

# The Useless Data Point

## A Struggling Path to Data Literacy

29<sup>th</sup> of October 2018



Francesco Costigliola



# Fidelidade Group includes both insurance companies and insurance related services companies

## Insurance companies

Health



Non-Life



Life

Portugal,  
Angola (Universal),  
Mozambique,  
Cape Verde (Garantia),  
Spain, France and  
Macao

Traditional

Direct/Online



## Service companies



Healthcare



Assistance



Health & Safety



Real Estate



Loss Adjusters



Auto Repair



# Fidelidade has been consolidating the leadership in Portuguese insurance market



**3,8Bn€**

Premiums Written,  
amongst largest  
players in Iberia



**15,5Bn€**

Assets under  
Management, amongst  
largest investors in PT



**230Mn€**

Net Income



**31%**

Overall market  
share in Portugal



**3.600+**

Workers in Fidelidade  
Group Insurance  
Companies



**2,2Mn**

Customers including  
individuals, SME and  
Corporates

A dark server room with rows of server racks. The racks have glowing blue lights at the bottom and small white lights along the top edge. The ceiling has several rectangular light fixtures.

Data is...



**Data... data is raw!**

**It simply exists and has no significance beyond its existence**

89

Is this clear?

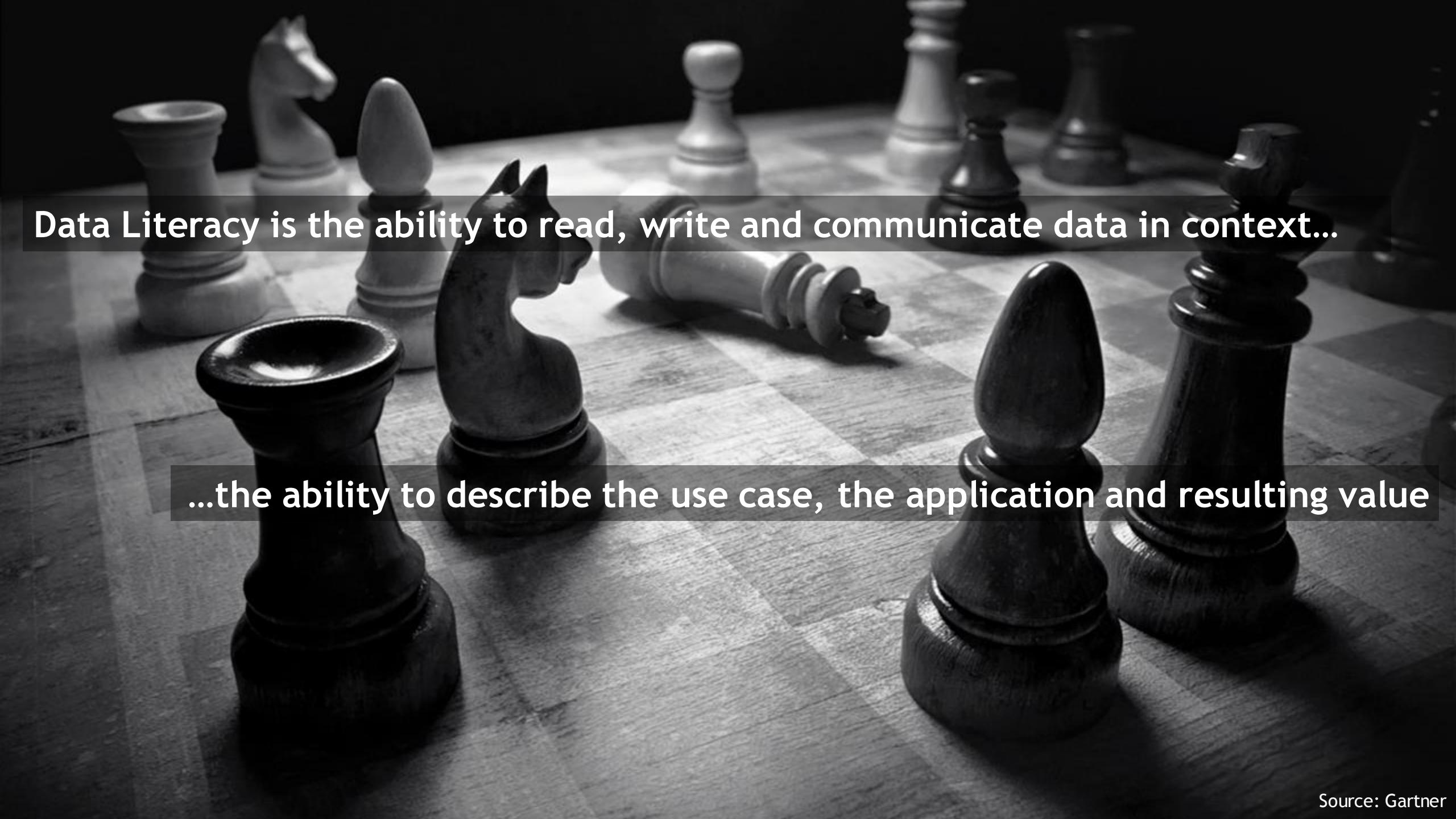
... and now?





70

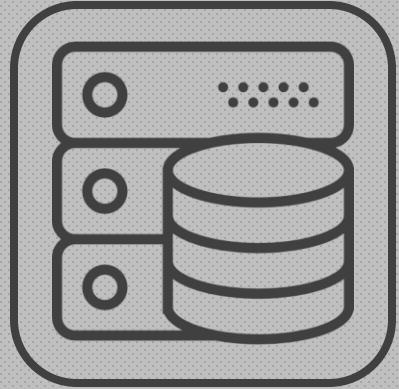
Context about data is key!



**Data Literacy is the ability to read, write and communicate data in context...**

**...the ability to describe the use case, the application and resulting value**

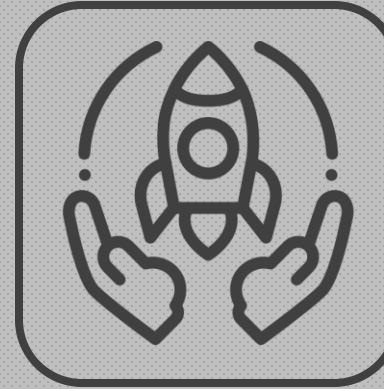
# Road to Data Literacy



Analytics  
Information  
Management



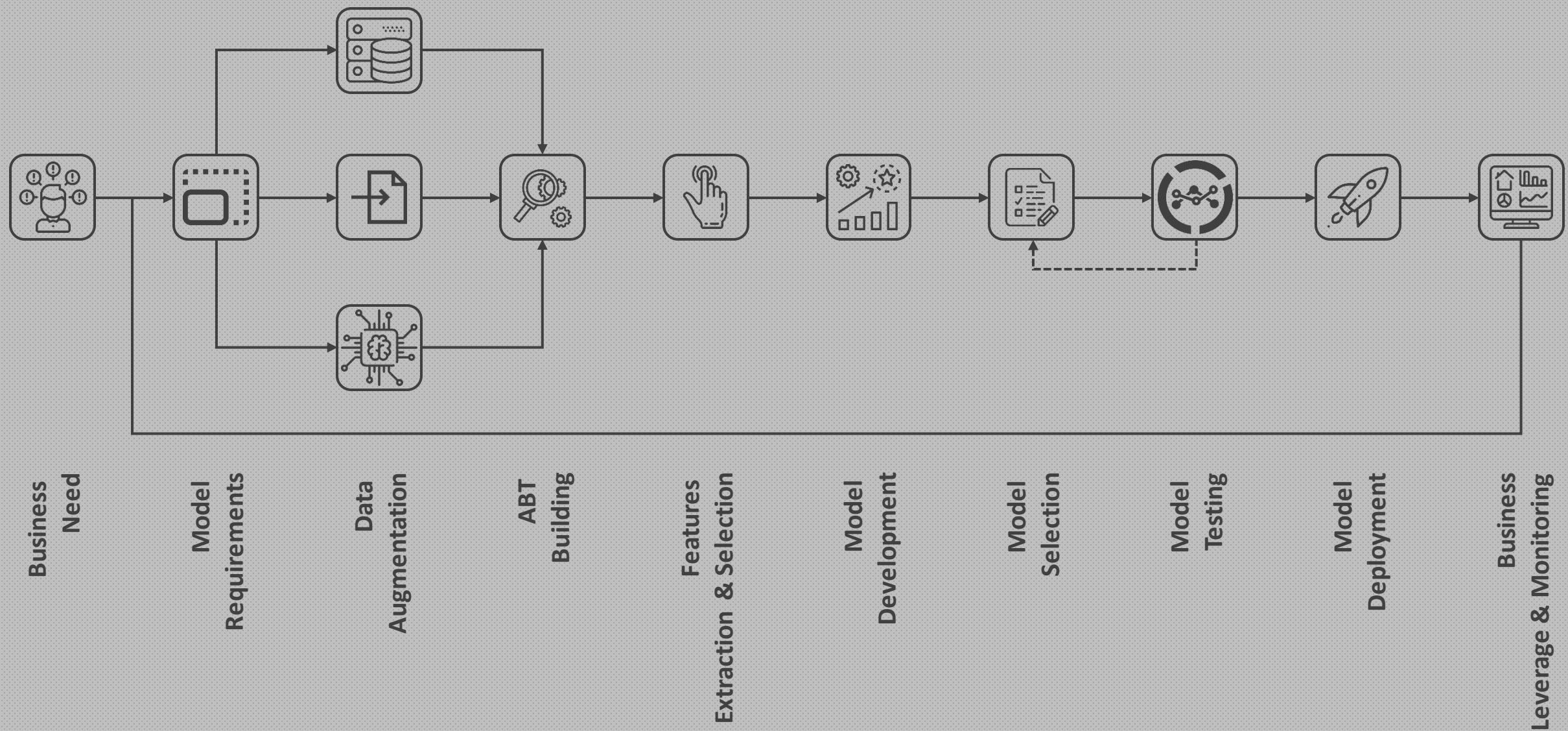
Analytics  
&  
Data Science



Integration  
& Business  
Leverage



# From Business Needs to Measurable Results



# Next Best Offer for Inbound



# Intelligent Process Automation



A close-up photograph of a child's hands playing with sand in a wooden sandbox. In the foreground, a pink plastic bucket lies on its side, partially buried in the sand. Behind it, a green bucket with a red letter 'T' on it contains some sand. The child's legs and feet are visible on the left, wearing grey pants and white sneakers. The background shows more of the sandy area and a wooden fence.

# Embrace Controlled Failure

# Our Bulletproof Learning Method!



We build it better!

# We decided to build our own OCR!

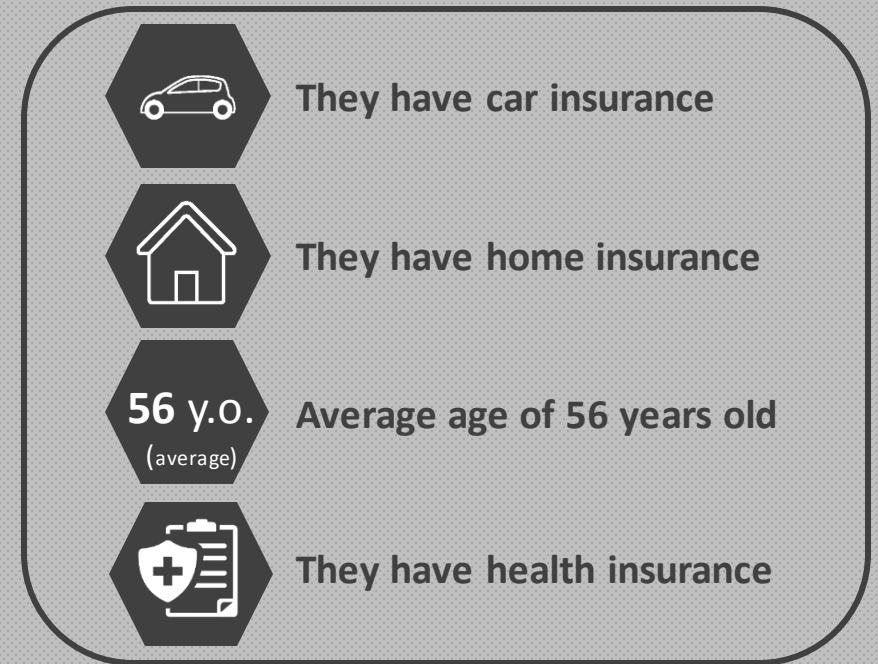
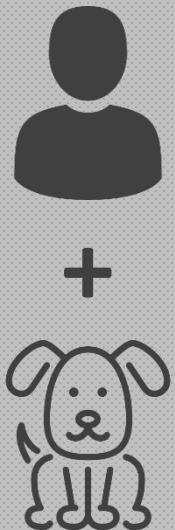
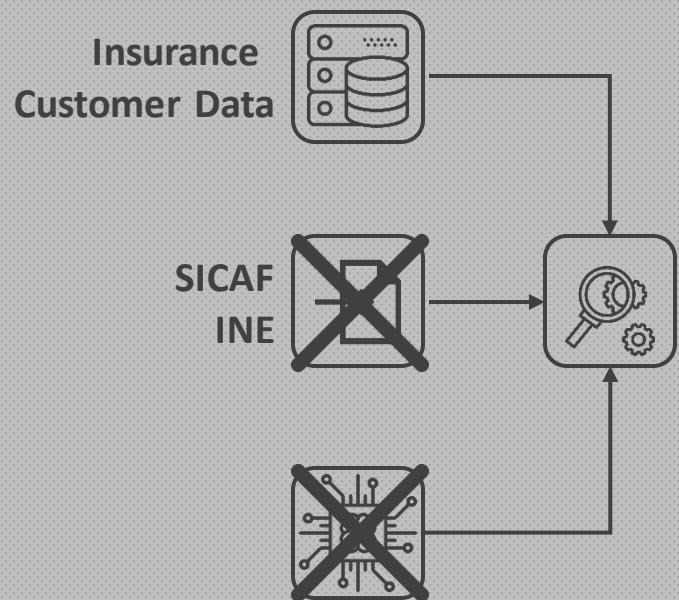
Don't reinvent the wheel!



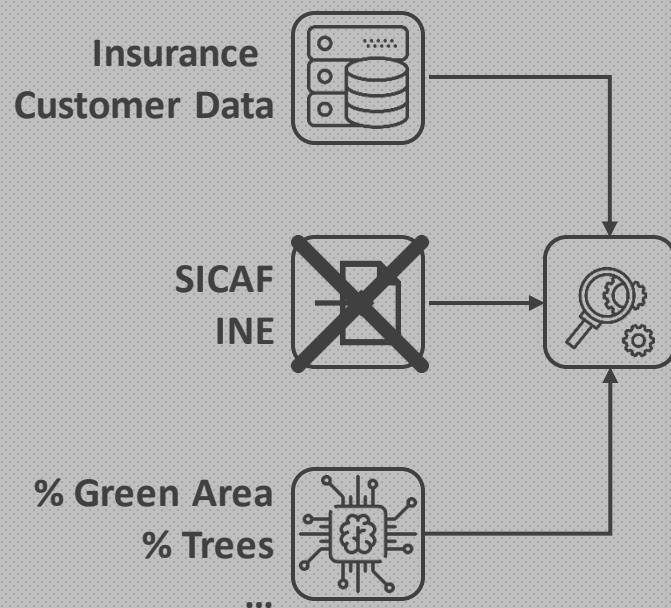
A dark, star-filled night sky serves as the background for this image. In the lower center, the silhouette of a person stands on a small, rocky outcrop. The horizon line is visible at the bottom, showing the dark outlines of trees or bushes. The sky is filled with numerous stars of varying brightness, with a denser concentration towards the center.

# How Incomplete is Your Data?

# Propensity to own a Dog!

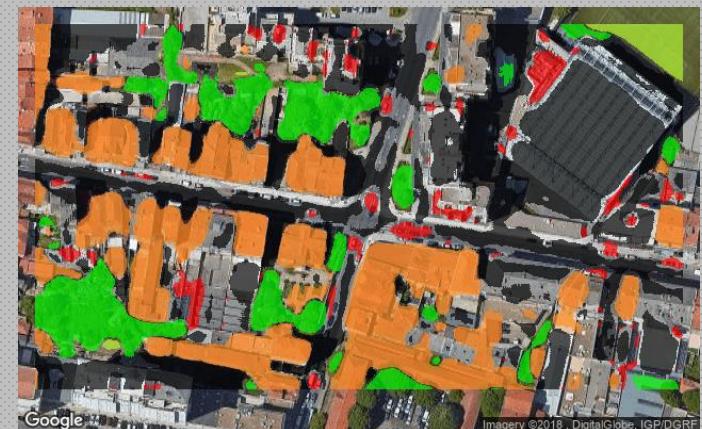


# Propensity to own a Dog!



% Green Area

85%

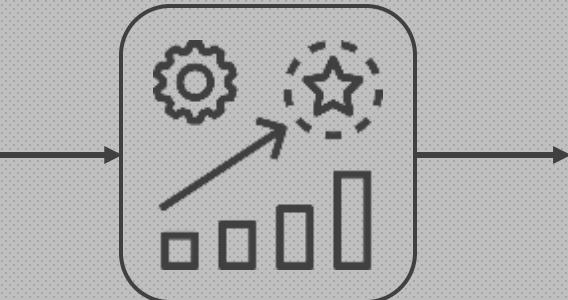


% Green Area

23%

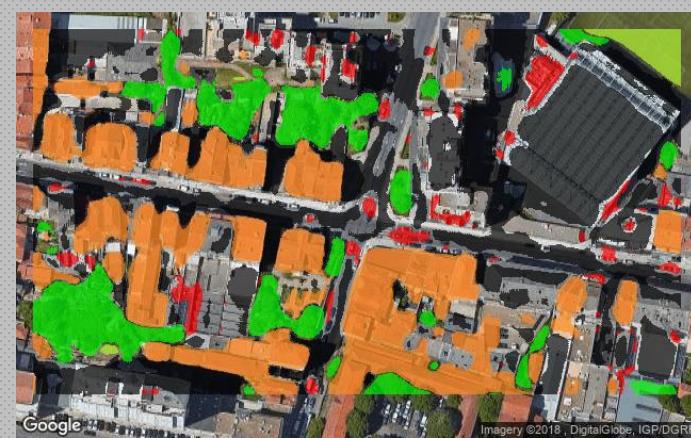
# Propensity to own a Dog!

Training Data



Keras using a  
**Full Convolutional Neural  
Network Architecture**  
(Conv Layer, Max-Pooling, ...)

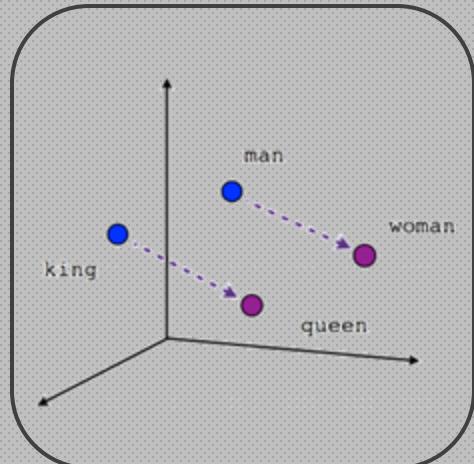
Classification Results



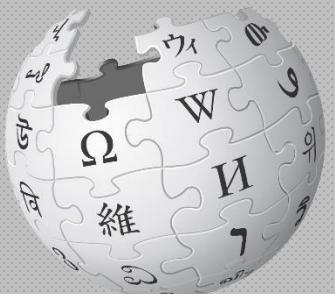
It looks like you ate the whole WIKIPEDIA!!!

L<sub>1</sub> E<sub>1</sub> A<sub>1</sub> R<sub>1</sub> N<sub>1</sub>

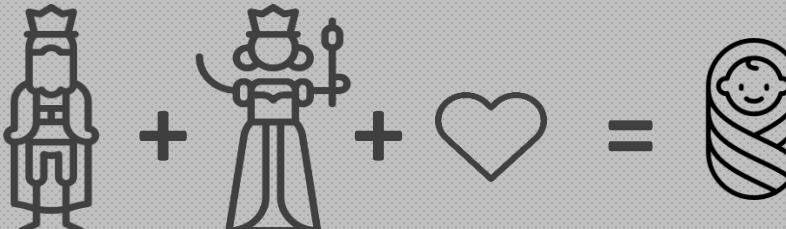
# Topics Classification using Word Embedding



Word Embedding  
GENSIM



```
In [59]: model_w2v.most_similar(positive=['rei', 'mulher'], negative=['homem'])[0]  
Out[59]: ('rainha', 0.7111630439758301)
```



```
In [57]: model_w2v.most_similar(positive=['rei', 'rainha', 'sexo'], negative=[])[0]  
Out[57]: ('herdeiro', 0.7049692273139954)
```





Problem



claimer-The speaker

# Thank you!

29<sup>th</sup> of October 2018



Francesco Costigliola