

Big Data Processing

L02: Big Data Motivation

Dr. Ignacio Castineiras
Department of Computer Science

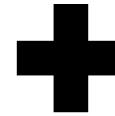
Outline

1. Artificial Intelligence Upside.
2. Artificial Intelligence Downside.

Outline

1. Artificial Intelligence Upside.
2. Artificial Intelligence Downside.

Artificial Intelligence: my view



Artificial Intelligence Upside

Example1:

Google HashCode'18: Self-driving rides

https://storage.googleapis.com/coding-competitions.appspot.com/HC/2018/hashcode2018_qualification_task.pdf



Artificial Intelligence Upside

Example1:

Google HashCode'18: Self-driving rides

https://storage.googleapis.com/coding-competitions.appspot.com/HC/2018/hashcode2018_qualification_task.pdf



This is a problem we tackled 2 years ago, as part of the CIT Programming Society

Artificial Intelligence Upside

Problem:

Given some passenger petitions given in advance,
can we schedule our taxi fleet rides to maximise the service?



Artificial Intelligence Upside

More in detail...

1. We are managing a company of taxis.
 - Typically, each taxi driver would like to maximise its own revenue.
 - However, let's assume we are in a model with self-driving cars. On it, we don't care about the revenue of a particular taxi; instead we try to maximise the revenue produced by the whole fleet of taxis.



Artificial Intelligence Upside

More in detail...

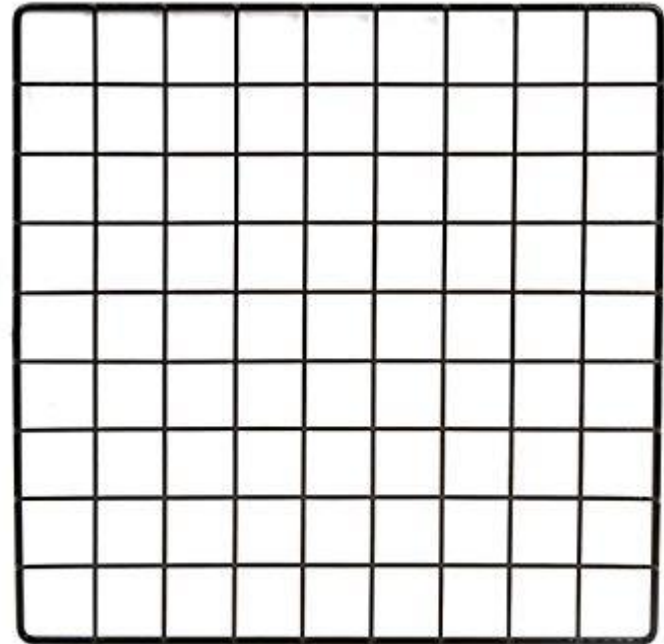
2. Let's imagine our taxis operate in a city.



Artificial Intelligence Upside

More in detail...

2. The streets of this city can be modelled as a grid.



Artificial Intelligence Upside

More in detail...

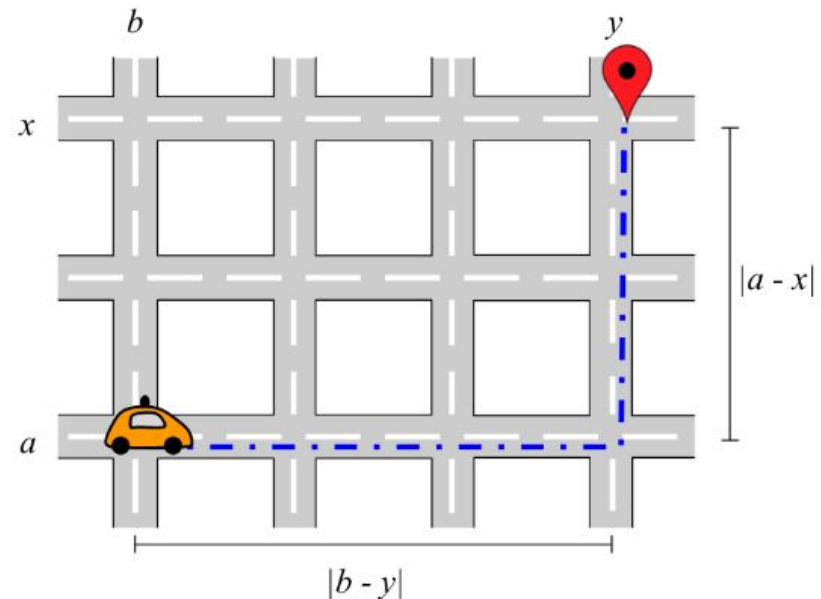
3. Let's imagine we have a set of customers, who have requested our taxi service in advance!



Artificial Intelligence Upside

More in detail...

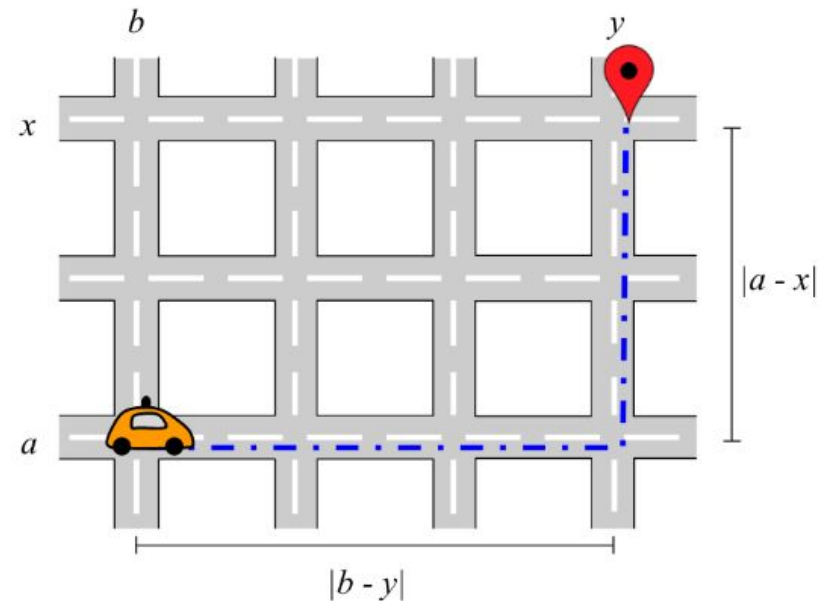
4. Each of these petitions can be represented via:
- The point to depart from.
 - The point to reach to.
 - The time of departure.



Artificial Intelligence Upside

More in detail...

- Likewise, the revenue for a trip can be modelled as the Euclidean distance between the start and end points, with a bonus for trips made on time.



Artificial Intelligence Upside

More in detail...

6. Let's revisit the problem statement now:

- You are the manager of MyTaxi.
- You control a fleet of taxis.
- It's early in the morning, and you know in advance all the customer petitions MyTaxi is going to have during the day.

The problem is:

- Create an algorithm to schedule the customer trips taken by the taxis, with the goal of maximising the revenue made by the entire fleet of taxis.

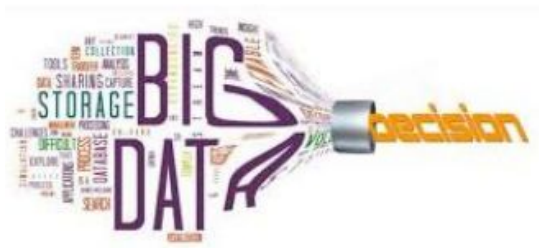


Artificial Intelligence Upside

Example1:

Clearly this is a Combinatorial Optimisation Problem,
and it thus belong to the third component of the
Artificial Intelligence view described before:

Big
Data



Machine
Learning



Combinatorial
Optimisation



Artificial Intelligence Upside

Example1:

But, wait a moment...

Did we just claim that, early in the morning,
we know **in advance**
all the taxi petitions we are going to receive during the day??!!



Artificial Intelligence Upside

Example1:

But, wait a moment...

Did we just claim that, early in the morning,
we know **in advance**
all the taxi petitions we are going to receive during the day??!!

Is this science fiction?

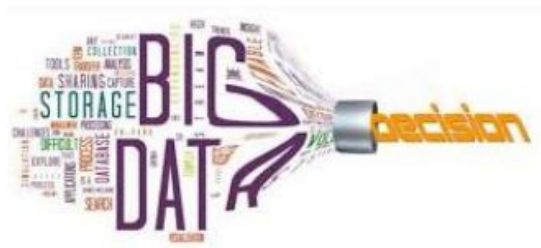


Artificial Intelligence Upside

Example1:

Let's start reasoning about our problem with the other two Artificial Intelligence components in mind:

Big
Data



Machine
Learning



Combinatorial
Optimisation



Artificial Intelligence Upside

Example1:

What would you do if you were working for MyTaxi, and you had to estimate in advance the amount of taxi petitions you are going to receive during the day?

Artificial Intelligence Upside

Example1:

What would you do if you were working for MyTaxi, and you had to estimate in advance the amount of taxi petitions you are going to receive during the day?



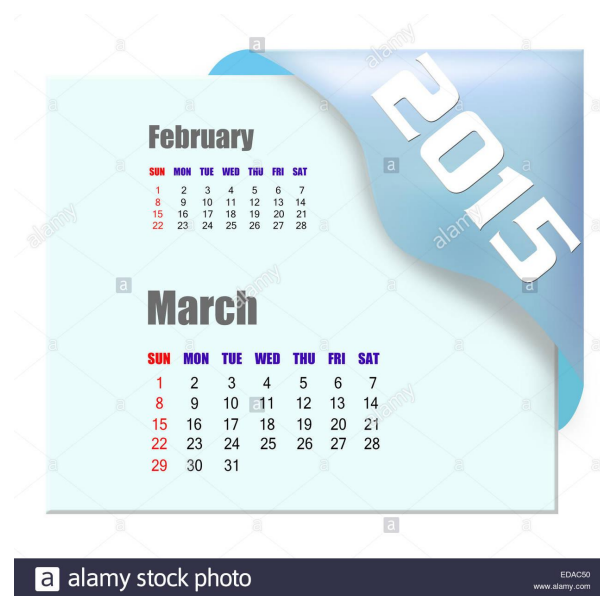
I tell you: struggling! That's what you will do :)

Artificial Intelligence Upside

Example1:

It is certainly not an easy task, but:

- Maybe you want to take a look at MyTaxi log files from past years.
- How many petitions did we receive in that year?

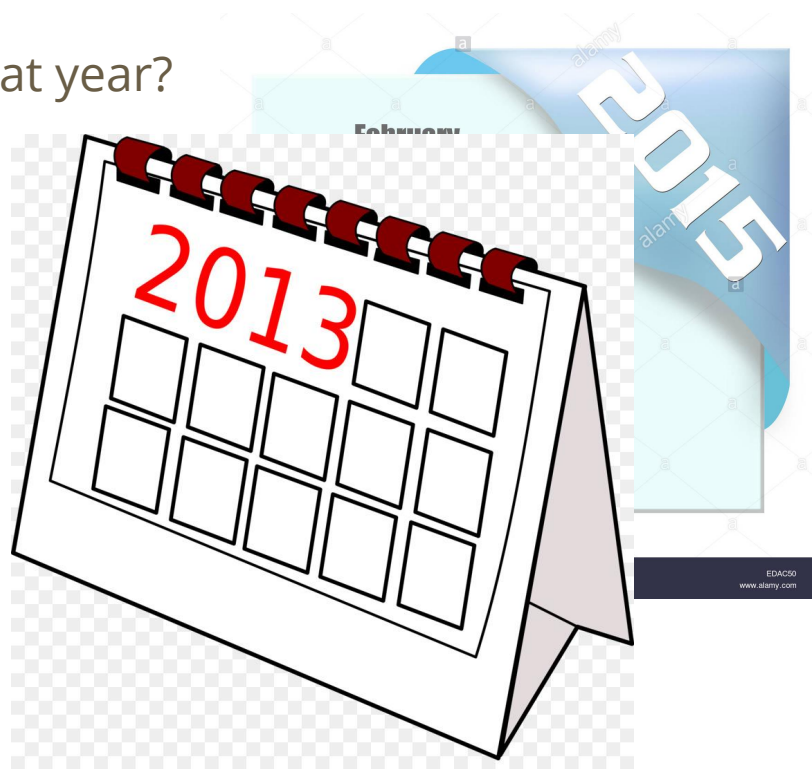


Artificial Intelligence Upside

Example1:

It is certainly not an easy task, but:

- Maybe you want to take a look at MyTaxi log files from past years.
- How many petitions did we receive in that year?
- And in that other year?



Artificial Intelligence Upside

Example1:

It is certainly not an easy task, but:

- Maybe you want to take a look at MyTaxi log files from past years.
- How many petitions did we receive in that year?
- And in that other year?
- The more the merrier, give me as many years as you can!



Artificial Intelligence Upside

Example1:

It is certainly not an easy task, but:

- But that's not enough.
What if the weather also has an impact in the amount of taxis being taken?
- What was the weather like in these last days?

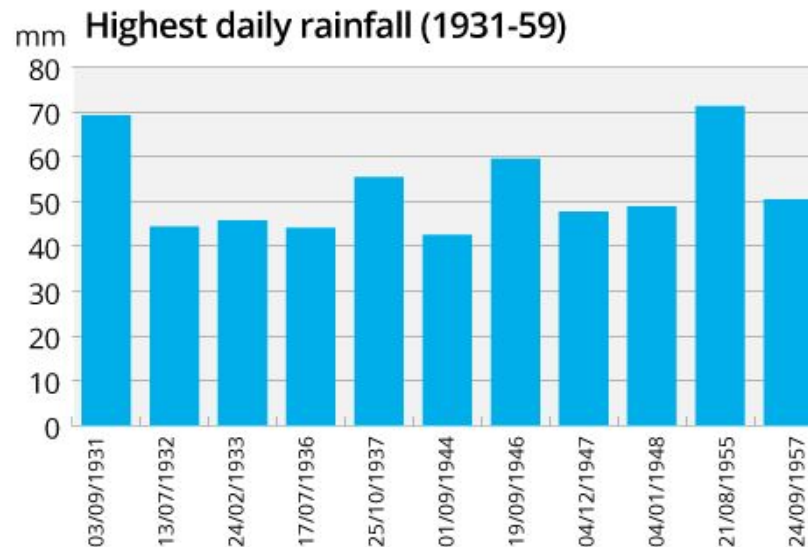


Artificial Intelligence Upside

Example1:

It is certainly not an easy task, but:

- But that's not enough.
What if the weather also has an impact in the amount of taxis being taken?
- What was the weather like in these last days?
- And, by these last days I mean the last thousands of days :)



Artificial Intelligence Upside

Example1:

It is certainly not an easy task, but:

- Is that enough? Maybe the income of the people matters too...
- How many people own a car these days?



Artificial Intelligence Upside

Example1:

It is certainly not an easy task, but:

- Is that enough? Maybe the income of the people matters too...
- How many people own a car these days?
- And how many did in the past?



Artificial Intelligence Upside

Example1:

It is certainly not an easy task, but:

- And is it that all?
- What about climate awareness?



ZERO 
EMISSION

Artificial Intelligence Upside

Example1:

It is certainly not an easy task, but:

- And is it that all?
- What about the last bus monthly pass offer?



Artificial Intelligence Upside

Example1:

It is certainly not an easy task, but:

- And is it that all?
- And what about...?
- And also...?
- Don't forget about...?

Artificial Intelligence Upside

Example1:

It is certainly not an easy task, but:

- And is it that all?
- And what about...?
- And also...?
- Don't forget about...?



I told you: struggling! That's what you will do :)

Artificial Intelligence Upside

Example1:

In other words,
you want to put your hands into tons of data



Artificial Intelligence Upside

Example1:

In other words,
you want to put your hands into tons of data
and make sense of it!

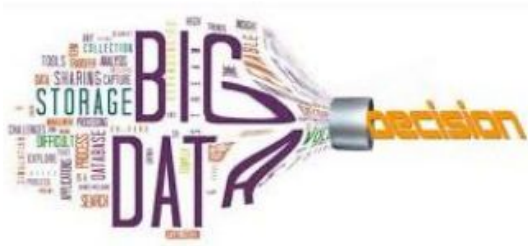


Artificial Intelligence Upside

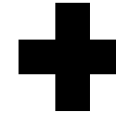
Example1:

This is the role of the Big Data component!
You analyse the data, looking for patterns, correlations, aggregations, and ultimate conclusions.

Big
Data



Machine
Learning



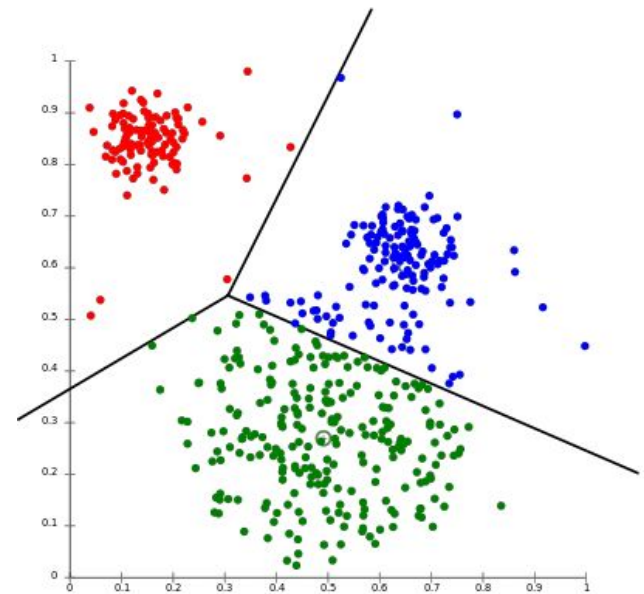
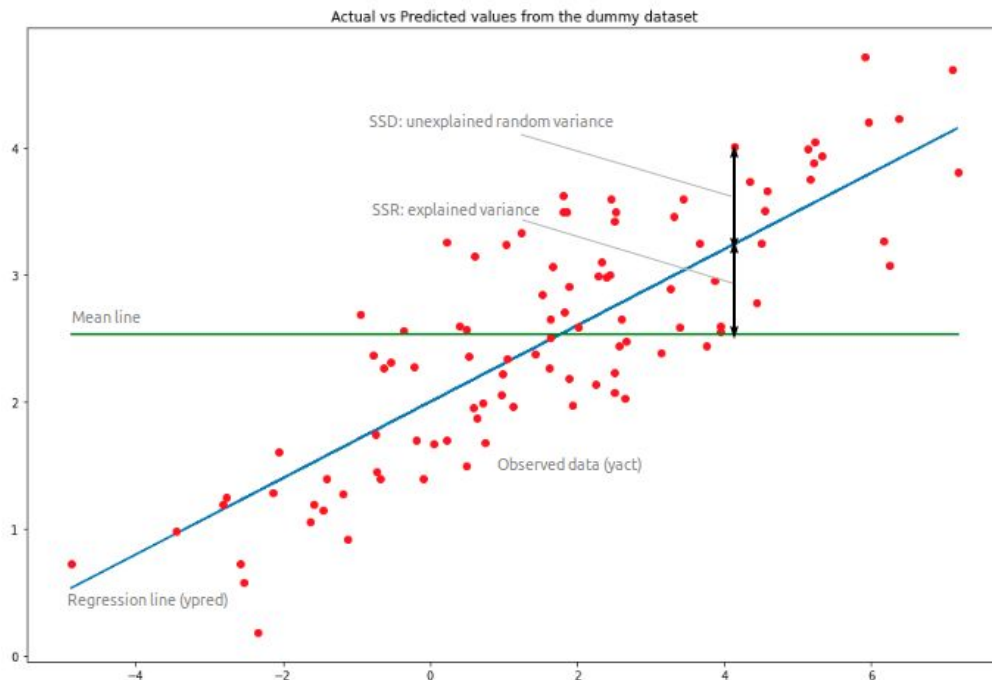
Combinatorial
Optimisation



Artificial Intelligence Upside

Example1:

Then you pass on these patterns, correlations, aggregations, and conclusions so as to build prediction models.

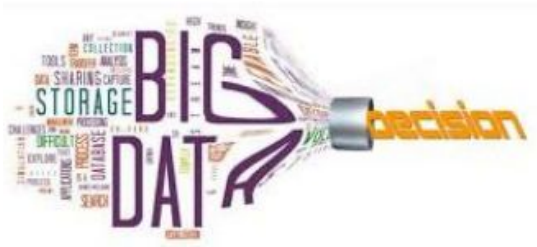


Artificial Intelligence Upside

Example1:

This is done by the Machine Learning component!

Big
Data



Machine
Learning



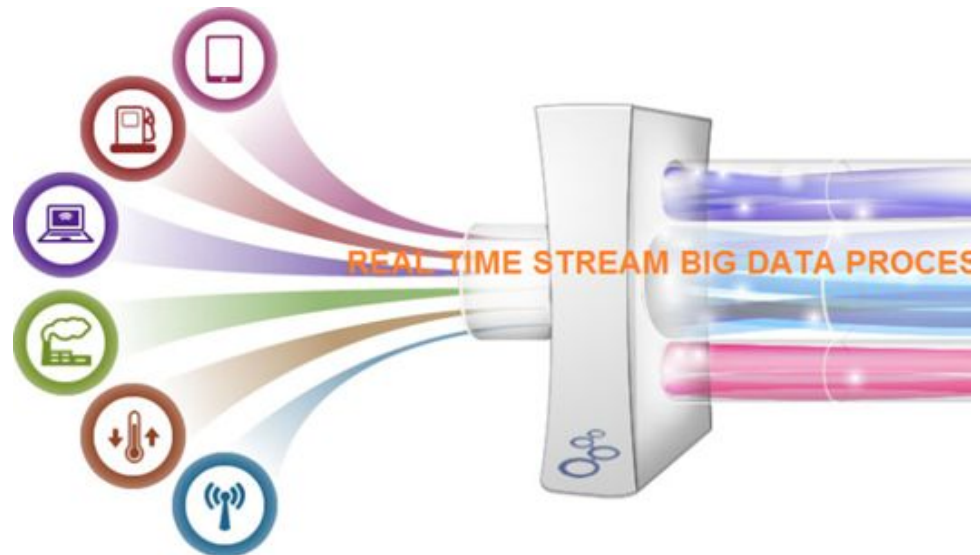
Combinatorial
Optimisation



Artificial Intelligence Upside

Example1:

Now, you collect novel data,
perhaps obtained every new hour, or even every new minute,
or even every new second.



Artificial Intelligence Upside

Example1:

This is the role of the Big Data Streaming Analysis component again!

This component keeps up to date with the data being received and processes it, whatever the ingestion pace is.

Big
Data



Machine
Learning



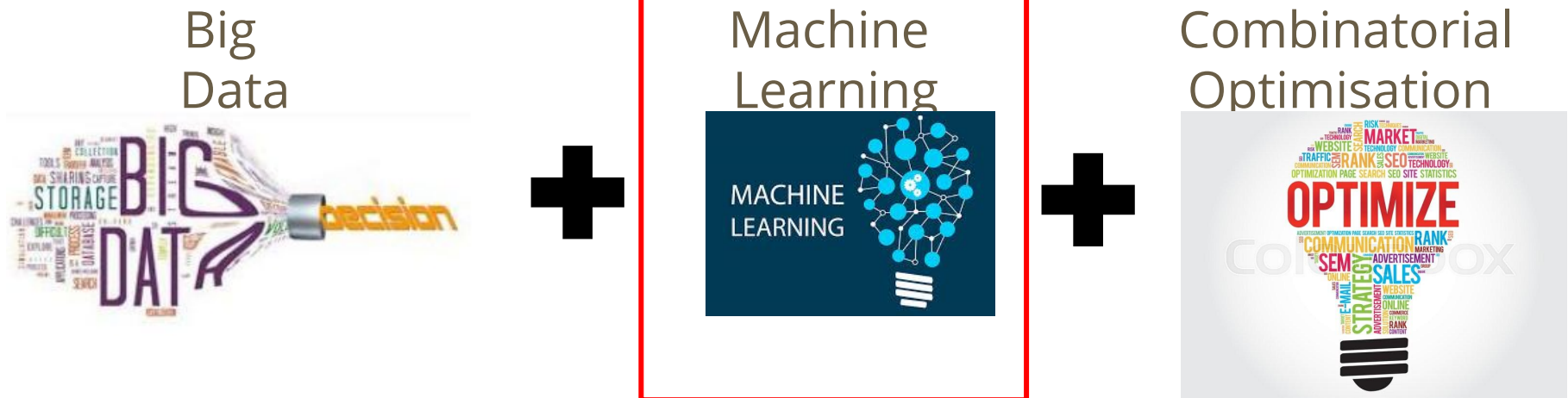
Combinatorial
Optimisation



Artificial Intelligence Upside

Example1:

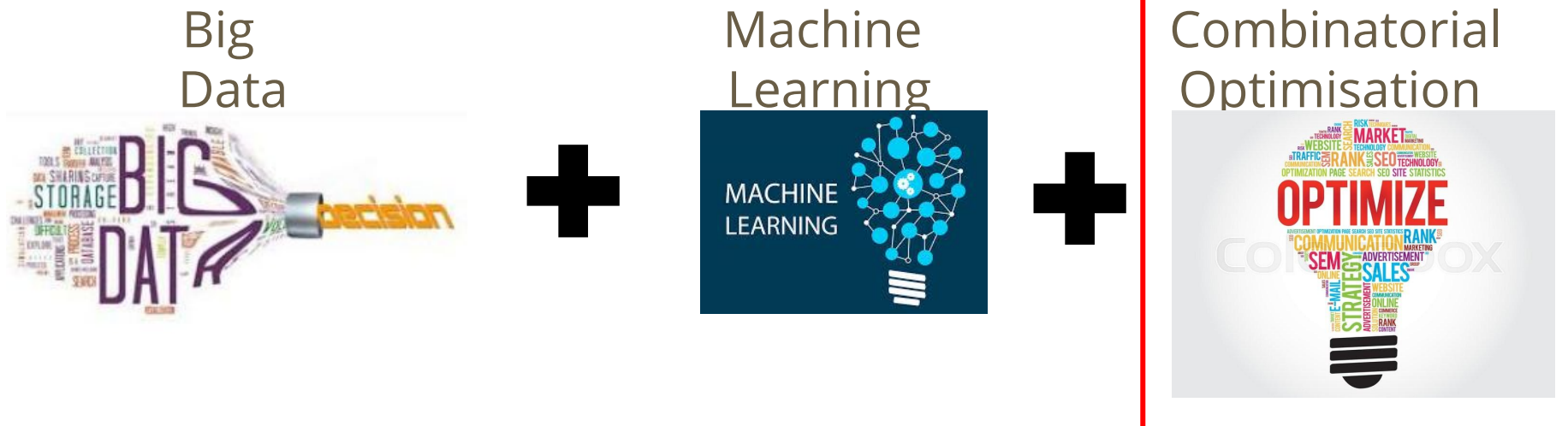
And this novel data, once analysed, can be fed into our model predictor previously created, so as to provide us with an updated prediction for the novel data just received.



Artificial Intelligence Upside

Example1:

And now yes, finally with this prediction for the novel data we can use combinatorial optimisation so as to enable better decisions.

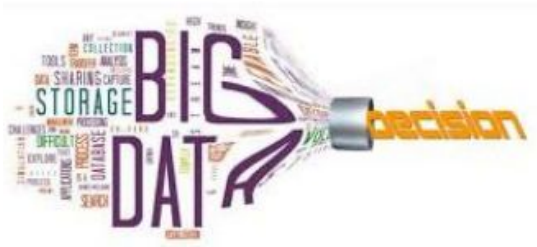


Artificial Intelligence Upside

Example1:

In our case, once we can predict the taxi customer petitions for the new day, we can come back to our Google Hash Code'18 constraint optimisation problem and solve it maximising the revenue generated by the taxi fleet.

Big
Data



Machine
Learning



Combinatorial
Optimisation



Artificial Intelligence Upside

Example1:

And this is just one example of how Artificial Intelligence can make our society better by providing us with a better taxi service, that ultimate reduces the number of private cars and improves the CO2 levels.

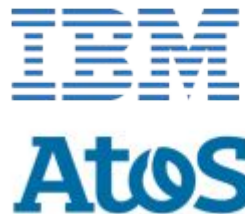


Artificial Intelligence Upside

Example2:

Years ago I worked for the European Union funded research project GENiC:
Globally Optimised Energy Efficient Data Centres

<http://projectgenic.eu/>

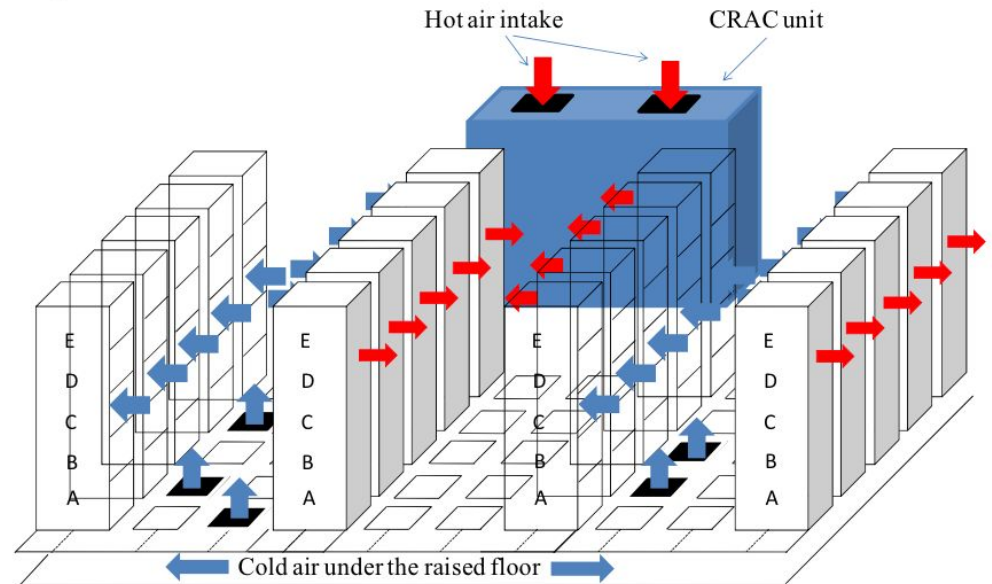


Artificial Intelligence Upside

Example2:

The project fell under the research areas of:

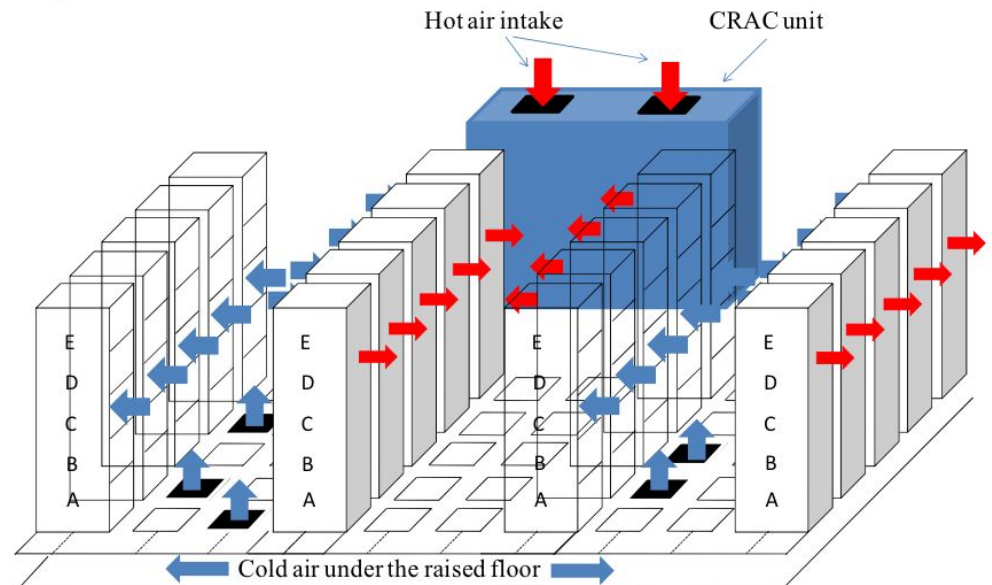
- Green computing.
- Sustainable Data Centres.
- Renewable energy sources.



Artificial Intelligence Upside

Example2:

- On it, we had a set of servers (which play the role of the fleet of taxis).



Artificial Intelligence Upside

Example2:

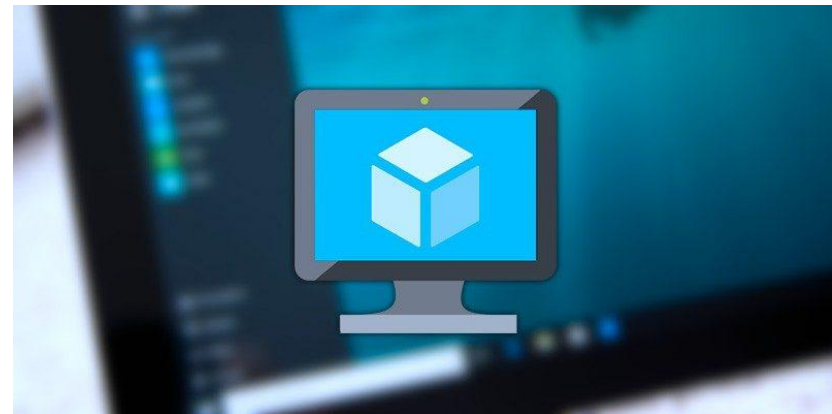
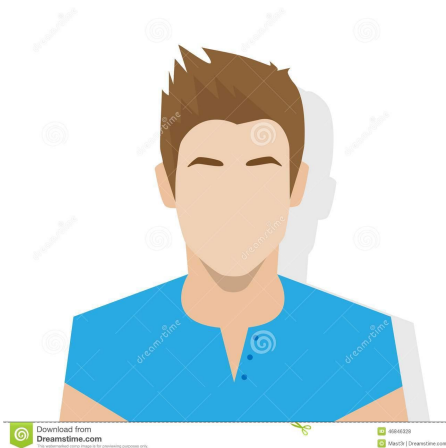
- We had a set of customer petitions (e.g., using Gmail).



Artificial Intelligence Upside

Example2:

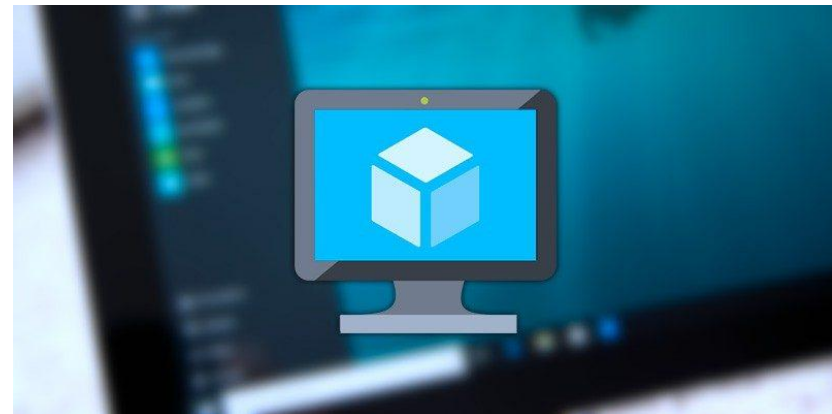
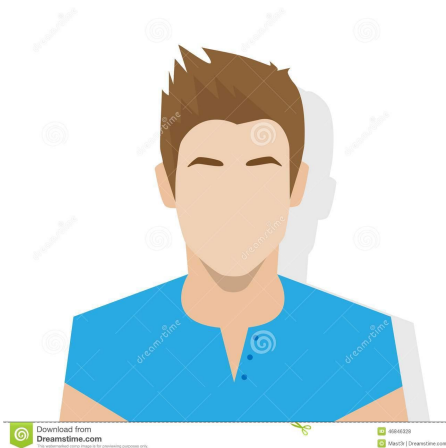
- Each customer petition was identified by the time and the amount of resources needed. This petition was processed by a Virtual Machine (VM) consolidated in a server.



Artificial Intelligence Upside

Example2:

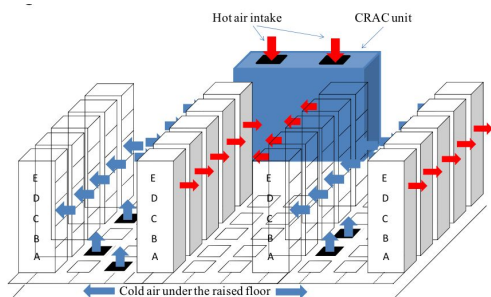
- Each customer petition was identified by the time and the amount of resources needed. This petition was processed by a Virtual Machine (VM) consolidated in a server.
- This played the role of our customer asking for taxis. A VM can satisfy multiple customer petitions (as long as it has the resources for it).



Artificial Intelligence Upside

Example2:

- The goal was to minimise the power consumption for running the customer petitions by wisely consolidating them in concrete VMs of the Data Centre.

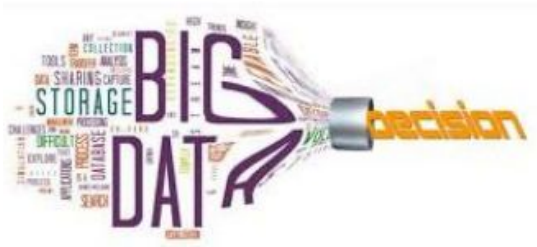


Artificial Intelligence Upside

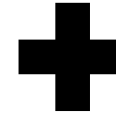
Example2:

Clearly this is a Combinatorial Optimisation Problem, and it thus belong to the third component of the Artificial Intelligence view described before:

Big
Data



Machine
Learning



Combinatorial
Optimisation



Artificial Intelligence Upside

Example2:

However, at this stage I hope you have found the similarity between this example and the previous one.

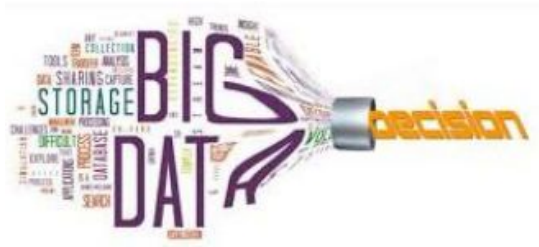
How on earth did we get to know the customer petitions **in advance**??!!

Artificial Intelligence Upside

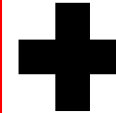
Example2:

In this case we will avoid reasoning about the other two components of Artificial Intelligence, but you can just guess their role from the previous example.

Big
Data



Machine
Learning



Combinatorial
Optimisation



Artificial Intelligence Upside

When you look at these examples and the way Artificial Intelligence is shaping our society you might wonder...

When did all of this start?

Artificial Intelligence Upside

This is a very difficult question to answer
as it has been a race with many hints.

But let me highlight the following one (see video from 1:20 to 3:00):

<https://www.youtube.com/watch?v=9hUlxyE2Ns8&t=187s>



Outline

1. Artificial Intelligence Upside.
2. Artificial Intelligence Downside.

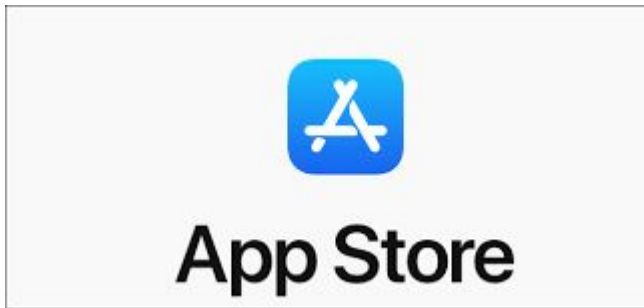
Artificial Intelligence Downside

- I remark the arrival of the SmartPhone as it comes with plenty of sensors:
 - Wifi Connection.
 - Accelerometer.
 - GPS.



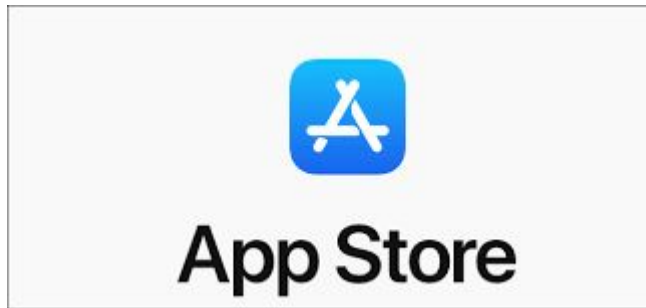
Artificial Intelligence Downside

- Ever since people started downloading mobile apps.



Artificial Intelligence Downside

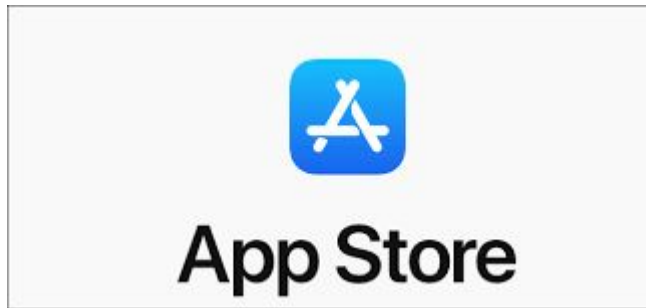
- Ever since people started downloading mobile apps.



- Most of these mobile apps are *free*
(in the sense that one does not need to pay for downloading them).

Artificial Intelligence Downside

- Ever since people started downloading mobile apps.



- Most of these mobile apps are *free*
(in the sense that one does not need to pay for downloading them).
- However, nothing is really for free, and we just *pay* the price of these apps with our data, with the information we generate.

Artificial Intelligence Downside

- Just think for a second of 1 permission most apps require from us:

ACCESS_FINE_LOCATION

```
1 <manifest
2   xmlns:android="http://schemas.android.com/apk/res/android"
3   package="____"
4   android:versionCode="1"
5   android:versionName="1.0">
6
7   <uses-permission android:name="android.permission.INTERNET" />
8   <uses-permission android:name="android.permission.SYSTEM_ALERT_WINDOW" />
9
10  <uses-permission android:name="android.permission.ACCESS_FINE_LOCATION" />
11  <uses-permission android:name="android.permission.ACCESS_COARSE_LOCATION" />
12  <application
13    android:name=".MainApplication"
14    android:label="@string/app_name"
```



Artificial Intelligence Downside

Basically, this means answering to the question:

Where are you?

Artificial Intelligence Downside

- Of course accessing our location is crucial for the service offered by some apps:



Artificial Intelligence Downside

- But, there are many other apps requiring `ACCESS_FINE_LOCATION` too. And, for most of them, it is unclear which service(s) they won't be able to offer us in case the permission is declined.



Artificial Intelligence Downside

- Why is it so much important to know your location?

Artificial Intelligence Downside

- Why is it so much important to know your location?

Because knowing where you are 24/7
is a way of inferring the answer
to a much more powerful question:

Artificial Intelligence Downside

- Why is it so much important to know your location?

Because knowing where you are 24/7
is a way of inferring the answer
to a much more powerful question:

Who are you?

Artificial Intelligence Downside

Let's think together of the information
that can be inferred from your location.

This resonates much more (as it makes the metaphor explicit)
if you are watching this class in your mobile phone:



Artificial Intelligence Downside

Let's think together of the information
that can be inferred from your location.

The examples presented here are from a talk of Chema Alonso, Chief Data Officer (CDO) of Telefonica, one of the strongest telecommunication companies in the world.



Artificial Intelligence Downside

Let's think together of the information
that can be inferred from your location.

1. Information derived from your location:

Where do you live?

Where do you go on holidays?

Where do you work?

Where do you do your shopping?



Artificial Intelligence Downside

Let's think together of the information
that can be inferred from your location.

2. Information derived from how your location changes over time:

Do you walk?

Do you go by car?

By bike perhaps?

Maybe a combination of them?



Artificial Intelligence Downside

Let's think together of the information
that can be inferred from your location.

3. **Let's now cross your location with a database of points of interest:**
These are databases with millions of points, updated automatically and fixed manually, everyday.

Do you remember the movie *Dude, where's my car?*
So, with Siri, this problem can no longer happen.



Artificial Intelligence Downside

Let's think together of the information
that can be inferred from your location.

3. **Let's now cross your location with a database of points of interest:**

The app detects that you are going at car speed.

Then, you stop in a location which stands for a restaurant in its point of interests database.

You spend about an hour there.

Finally, when you go out and start walking again...

...the app informs you, automatically, where did you park your car.



Artificial Intelligence Downside

Let's think together of the information
that can be inferred from your location.

4. **Crossing your location with the time of the day it is:**
Where do you sleep?
Do you sleep every night in the same place?



Artificial Intelligence Downside

Let's think together of the information
that can be inferred from your location.

4. Crossing your location with the time of the day it is:

Do you go to the gym?

With who?

And how often?

And for how long?



Artificial Intelligence Downside

Let's think together of the information
that can be inferred from your location.

4. Crossing your location with the time of the day it is:

Do you play sports?

Do you go to the pub? *(same questions as for playing sports apply here)*

And what about restaurants? Which ones do you like?



Artificial Intelligence Downside

Let's think together of the information
that can be inferred from your location.

5. And the list just goes on...

Do you go to the doctor?

Do you go to the hospital?

Do you visit somebody in jail?



Artificial Intelligence Downside

Let's think together of the information
that can be inferred from your location.

5. And the list just goes on...

Coming back to where you work and the exact location of your office:

- What is your role in your company?
- What is your salary?



Artificial Intelligence Downside

Let's think together of the information
that can be inferred from your location.

5. And the list just goes on...

Combining where do you live with a state property management database:

- Do you own your place or are you renting?



Artificial Intelligence Downside

Let's think together of the information
that can be inferred from your location.

5. And the list just goes on...

What is your close circle?

Who do you stay with?



Artificial Intelligence Downside

Let's think together of the information
that can be inferred from your location.

5. And the list just goes on...

Do you go to political meetings?

Or to demonstrations?

Do you have a faith? Do you go to church?



Artificial Intelligence Downside

All in all...

Where are you?  **Who are you?**

Artificial Intelligence Downside

But this is not all...

all the above examples were based on the exterior,
in actual actions performed by individuals...



Artificial Intelligence Downside

...but what if we change slightly the perspective...

...and start looking at the interior,
at the thoughts, sensations and emotions of people
(that is, at their feelings, at their personality)



Artificial Intelligence Downside

One of the websites I am most fascinated by is:

<https://www.16personalities.com/>

Analysts



Architect

INTJ-A / INTJ-T



Logician

INTP-A / INTP-T



Commander

ENTJ-A / ENTJ-T



Debater

ENTP-A / ENTP-T

Artificial Intelligence Downside

One of the websites I am most fascinated by is:

<https://www.16personalities.com/>

Diplomats



Advocate

INFJ-A / INFJ-T



Mediator

INFP-A / INFP-T



Protagonist

ENFJ-A / ENFJ-T



Campaigner

ENFP-A / ENFP-T

Artificial Intelligence Downside

One of the websites I am most fascinated by is:

<https://www.16personalities.com/>

Sentinels



Logistician

ISTJ-A / ISTJ-T



Defender

ISFJ-A / ISFJ-T



Executive

ESTJ-A / ESTJ-T



Consul

ESFJ-A / ESFJ-T

Artificial Intelligence Downside

One of the websites I am most fascinated by is:

<https://www.16personalities.com/>

Explorers



Virtuoso

ISTP-A / ISTP-T



Adventurer

ISFP-A / ISFP-T



Entrepreneur

ESTP-A / ESTP-T



Entertainer

ESFP-A / ESFP-T

Artificial Intelligence Downside

One of the websites I am most fascinated by is:

<https://www.16personalities.com/>

- Besides how accurate I found the stuff of this website, the two things blowing my mind the most are:

Artificial Intelligence Downside

One of the websites I am most fascinated by is:

<https://www.16personalities.com/>

- Besides how accurate I found the stuff of this website, the two things blowing my mind the most are:
 1. According to their model, people are not that much different. Indeed, just 4 traits and 16 personalities serve to represent the whole mass.

Artificial Intelligence Downside

One of the websites I am most fascinated by is:

<https://www.16personalities.com/>

- Besides how accurate I found the stuff of this website, the two things blowing my mind the most are:
 1. According to their model, people are not that much different. Indeed, just 4 traits and 16 personalities serve to represent the whole mass.
 2. According to their model, it does not take much to match an individual to a personality, just quick 90 questions (that can be filled out in about 10 minutes) do the job.

Artificial Intelligence Downside

And, needless to be said, knowing who everybody is, understanding each and everyone personalities, brings an incredible amount of opportunities...

Artificial Intelligence Downside

And, needless to be said, knowing who everybody is, understanding each and everyone personalities, brings an incredible amount of opportunities...

...together with an incredible amount of risks too!

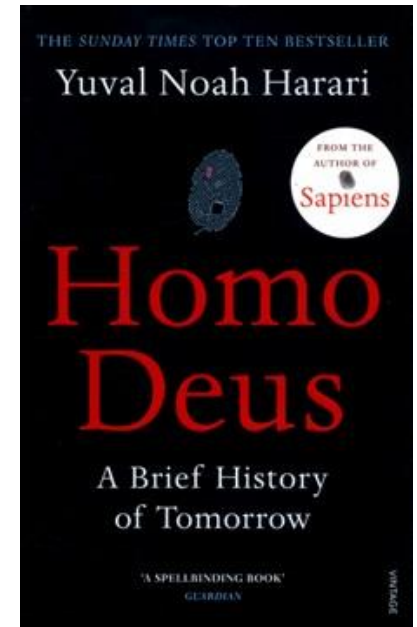
Artificial Intelligence Downside

- Recently, as a result of the scandal of Cambridge Analytica, the media has turned its attention to these risks:



Artificial Intelligence Downside

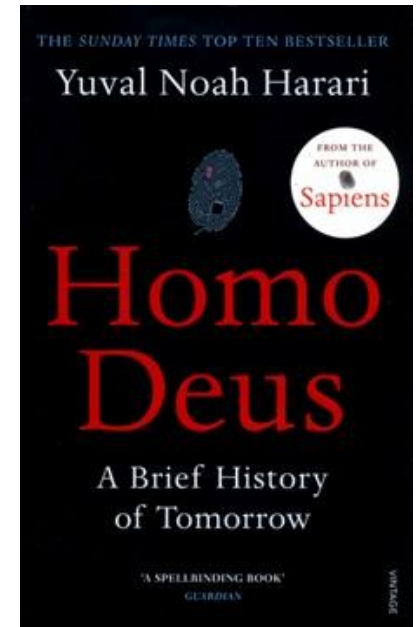
- But, in my humble opinion, the best way to reflect about the future that Artificial Intelligence and Biotechnology will bring us to is by reading the book Homo Deus, from professor Yuval Noah Harari.



Artificial Intelligence Downside

- But, in my humble opinion, the best way to reflect about the future that Artificial Intelligence and Biotechnology will bring us to is by reading the book Homo Deus, from professor Yuval Noah Harari.

Let's finish this section with the last paragraph of his book,
which I found so inspiring:



Artificial Intelligence Downside

Professor Yuval Noah Harari, Homo Deus.

- *If we think in term of months, we had probably focus on immediate problems such as the turmoil in the Middle East, the refugee crisis in Europe and the slowing of the Chinese economy.*
- *If we think in terms of decades, then global warming, growing inequality and the disruption of the job market loom large.*
- *Yet if we take the really grand view of life, all other problems and developments are overshadowed by three interlinked processes:*
 1. *Science is converging on an all-encompassing dogma, which says that organisms are algorithms, and life is data processing.*
 2. *Intelligence is decoupling from consciousness.*
 3. *Non-conscious but highly intelligent algorithms may soon know us better than we know ourselves.*

Artificial Intelligence Downside

Professor Yuval Noah Harari, Homo Deus.

- *These three processes raise three key questions, which I hope will stick in your mind long after you have finished this book:*
- 1. Are organisms really just algorithms, and is life really just data processing?*
 - 2. What's more valuable – intelligence or consciousness?*
 - 3. What will happen to society, politics and daily life when non-conscious but highly intelligent algorithms know us better than we know ourselves?*

Outline

1. Artificial Intelligence Upside.
2. Artificial Intelligence Downside.

Thank you for your attention!