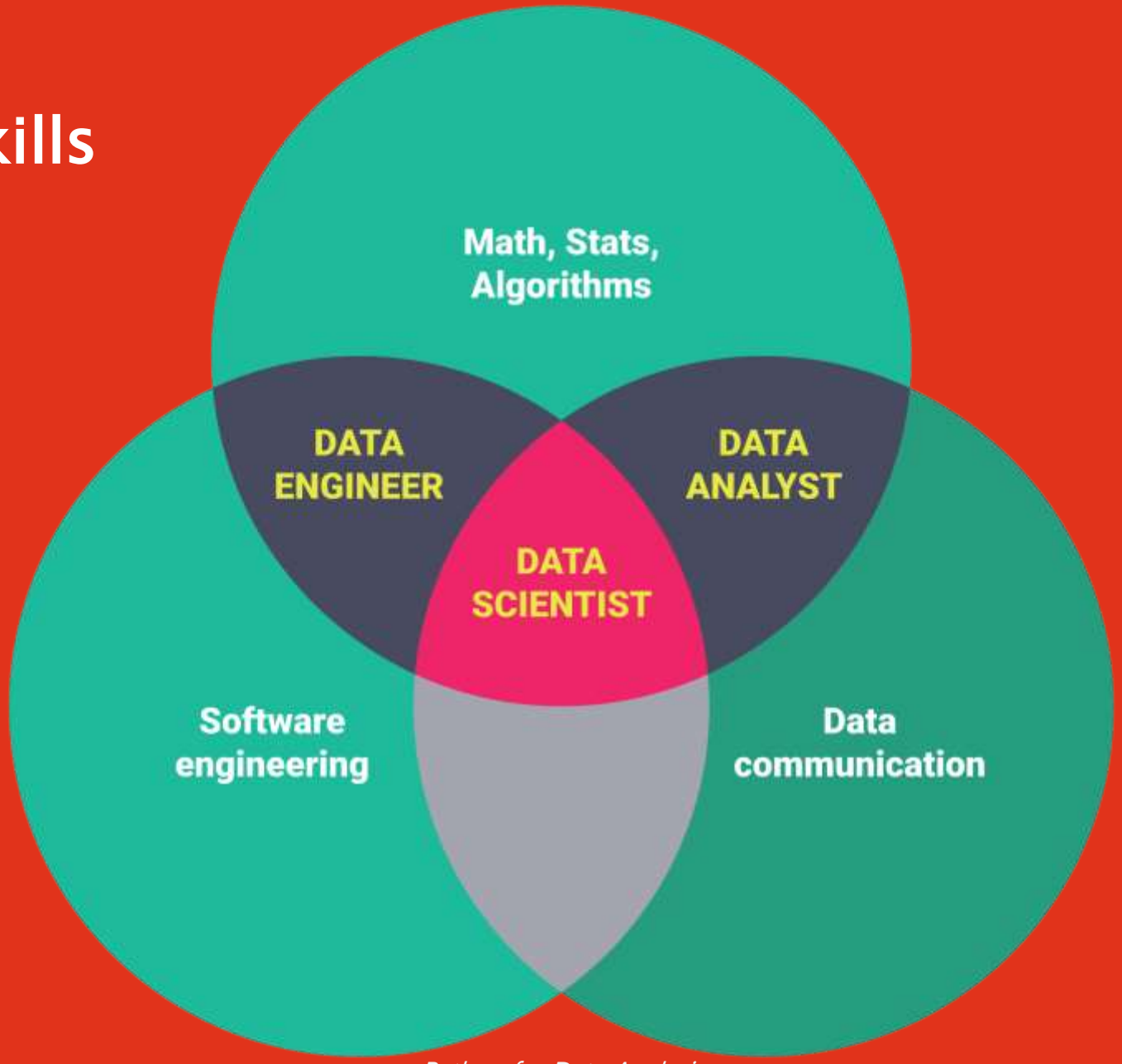
An algorithm has learned how to identify imaging patterns correlating to heart attack and death in cardiac patients and can predict the occurrence of these events with superior accuracy to human doctors, according to a study presented at 2019's International Conference on Nuclear Cardiology and Cardiac CT (ICNC) in Lisbon, Portugal.

In current medical practice, doctors will use risk scores to make treatment decisions for their cardiac patients. These are based on a series of variables like weight, age and lifestyle. However, they do not always have the desired levels of accuracy. A 2015 study published in *Annals of Internal Medicine* found that four of five standard risk scoring tests overestimated the risk of cardiovascular disease in a control group, which researchers suggested could lead to adverse outcomes such as the prescribing of statin therapies to patients who do not need them.

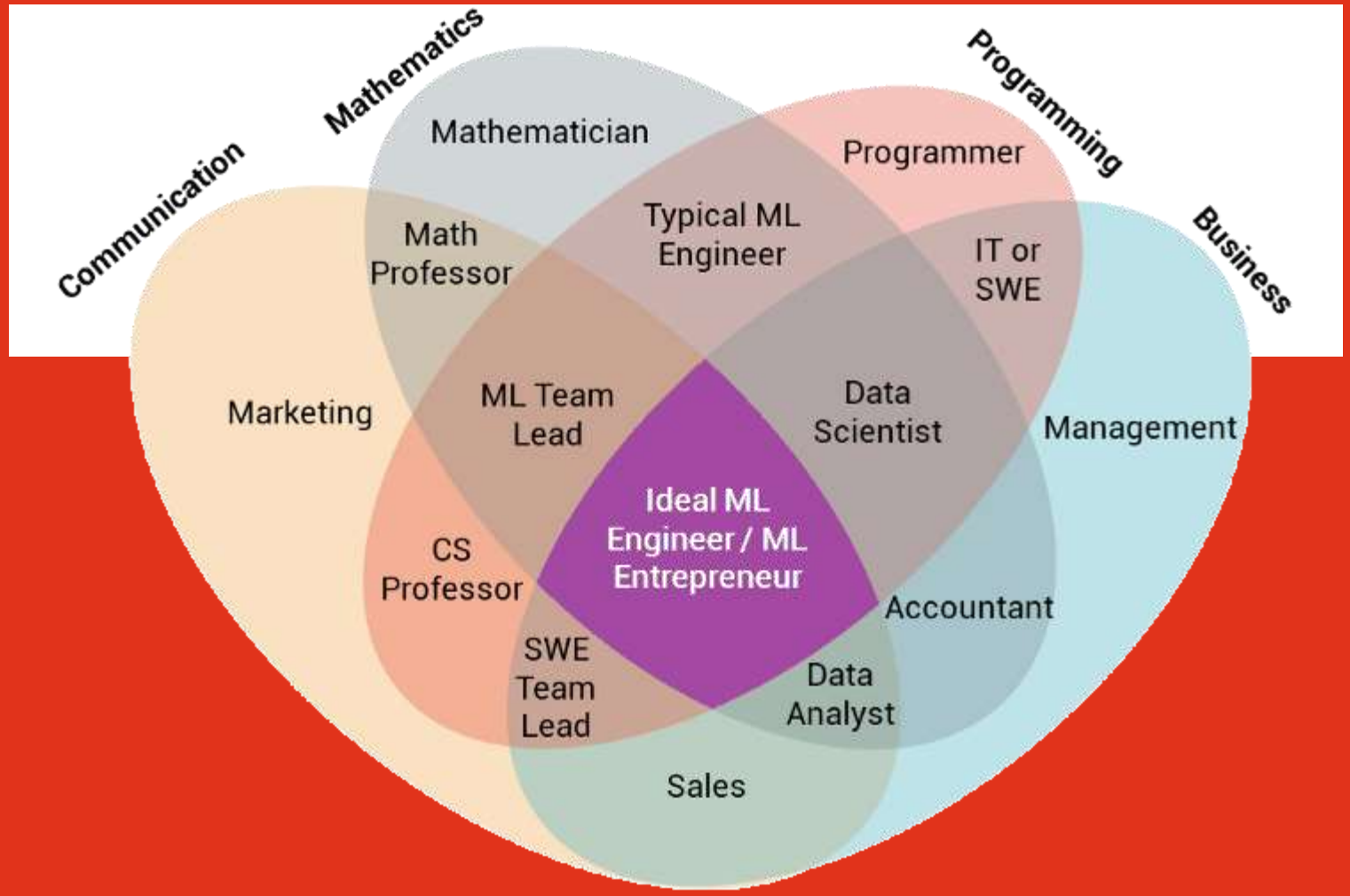
Profiles in Data Areas

1. Data Architecture: Which systems? Dessign of the HW & SW
2. Data Analysis: They speak the language of business and technical staff
3. Data Science: They transform data into information (value)
4. Data Engineering: ETLs (extract, transform and load data)
5. Data Visualization: Mapping output for business reports
6. Data Governance: Uniqueness of the data, validation of sources, permissions
7. DevOps: Putting everything in Production, deploying in real scenarios
8. Data Security: Defending our data

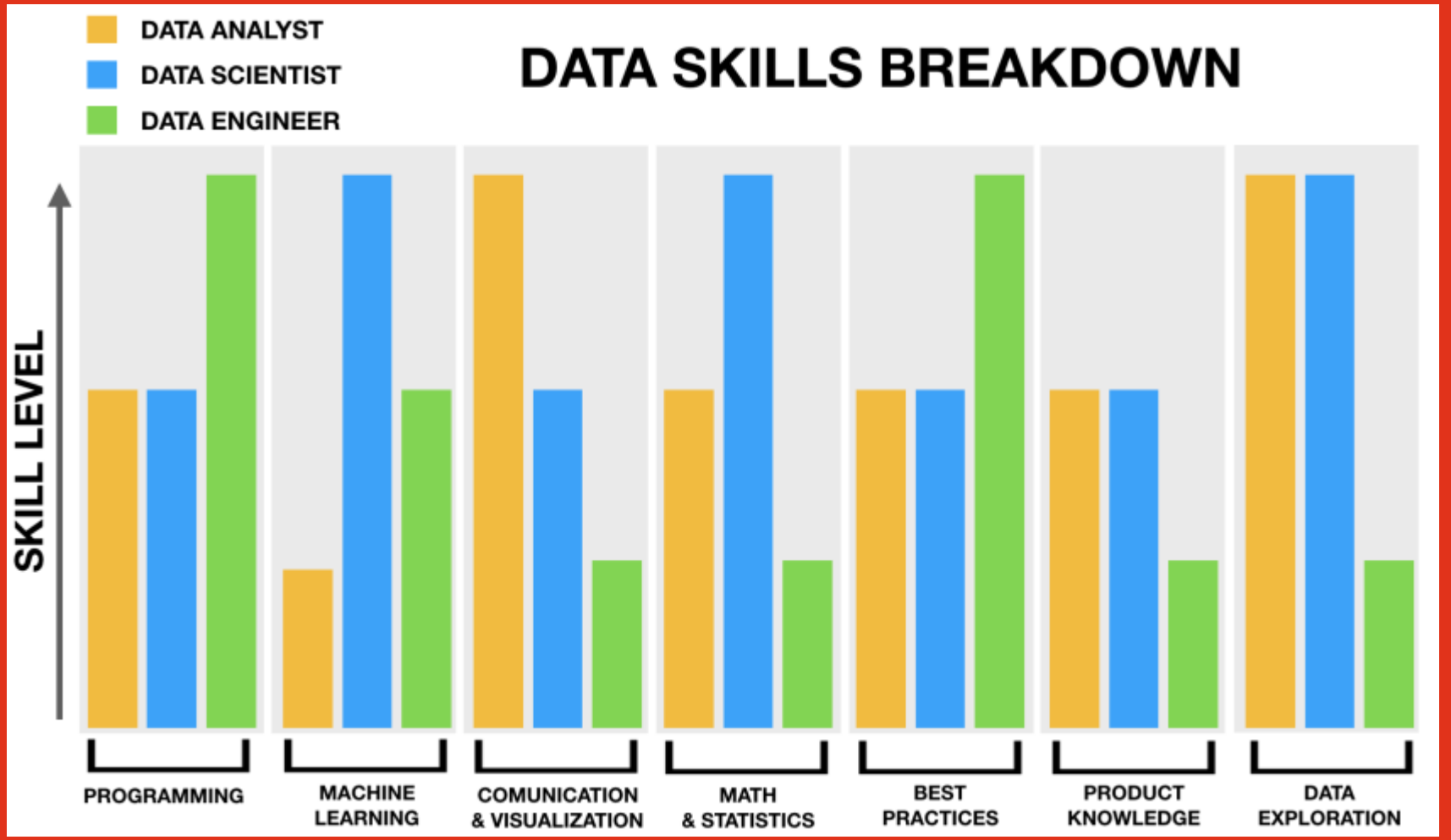
Data Skills



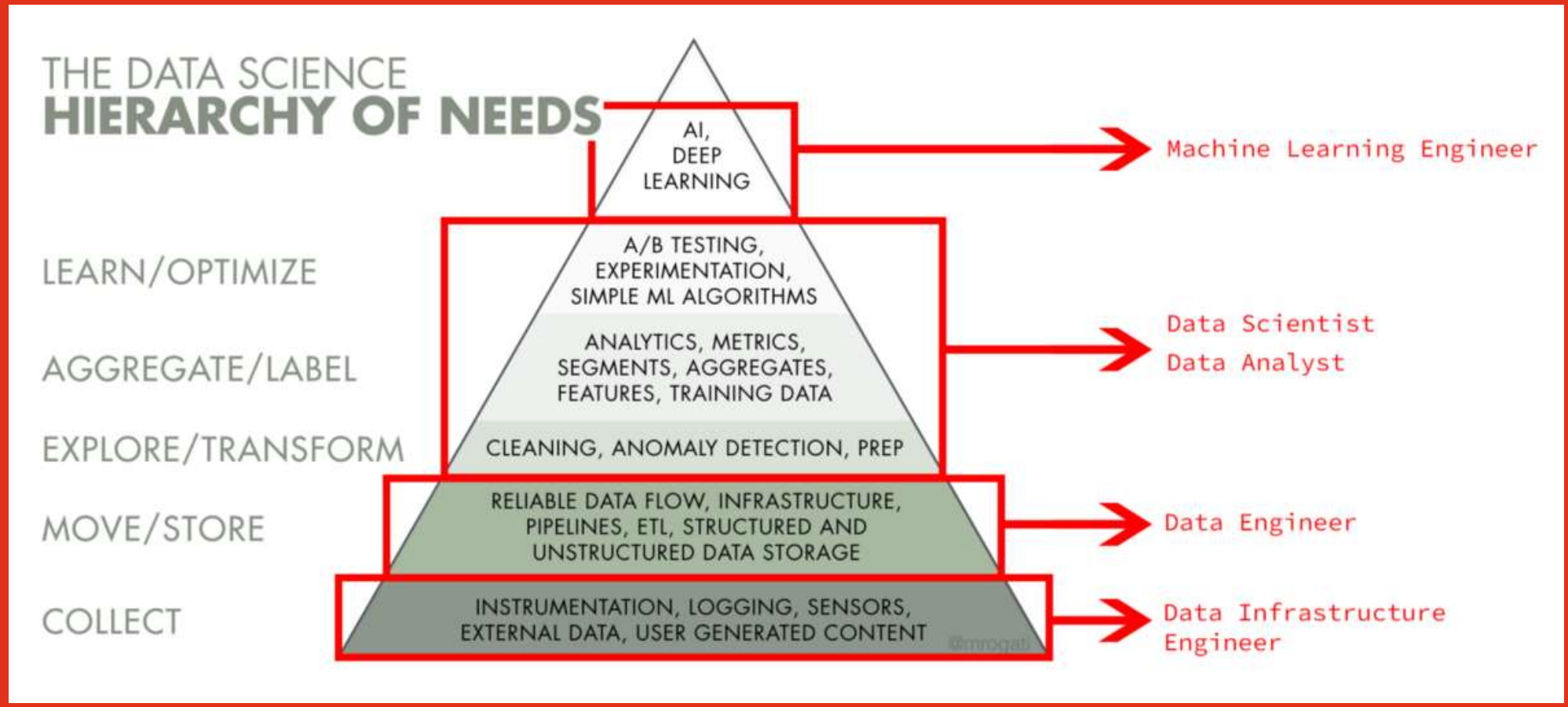
Data Skills in detail



Summing up the skills



Intro to the current picture of the bursting of data



Data Security

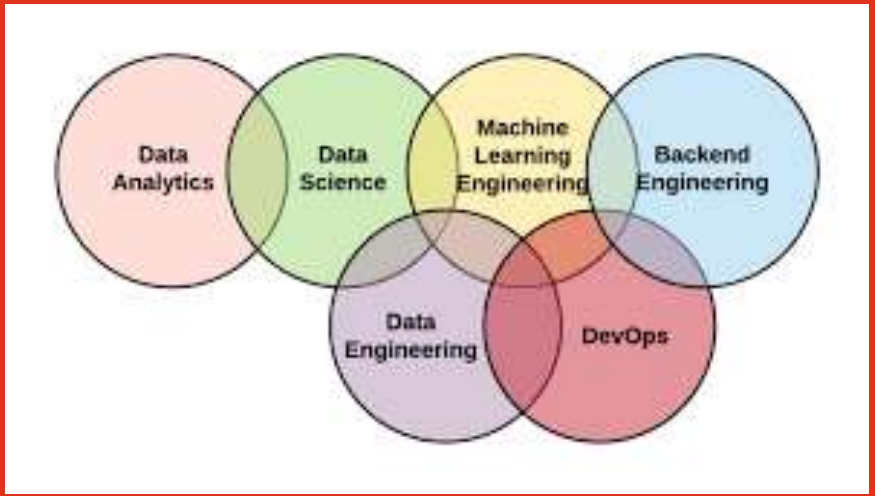
Data Governance

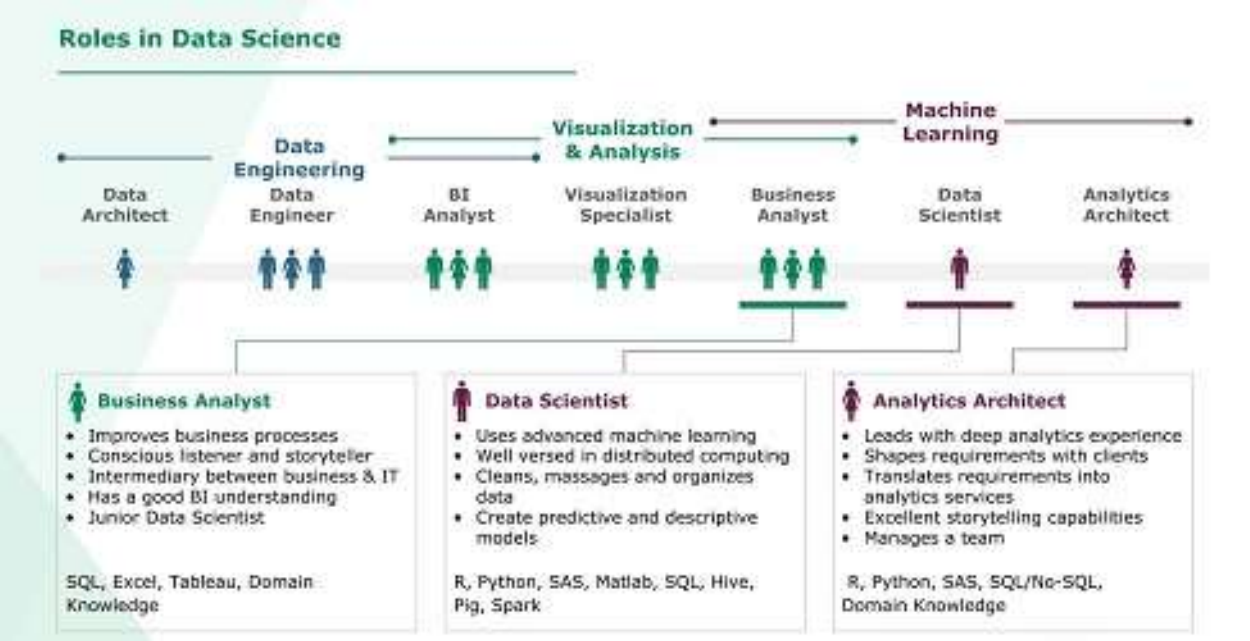
Arch → DE → DA → DS → DevOps

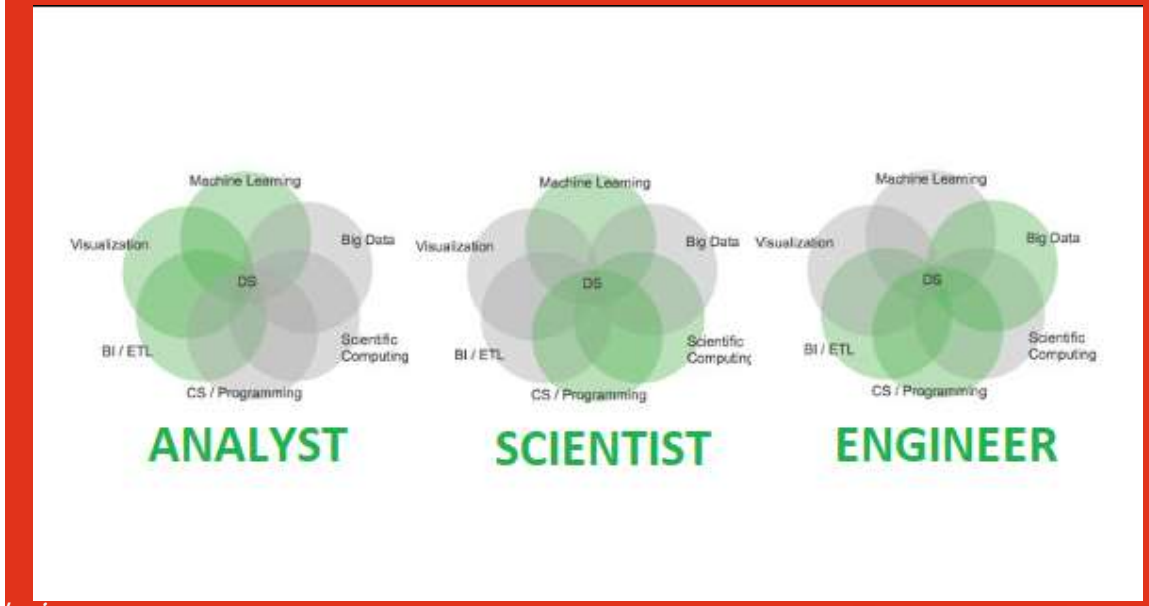
Testing Environment

PreProduction Environment

Production Environment

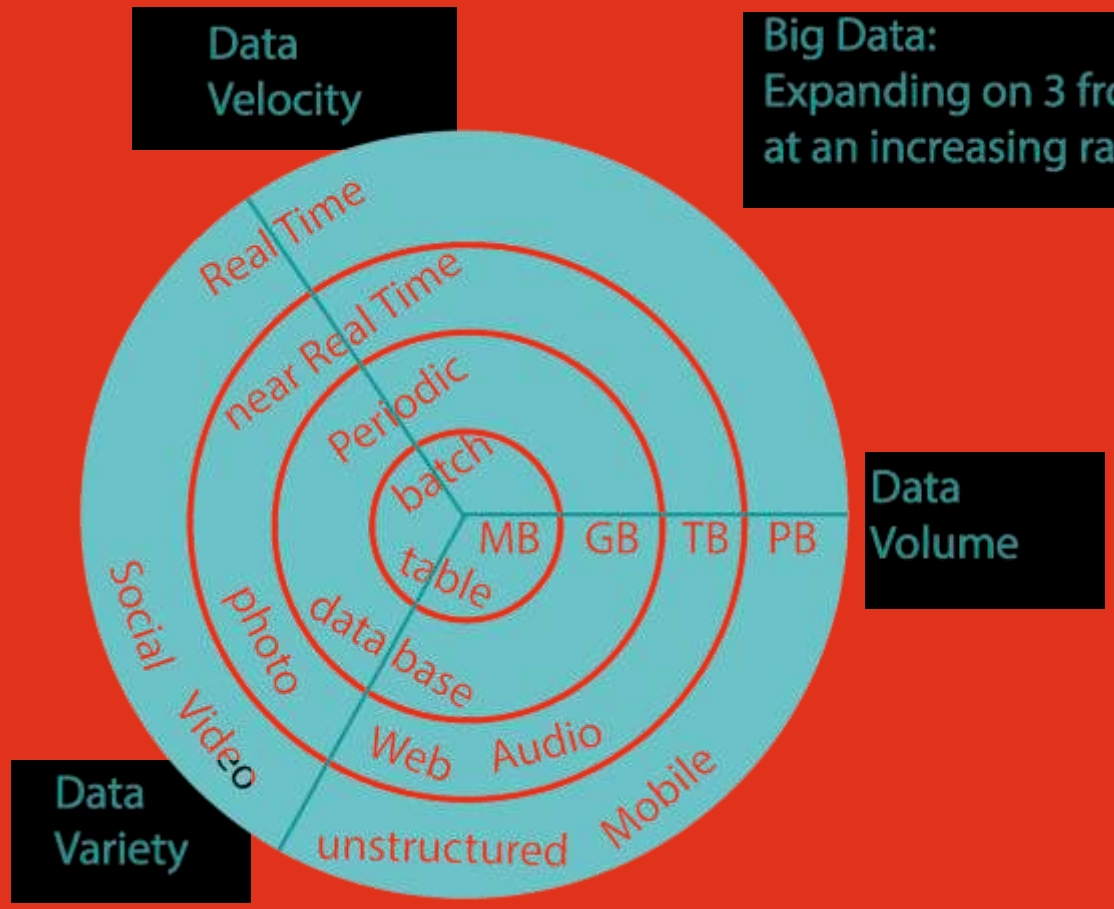




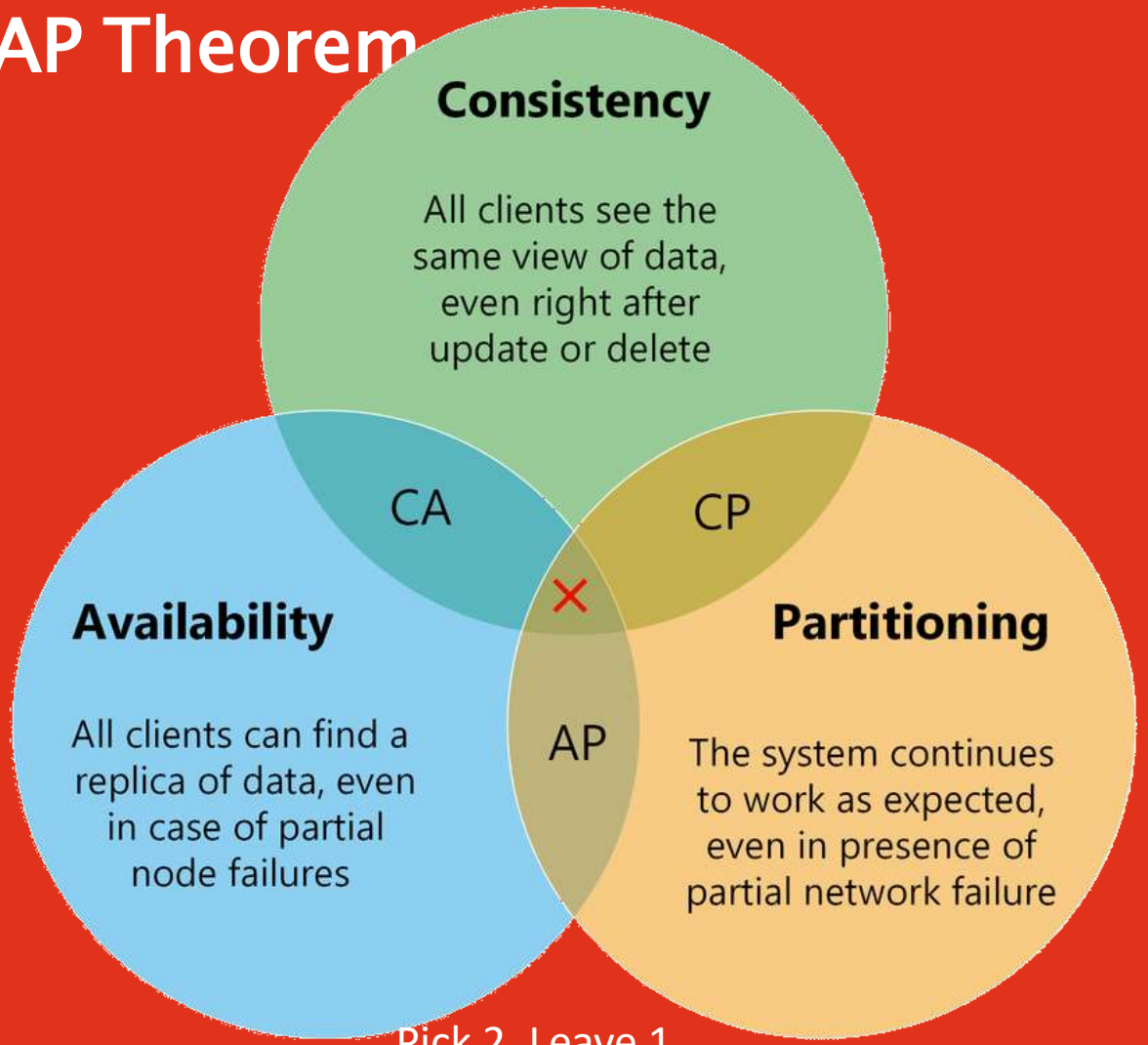


lysis

The 3 Vs of Big Data

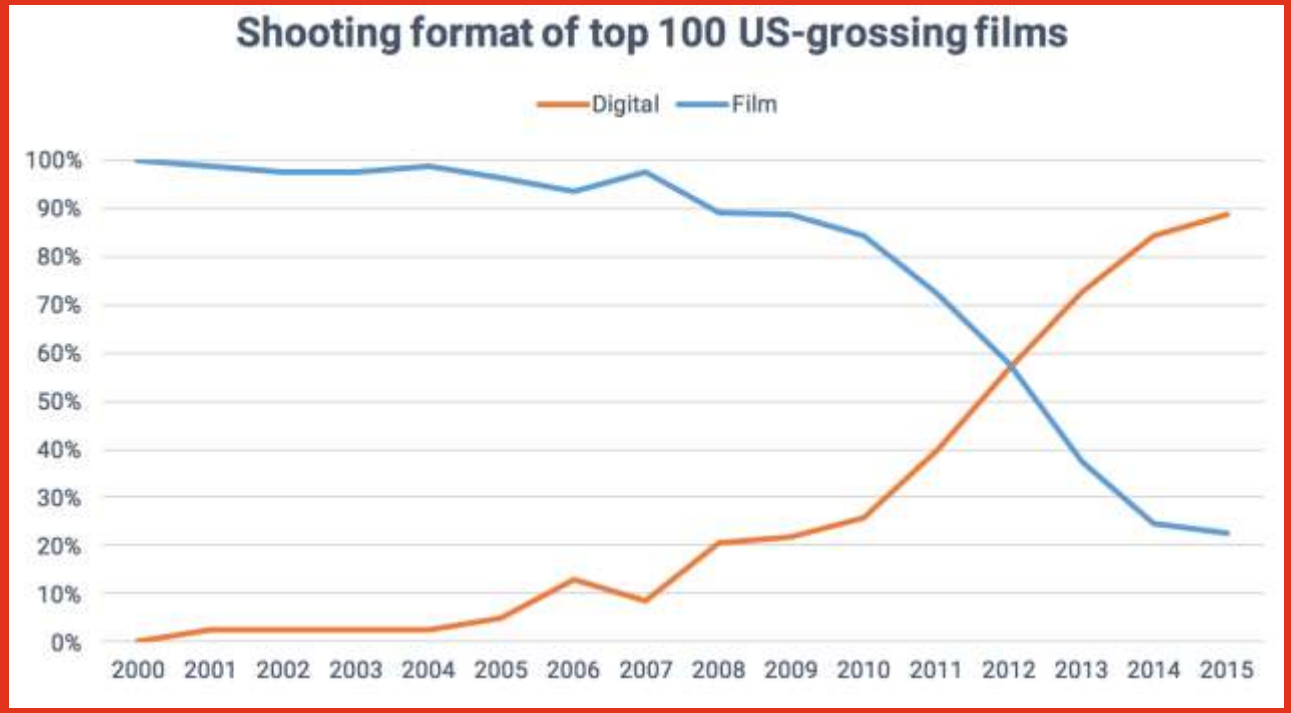


CAP Theorem



Pick 2, Leave 1
IMPOSSIBLE to pick 3

The beauty of data: story telling



CTRL + CLICK on the image for a bit of magic

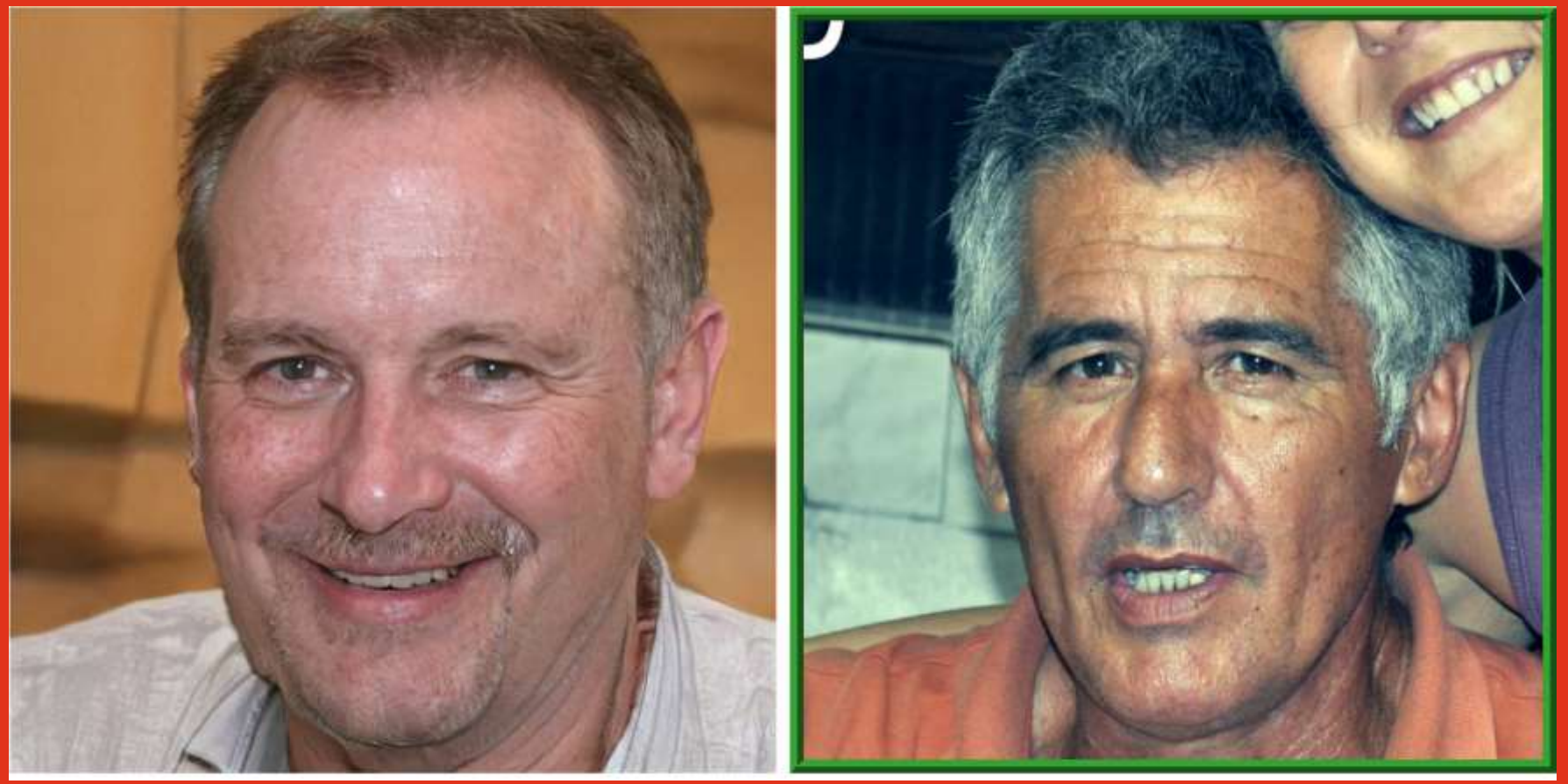
OK, so…how far will data take me?

You might think that the filming example is stunning but that traditional data processing can show less impressive results and also work as well.

However, the state-of-the-art in Data Science really needs all the skills that we have discussed. Take a look.

One picture is real and one picture was created from scratch with AI. The fake picture does not exist or is a mix of pictures. An algorithm plotted the face pixel by pixel after learning what a possible human face is.

The first face has never existed.



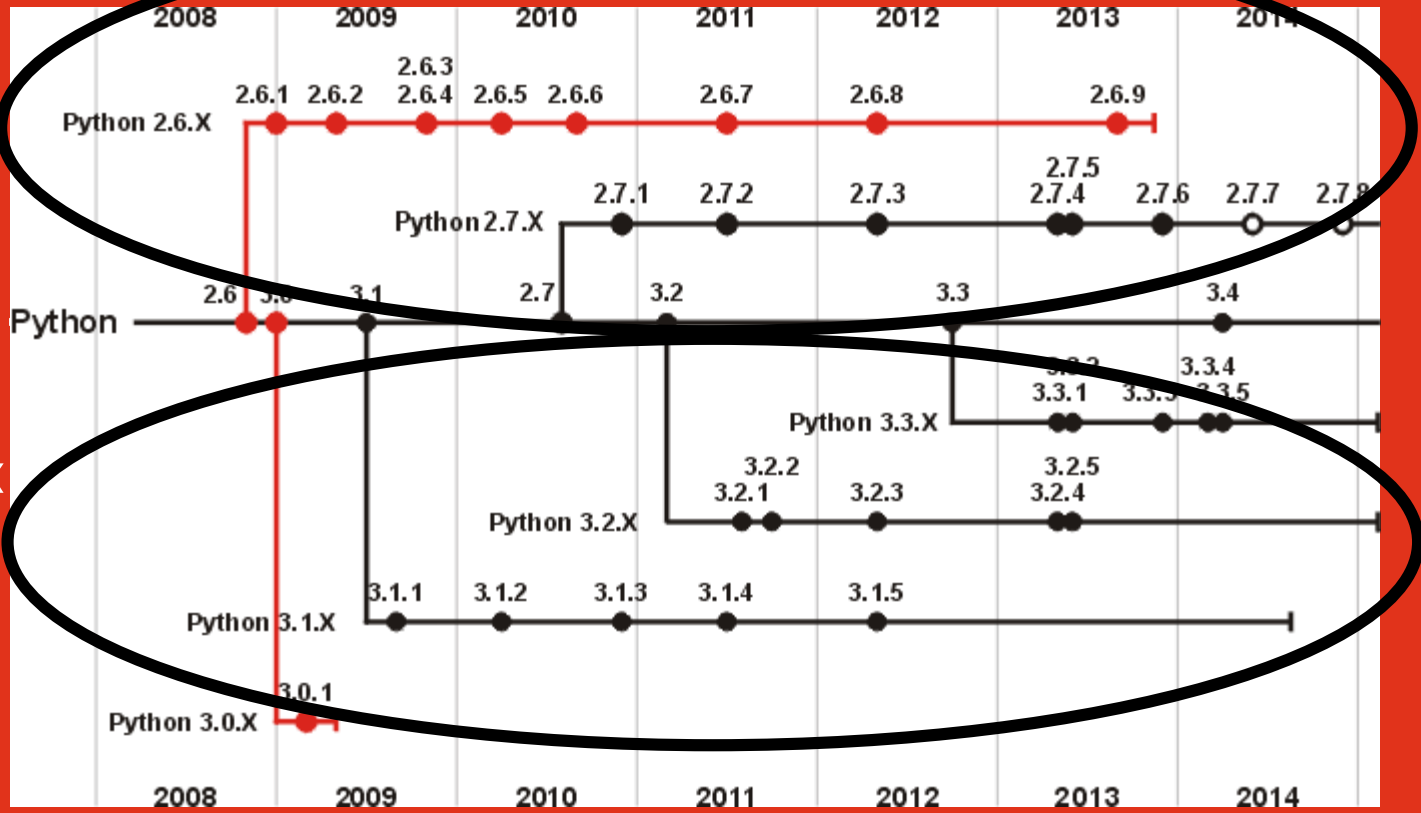
OK, I want to educate myself to become more digital.

You should speak Python!

Versions of Python

1. Python has 2 versions
2. Version 2.x and 3.x
3. They are different 'families'
4. **Version 2.x IS NOT under**
5. Families with parallel lives
6. **2.x does NOT work in 3.x**
7. We will learn 3.x
8. **Syntax and commands are different!**

Family 2.x



Family 3.x

Any question?

Thank you!