

SOMACHINE

Machine Learning, Big Data, and Deep Learning in Astronomy

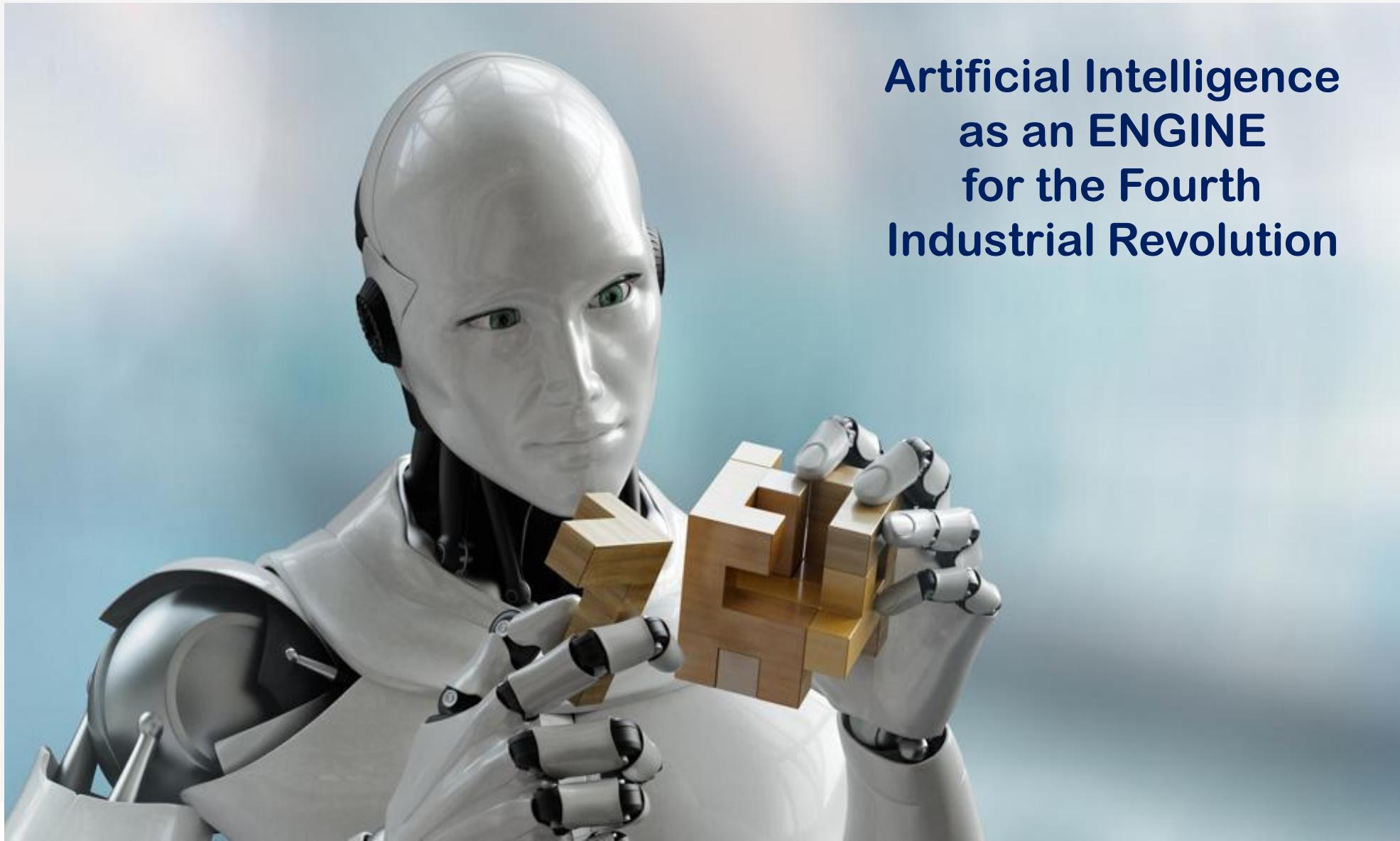
A Severo Ochoa School of the Instituto de Astrofísica de Andalucía (CSIC)

Placing Artificial Intelligence and Machine Learning in the context

Francisco Herrera



Artificial Intelligence: Fourth Industrial Revolution



Placing Artificial Intelligence and Machine Learning in the context. A personal perspective

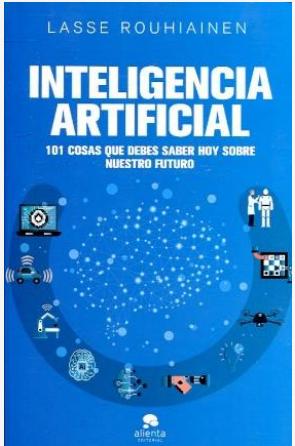
- ARTIFICIAL INTELLIGENCE:
FROM THEORY TO PRACTICE 70 YEARS OF HISTORY
- ARTIFICIAL INTELLIGENCE: FROM DATA TO ALGORITHM DESIGN
- ARTIFICIAL INTELLIGENCE: AI AND INDUSTRY. PERSPECTIVES FOR THIS DECADE
- ARTIFICIAL INTELLIGENCE: A DEEP LOOK AT SMART DATA, BIG DATA, DEEP LEARNING, AND TRUSTWORTHINESS (INTERPRETABILITY, ETHICS ...)

Placing Artificial Intelligence and Machine Learning in the context. A personal perspective

**ARTIFICIAL INTELLIGENCE: FROM THEORY TO PRACTICE 70
YEARS OF HISTORY**

"Artificial intelligence is the new electricity" A. Ng

Artificial Intelligence: " Increasingly capable non-thinking machines"

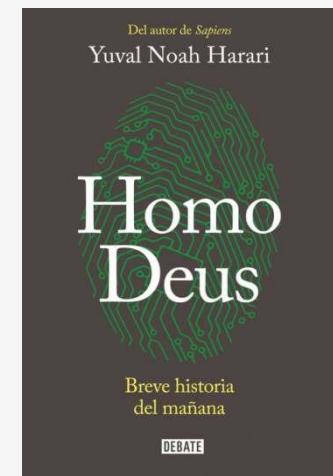


"The world will change thanks to the growth of Artificial Intelligence"

Lasse Rouhiainen
(*Inteligencia Artificial*, Alienta Ed, 2018)

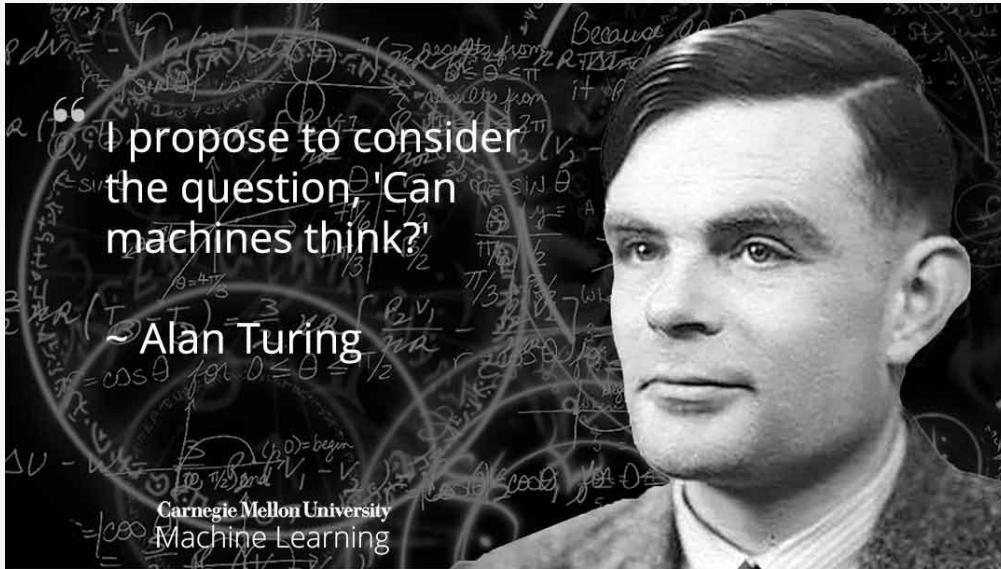


Yuval Noah Harari (Oct, 2016):
"The world is going to change radically thanks to algorithms, big data and artificial intelligence"



Artificial Intelligence: 70 YEARS OF HISTORY

1947-50



I propose to consider
the question, 'Can
machines think?'

~Alan Turing

1955-56



John McCarthy (1955) (Stanford)
Dartmouth Conference (1956)

**Artificial Intelligence: "the science
and engineering of making machines
that behave in a way that we would
call intelligent if humans had that
behaviour".**

Artificial Intelligence: Successful Technology

AI as a successful technology in the industry: 7 of the 10 largest companies in the world in terms of stock market capital are users or developers of artificial intelligence

Society > Economy

The 100 largest companies in the world by market capitalization in 2020
(in billion U.S. dollars)

Ranking of the companies rank 1 to 100	Market capitalization in billion U.S. dollars
Saudi Arabian Oil Company (Saudi Aramco) (Saudi Arabia)	1,684.8
Microsoft (United States)	1,359
Apple (United States)	1,285.5
Amazon (United States)	1,233.4
Alphabet (United States)	919.3
Facebook (United States)	583.7
Alibaba (China)	545.4
Tencent Holdings (China)	509.7
Berkshire Hathaway (United States)	455.4
Johnson & Johnson (United States)	395.3

Saudi Aramco (oil)

Microsoft

Apple

Amazon

Alphabet

Facebook

Alibaba

Tencent Holdings (Internet, Wechat, ...)

Berkshire Hathaway (manufactures)

Johnson & Johnson (Health, pharmaceutical, perfumes)

Artificial Intelligence: Successful Technology

Software is eating the world



facebook.

\$2.02T

\$1.60T

\$1.59T

\$1.01T

\$754B

~\$7 Trillions

ANDALUCÍA ES DIGITAL



~\$715 Billions

Artificial Intelligence: 70 YEARS OF HISTORY

From theory to practice. From algorithms to data



“Artificial intelligence is the new electricity” Andrew Ng



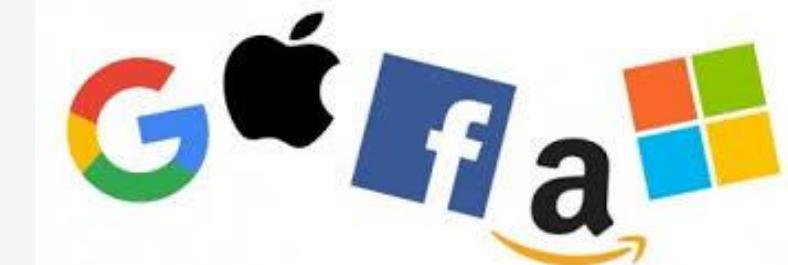
John McCarthy (1955) (Stanford)
Dartmouth Conference (1956)

Artificial Intelligence: "the science and engineering of making machines that behave in a way that we would call intelligent if humans had that behaviour".



Fourth Industrial Revolution

AI as a successful technology in the industry: 7 of the 10 largest companies in the world in terms of stock market capital are users or developers of artificial intelligence



Artificial Intelligence: 25 last years (1996 – 2020).

Paradigm change, from algorithm to data. The era of big data

1996 – 1997 Kasparov vs. Deep Blue



2005 - Driverless car



Stanley (Standord), DARPA grand Challenge - 2005 (S. Thrun)



2016: AlphaGo: Deep Learning, Learning playing to Go



DeepMind's groundbreaking artificial intelligence, [AlphaGo](#), defeated [Lee Sedol](#) 9p in the final game of the [Google DeepMind Challenge Match](#) on March 15, 2016, winning the five game match with a 4–1 score.



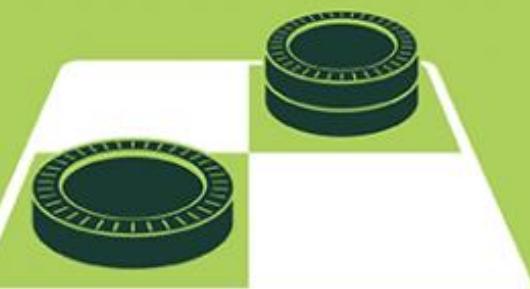
Artificial Intelligence: Paradigm change, from algorithm to data

A natural evolution towards AI and ML is taking place.

Ecloxion of Deep Learning

ARTIFICIAL INTELLIGENCE

Early artificial intelligence stirs excitement.



MACHINE LEARNING

Machine learning begins to flourish.



DEEP LEARNING

Deep learning breakthroughs drive AI boom.



1950's

1960's

1970's

1980's

1990's

2000's

2010's

Artificial Intelligence: Paradigm change, from algorithm to data

A natural evolution towards AI and ML is taking place.

Ecloxion of Deep Learning

ARTIFICIAL INTELLIGENCE

Early artificial intelligence stirs excitement.



1950's

1960's

1970's

1980's

1990's

2000's

2010's

MACHINE LEARNING

Machine learning begins to flourish.

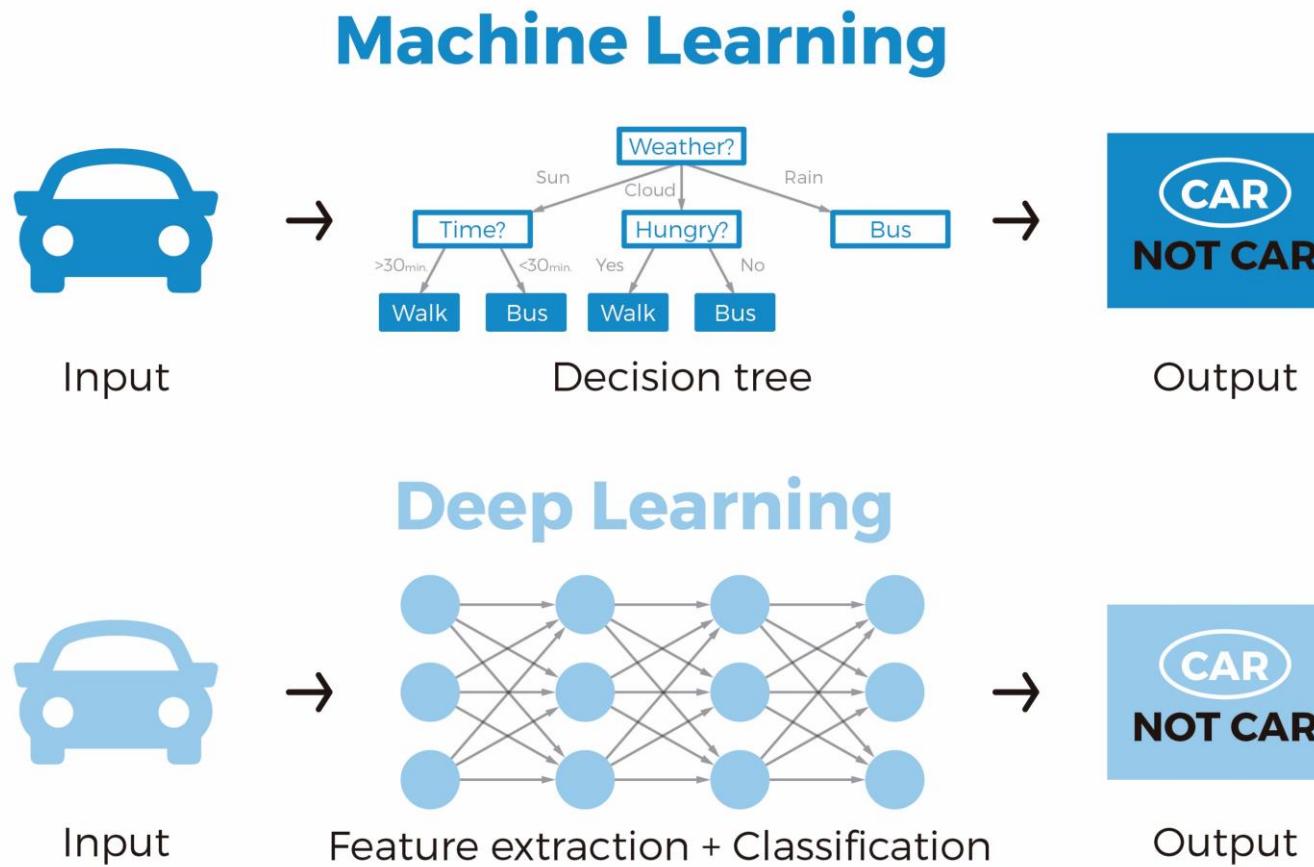


DEEP LEARNING

Deep learning breakthroughs drive AI boom.



Artificial Intelligence: Paradigm change, from algorithm to data



Artificial Intelligence: Paradigm change, from algorithm to data

**Big Data and Deep Learning emphasize
the leap from theory to practice,
focusing on the data/problem**



* Datos de septiembre de 2019
Fuente: CB Insights

Artificial Intelligence: Paradigm change, from algorithm to data

Big Data and Deep Learning emphasize the leap from theory to practice, focusing on the data/problem

Data is the central element in many problems/challenges (data are the new gold of the 21st century): The importance of data, data availability, labelling, ...



Artificial Intelligence: The era of big data

Data and Artificial Intelligence



Artificial Intelligence: The era of big data. Data and AI

The **Big Data** provides an enormous amount of data that feeds the algorithms of **Machine Learning (Artificial Intelligence)**, and allows the creation of more complex models with greater precision.



Intelligent systems powered by data to convert data into knowledge

Artificial Intelligence: Data and AI

A DEFINITION OF AI: MAIN CAPABILITIES AND DISCIPLINES

High-Level Expert Group on Artificial Intelligence

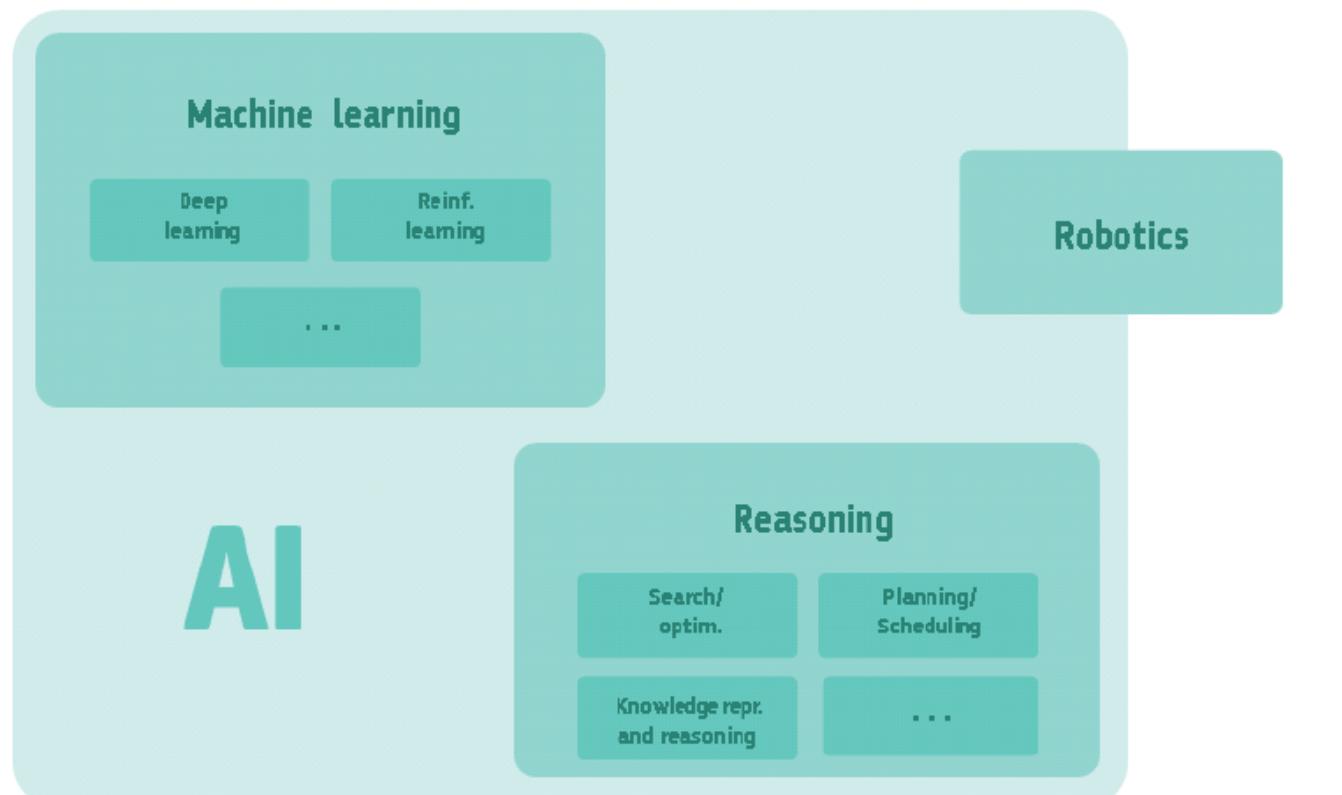
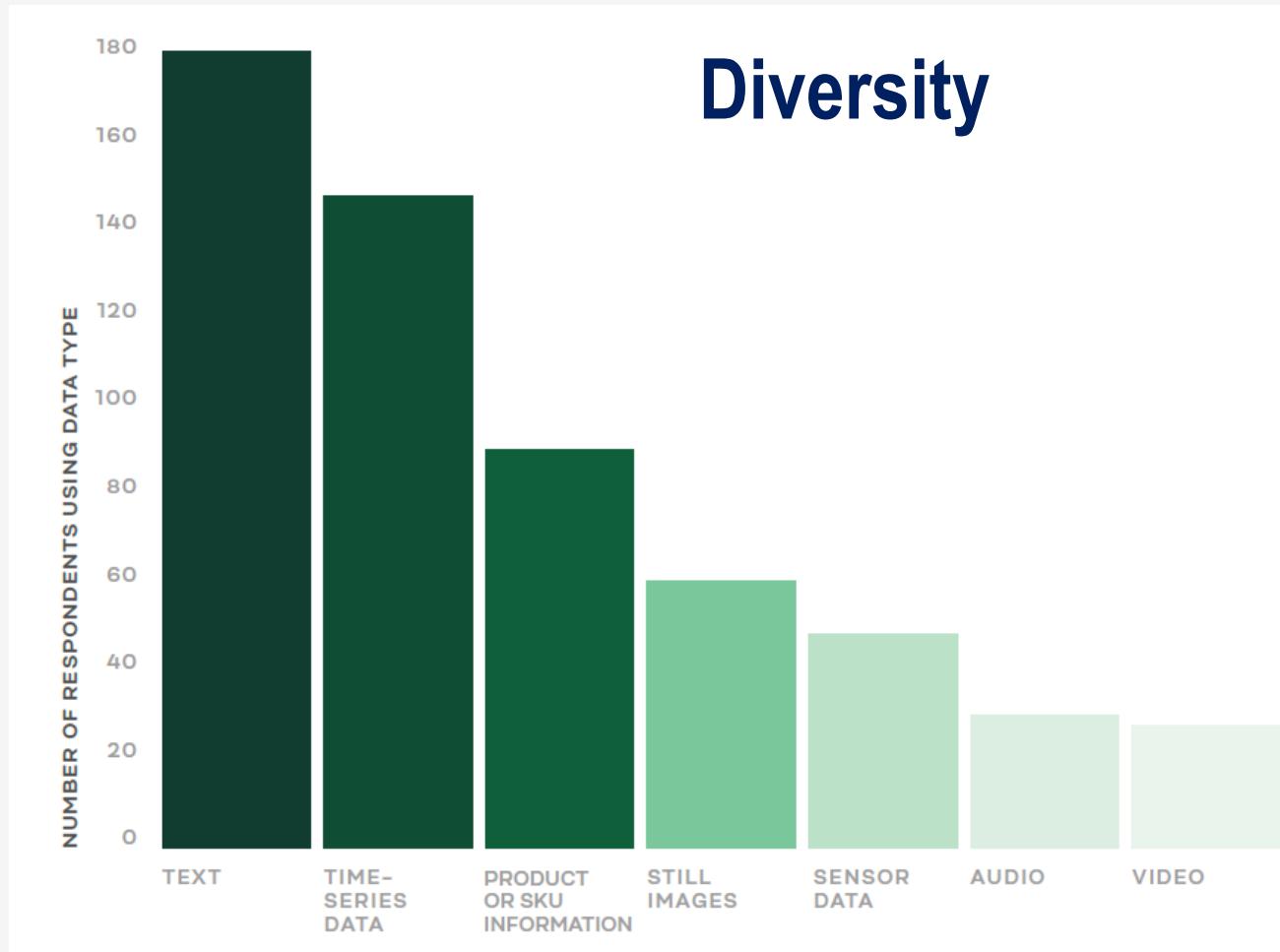


Figure 2: A simplified overview of AI's sub-disciplines and their relationship.

Both machine learning and reasoning include many other techniques, and robotics includes techniques that are outside AI. The whole of AI falls within the computer science discipline.

“Artificial intelligence (AI) systems are software (and possibly also hardware) systems designed by humans that, given a complex goal, act in the physical or digital dimension by perceiving their environment through **data acquisition**, interpreting the collected structured or unstructured data, **reasoning on the knowledge**, or processing the information, derived from this data and **deciding the best action(s)** to take to achieve the given goal.”

Artificial Intelligence: Data and AI

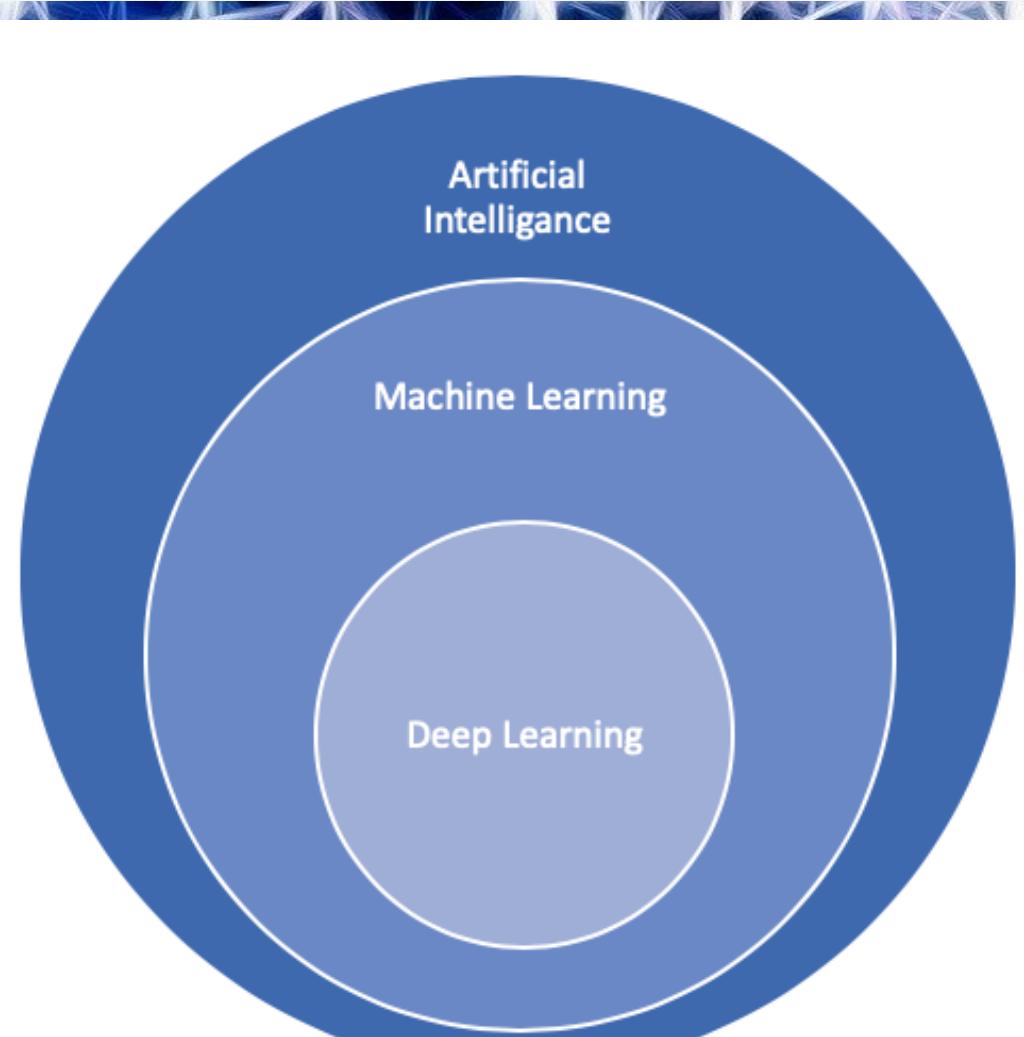


© Data Scientist report 2018.
COMMONLY USED DATA TYPES

Placing Artificial Intelligence and Machine Learning in the context. A personal perspective

ARTIFICIAL INTELLIGENCE: FROM DATA TO
ALGORITHM DESIGN

Artificial Intelligence: From data to algorithm design



Big Data (explosion of data from which to extract knowledge) and Deep Learning (automation of the process, and "summit" in algorithmic precision)

Artificial Intelligence: From data to algorithm design

From theory to practice

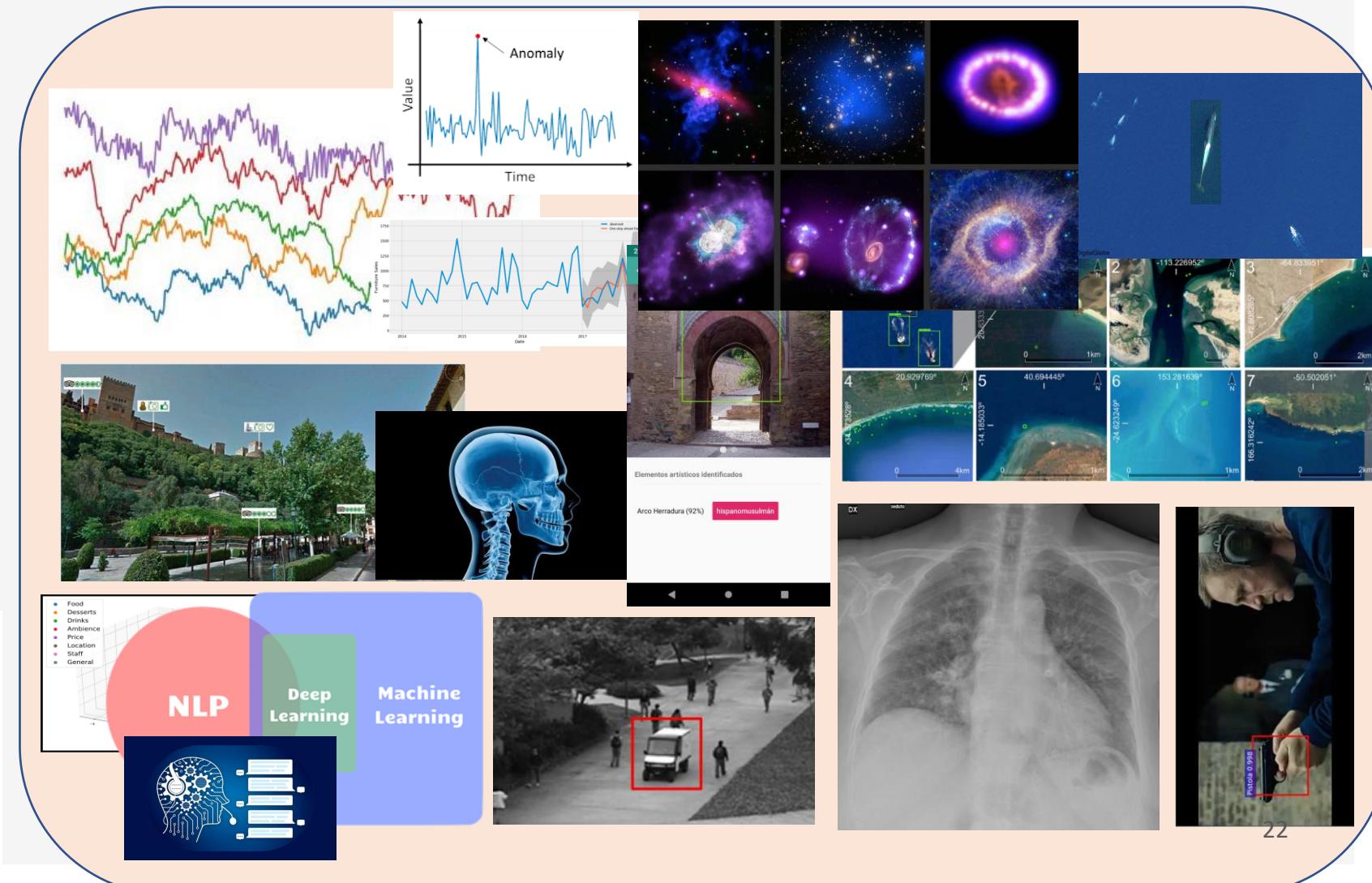


Information Fusion

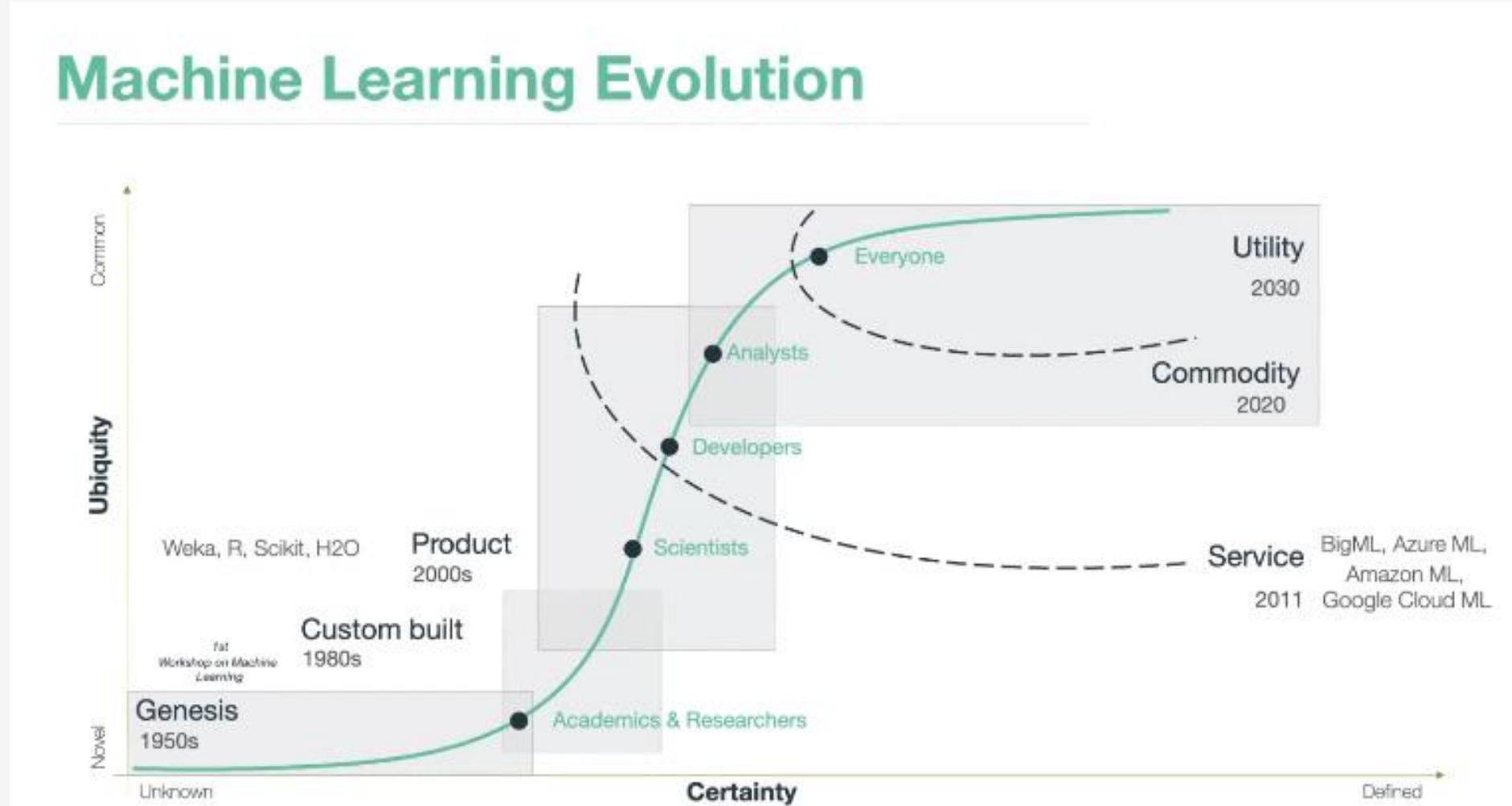
Volume 62, October 2020, Pages 73-80



MNIST-NET10: A heterogeneous deep networks fusion based on the degree of certainty to reach 0.1% error rate. Ensembles overview and proposal



Artificial Intelligence: From data to algorithm design



Artificial Intelligence: From data to algorithm design

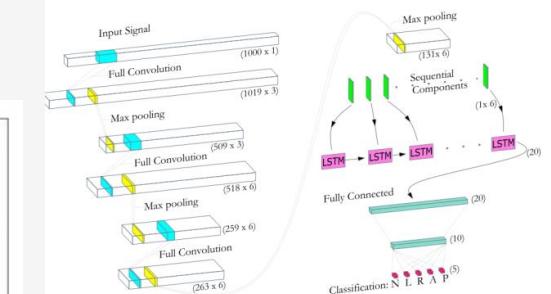
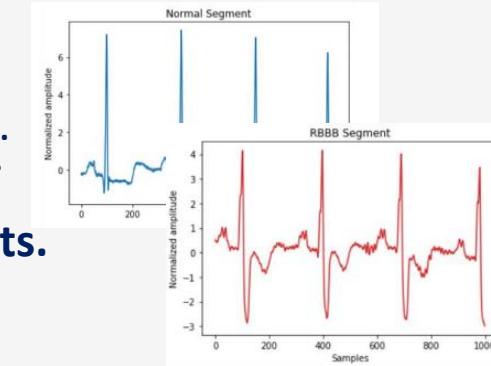
Case of Analysis

Shu Lih Oh, Eddie Y.K. Ng, Ru San Tan, U Rajendra Acharya.

Automated diagnosis of arrhythmia using combination of CNN and LSTM techniques with variable length heart beats.

Computers in Biology and Medicine 102 (2018) 278-287

<https://www.sciencedirect.com/science/article/abs/pii/S0010482518301446>



Abstract: Arrhythmia is a cardiac conduction disorder characterized by irregular heartbeats.

... we propose an automated system using a combination of convolutional neural network (CNN) and long short-term memory (LSTM) for diagnosis of normal sinus rhythm, left bundle branch block (LBBB), right bundle branch block (RBBB), atrial premature beats (APB) and premature ventricular contraction (PVC) on ECG signals. The novelty of this work is that we used ECG segments of variable length from the MIT-BIT arrhythmia physio bank database.

Very active area: More than 25 algorithms, a lot of applications

Placing Artificial Intelligence and Machine Learning in the context. A personal perspective

ARTIFICIAL INTELLIGENCE: AI AND INDUSTRY.
PERSPECTIVES FOR THIS DECADE

Artificial Intelligence: AI and Industry. Perspectives for this Decade

“Awareness of artificial intelligence is almost universal”

It is one thing to know that it exists and another to put it into practice.

Two out of five companies use at least one of these technologies, one intends to use it and two neither use nor plan to use it.

Artificial Intelligence: AI and Industry. Perspectives for this Decade

Survey of more than 9,000 businesses developed by Ipsos for the European Commission.

42% of European companies use at least one AI tool

Out of every five companies, two use at least one of these technologies, one intends to use it and two neither use nor plan to use it.

In this last group, Spain ranks above the European average. More than half of the companies surveyed (51%) have no intention of jumping on the bandwagon of artificial intelligence.

"Adoption is still relatively low, with each technology having adoption rates of 13% or less, but this also indicates the diversity of types of artificial intelligence technologies being used by European companies"

Artificial Intelligence: AI and Industry. Perspectives for this Decade

This study was carried out by Ipsos for the European Commission. It has found that 42% of enterprises currently use at least one AI technology, a quarter of them use at least two types, and 18% have plans to adopt AI technologies in the next two years.

Three key internal barriers to AI adoption are difficulties in hiring new staff with

- the right skills (57%),
- the cost of adoption (52%) and
- the cost of adapting operational processes (49%).

[European Commission](#) > [Strategy](#) > [Shaping Europe's digital future](#) > [Reports and studies](#) >

Shaping Europe's digital future

REPORT / STUDY | 28 July 2020

European enterprise survey on the use of technologies based on artificial intelligence

European Comision

Artificial Intelligence: AI and Industry. Perspectives for this Decade

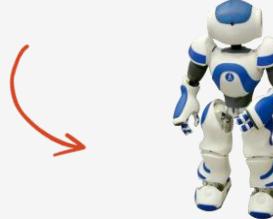
1 Automation and process optimization

Process automation



2 Autonomous machines

Autonomous machines



3 Artificial Vision and Deep Learning

Artificial Vision

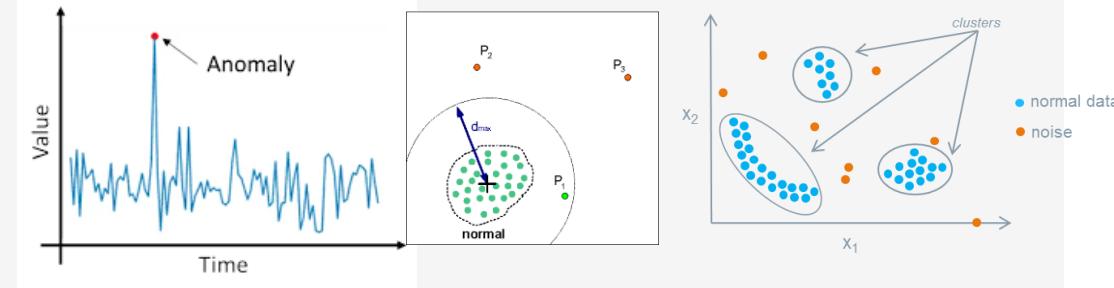


Artificial Intelligence: AI and Industry. Perspectives for this Decade

4

Anomaly Selection (IoT, sensors)

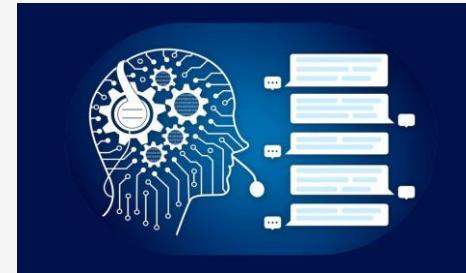
Anomalies/outliers



5

Sentiments Analysis and NLP

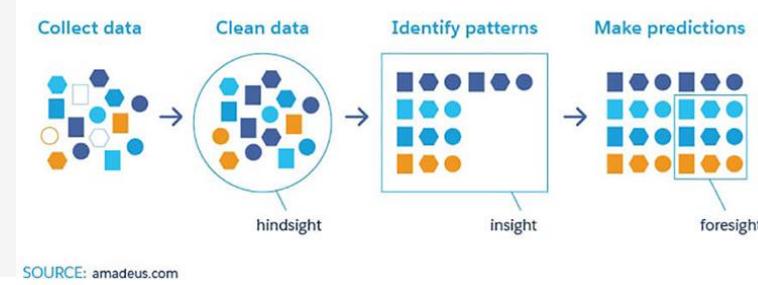
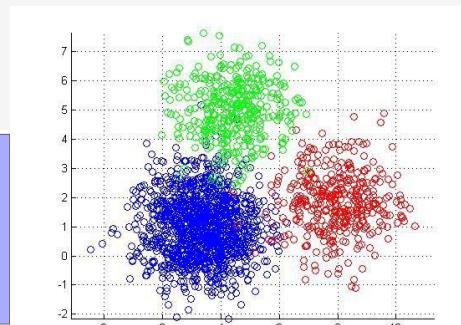
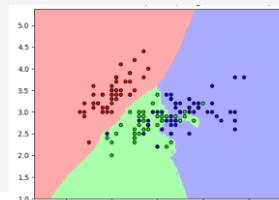
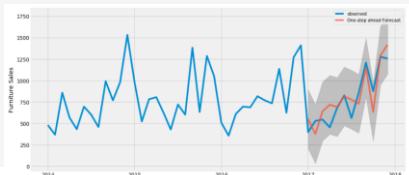
Sentiments Analysis and NLP



6

Prediction

Prediction (regression, classification)

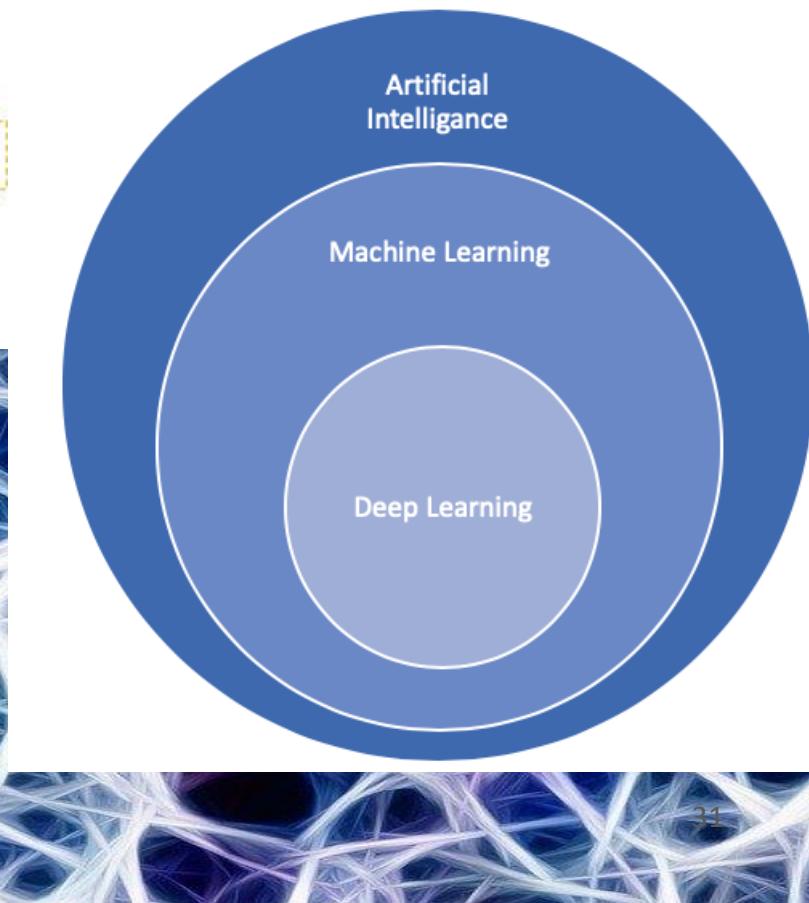
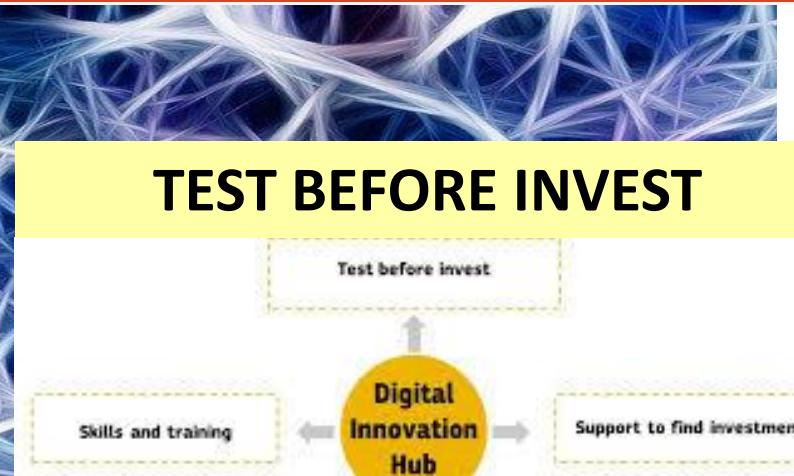


Artificial Intelligence: AI and Industry. European perspectives for this Decade



BIG DATA

Big data is the new gold.
The importance of data, data availability, labelling, ...



Artificial Intelligence: AI and Industry. European perspectives for this Decade

Three key internal barriers to AI adoption are difficulties in hiring new staff with

- the right skills (57%): **FORMATION**
- the cost of adoption (52%): **TEST BEFORE INVEST. CONCEPT AND VALIDATION LABS.**
- the cost of adapting operational processes (49%). **TEST BEFORE INVEST. CONCEPT AND VALIDATION LABS**



European Commission > Strategy > Shaping Europe's digital future > Reports and studies >

Shaping Europe's digital future

REPORT / STUDY | 28 July 2020

European enterprise survey on the use of technologies based on artificial intelligence

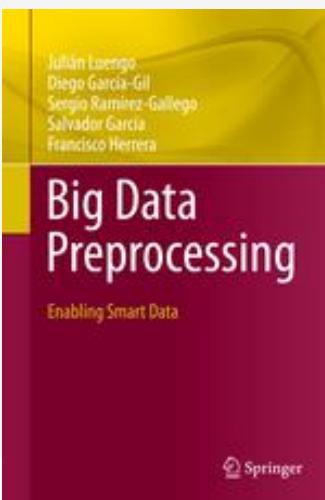
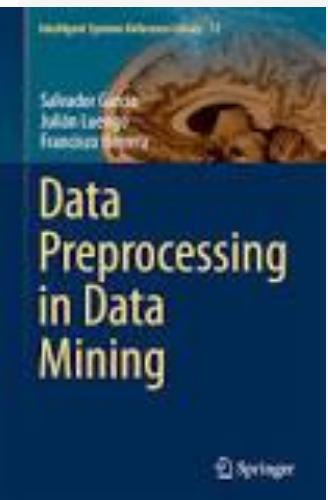


Placing Artificial Intelligence and Machine Learning in the context. A personal perspective

ARTIFICIAL INTELLIGENCE: A DEEP LOOK AT SMART
DATA, BIG DATA, DEEP LEARNING, AND
TRUSTWORTHINESS (INTERPRETABILITY, ETHICS ...)

Artificial Intelligence: Smart Data. Approaching Data

Challenge How To Ensure Data Quality For Machine Learning And AI Projects?



Driving Quality in Your AI Training Data

- ✓ Potential bias in datasets and how to prevent it
- ✓ Continuous improvement via machine learning
- ✓ Definition of quality for annotated data
- ✓ Ground truth and auditing to ensure quality
- ✓ Data Quality for singular problems

Artificial Intelligence: Big Data Insight. Approaching Data

Big Data Insight

The necessary binomial in big data beyond technology,
approaching data and algorithms

Going deeper into the data!
Finding now research goals!



Cloud/clusters
Hadoop Ecosystem: HDFS,
Spark, Flink, ...



Smart Data
Big Data Preprocessing



Big Learning: Machine Learning and Artificial Intelligence Scalable, efficient effective and explainable algorithms

Inteligencia Artificial: Big Data Insight. Approaching Data

Beyond technology, going deeper into the data!

Challenges of Big Data analysis

Jianqing Fan^{1,*}, Fang Han² and Han Liu¹

National Science Review
1: 293–314, 2014
doi: 10.1093/nsr/nwt032
Advance access publication 6 February 2014

What are the challenges of analyzing Big Data?

Big data are characterized by high dimensionality and large sample size.

These two features raise three unique challenges: (i) high dimensionality brings noise accumulation, spurious correlations and incidental

homogeneity, (ii) high dimensionality combined with large sample size creates computational cost and algorithm instability, (iii) data aggregated from multiple technologies. This creates irregularities and robust procedures.

Garbage In, Garbage Out!

Noise accumulation

Spurious correlation

Large sample brings
Redundancy
Irregularities
Imperfect data

A new look into data and the MAPS based analysis!



(Big) Data
preprocessing

Towards
Quality Data

Artificial Intelligence: Deep Learning Insight

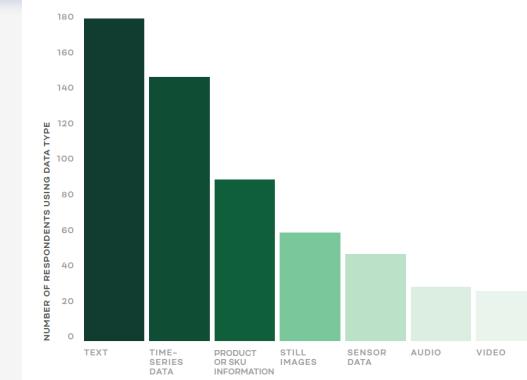
Approaching Data



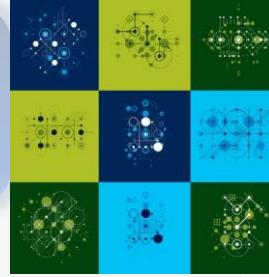
GPU



Images/IoT/NLP for Data
Data Augmentation and
Big Data Preprocessing



Neural Networks:
CNN, LTSM, ...



Artificial Intelligence: Trustworthy artificial intelligence

Europe (EU): 2019



Figure 2: Interrelationship of the seven requirements: all are of equal importance, support each other, and should be implemented and evaluated throughout the AI system's lifecycle

The Alan Turing Institute

Understanding artificial intelligence ethics and safety

A guide for the responsible design and implementation of AI systems in the public sector

Dr David Leslie
Public Policy Programme

Strategic Research, Innovation and Deployment Agenda for an AI PPP

A focal point for collaboration on Artificial Intelligence, Data and Robotics

Second Consultation Release
September 2019

A joint initiative by

BDV BIG DATA VALUE ASSOCIATION

eu ROBOTICS

A black and white photograph showing several people walking across a crosswalk. The background is a light-colored surface with dark diagonal stripes.

Artificial Intelligence: Trustworthy artificial intelligence. Approaching algorithms: explainability

Explainability

Place the **audience** as key aspect to be considered when explaining an AI model.

*functioning clear or
easy to understand.*

Cambridge Dictionary of English Language,
an explanation is “*the details or reasons
that someone gives to make something
clear or easy to understand*”



*Given an audience, an explainable AI is one that
produces details or reasons to make its functioning
clear or easy to understand.*

Information Fusion 58 (2020) 82–115



Contents lists available at ScienceDirect

Information Fusion

journal homepage: www.elsevier.com/locate/inffus



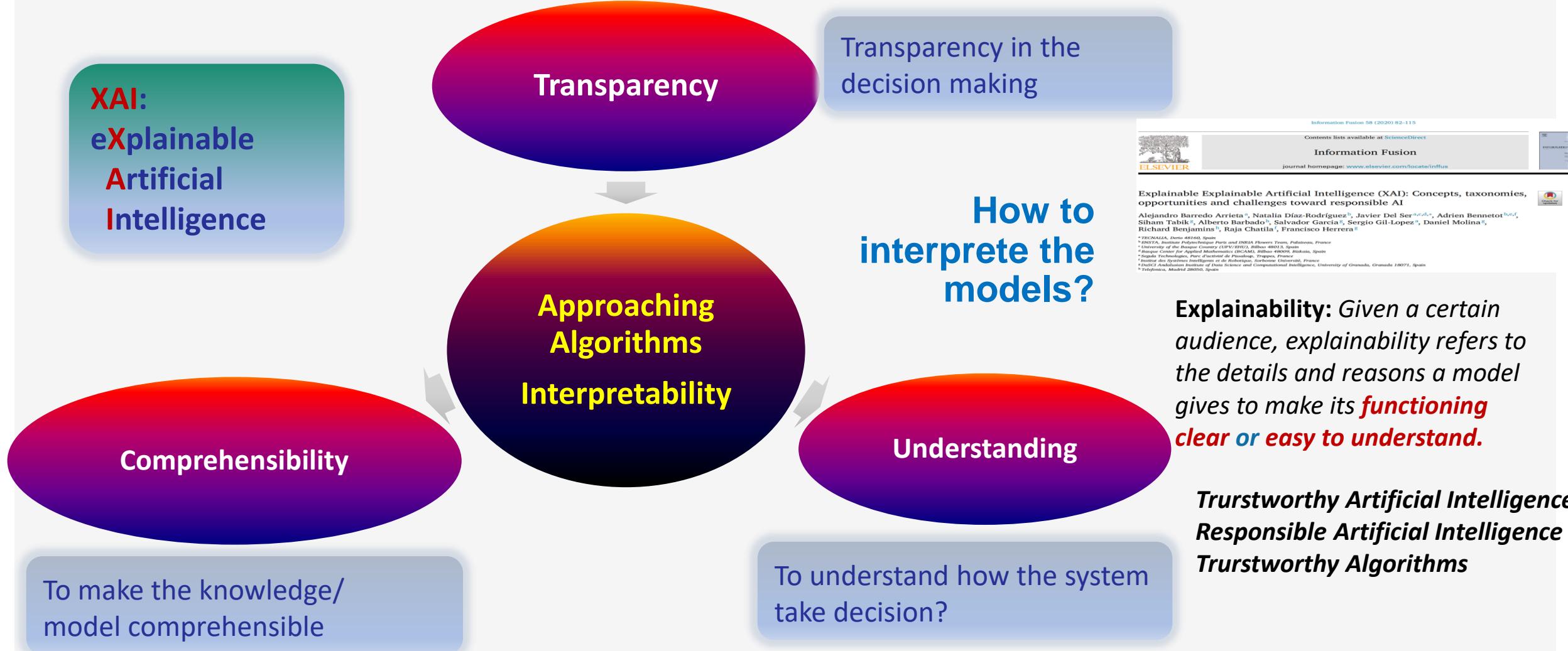
Explainable Artificial Intelligence (XAI): Concepts, taxonomies,
opportunities and challenges toward responsible AI

Alejandro Barredo Arrieta^a, Natalia Díaz-Rodríguez^b, Javier Del Ser^{a,c,d,*}, Adrien Bennetot^{b,e,f},
Siham Tabik^g, Alberto Barbado^h, Salvador Garcia^g, Sergio Gil-Lopez^a, Daniel Molina^g,
Richard Benjamins^h, Raja Chatila^f, Francisco Herrera^g

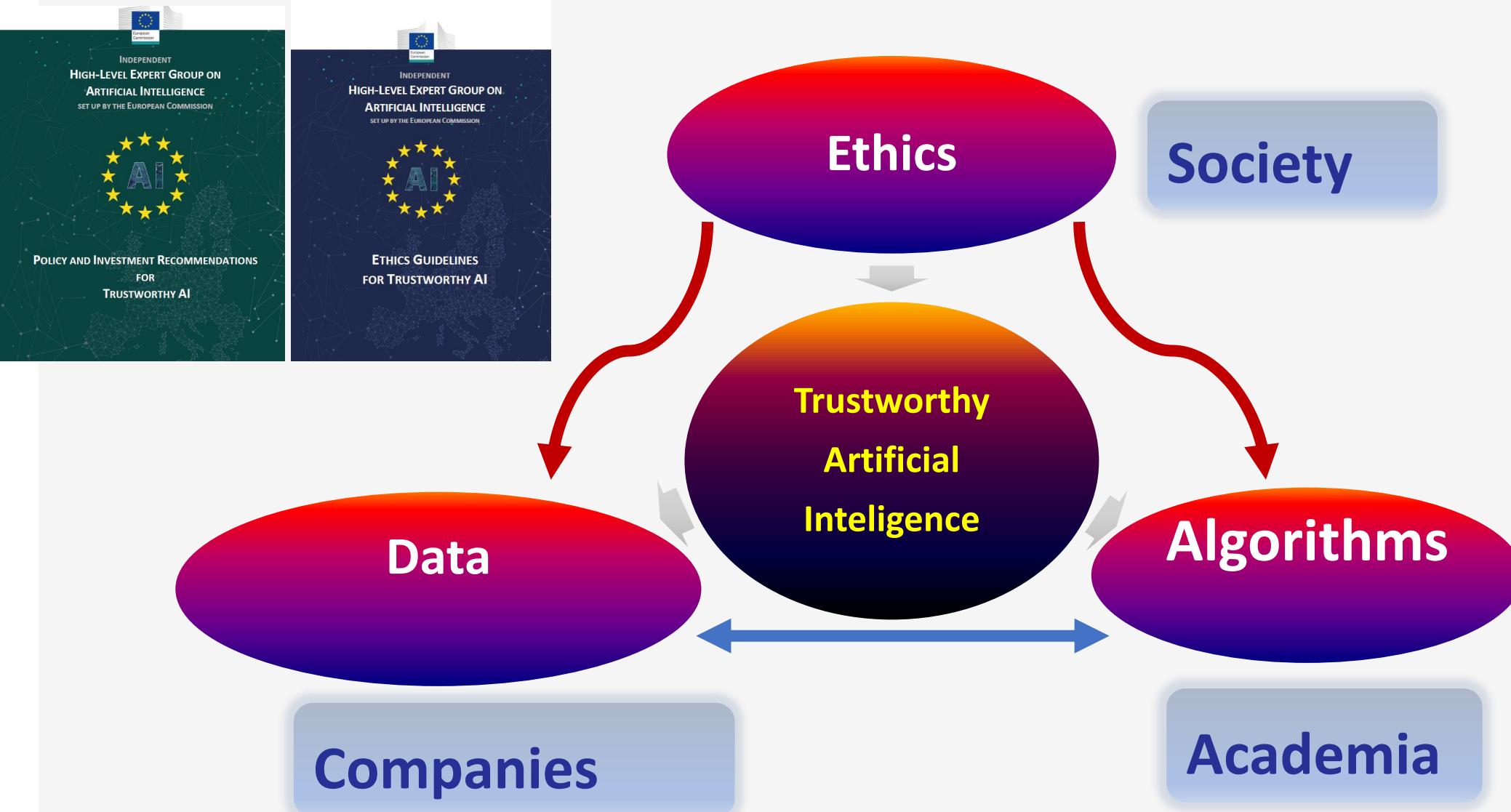


Artificial Intelligence: Trustworthy artificial intelligence

Paying attention to the Interpretability/explainability



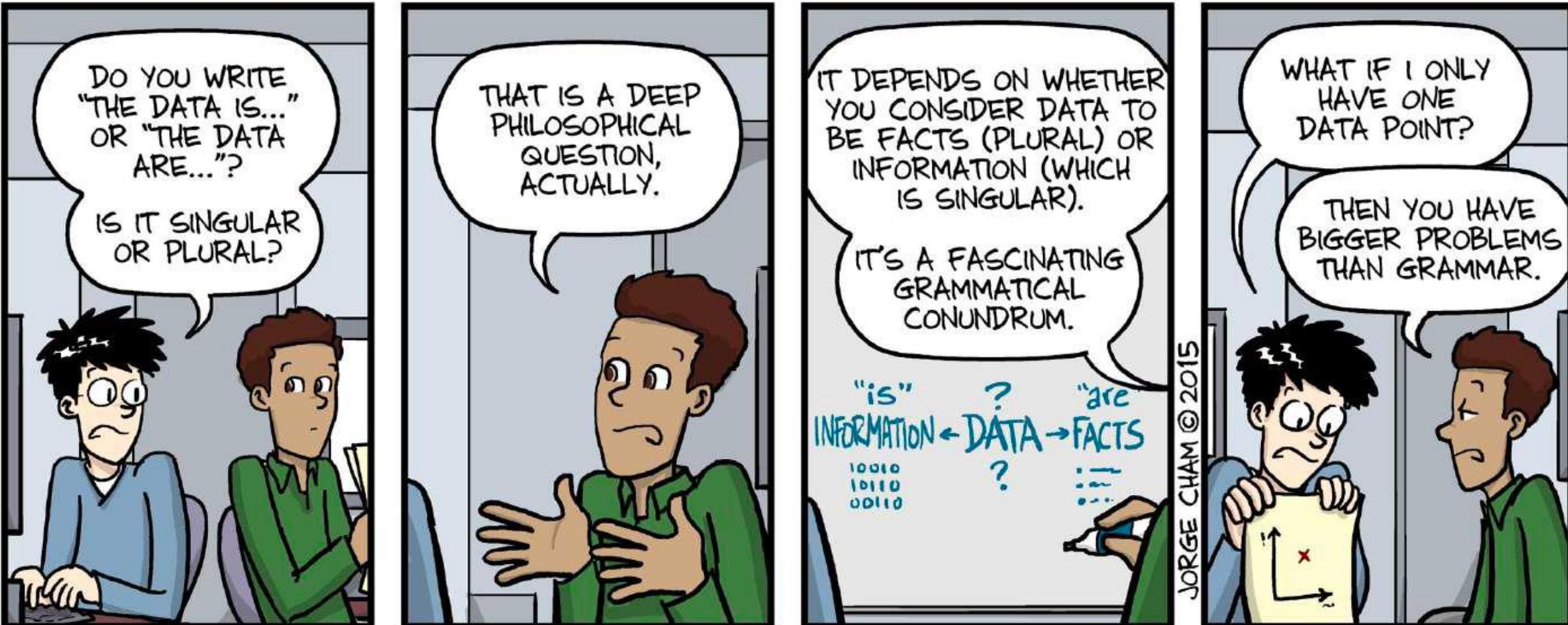
Artificial Intelligence: Trustworthy artificial intelligence



Placing Artificial Intelligence and Machine Learning in the context. A personal perspective

Final Comments

Artificial Intelligence: Final Comments. We need data



JORGE CHAM © 2015

WWW.PHDCOMICS.COM

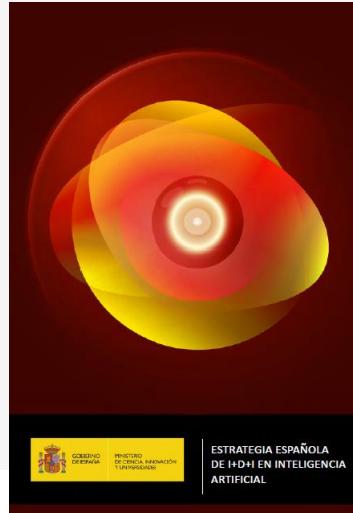
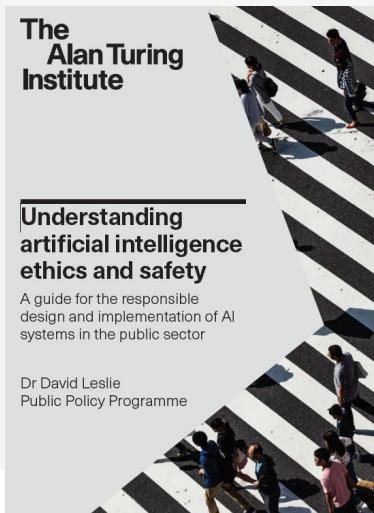
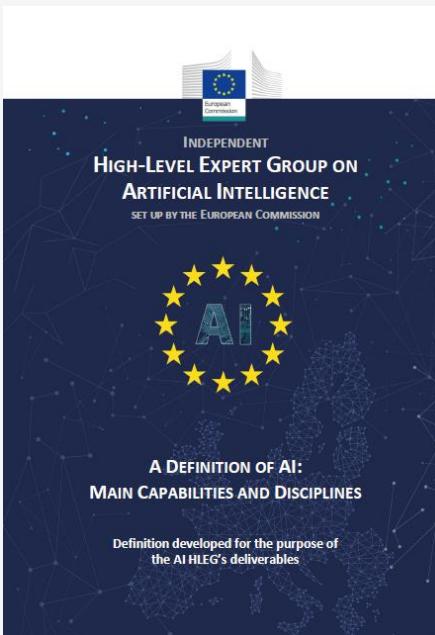
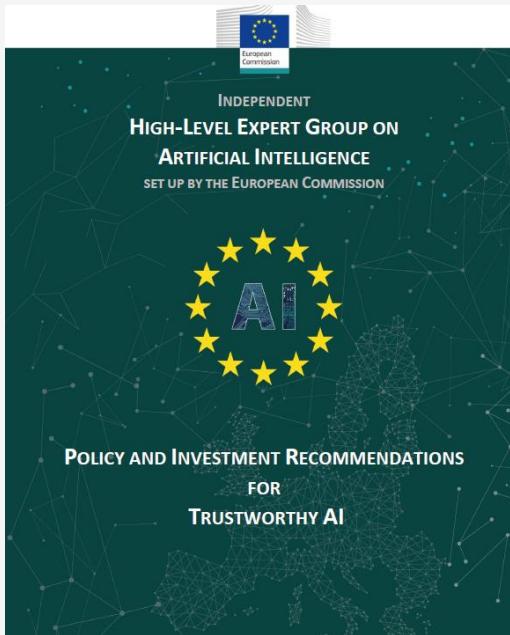
Artificial Intelligence: Final Comments. To put it into practice

Awareness of AI is almost universal.

It's one thing to know that it exists and another to put it into practice.



Artificial Intelligence: Final Comments. European vision



Artificial Intelligence: Final Comments. European vision

Europe (EU): 2019

A DEFINITION OF AI: MAIN CAPABILITIES AND DISCIPLINES

High-Level Expert Group on Artificial Intelligence



“Artificial intelligence (AI) systems are software (and possibly also hardware) systems designed by humans that, given a complex goal, act in the physical or digital dimension by perceiving their **environment** through **data acquisition**, **interpreting the collected** structured or unstructured **data**, reasoning on the knowledge, or processing the information, **derived from this data and** deciding the best action(s) to take to achieve the given goal. “

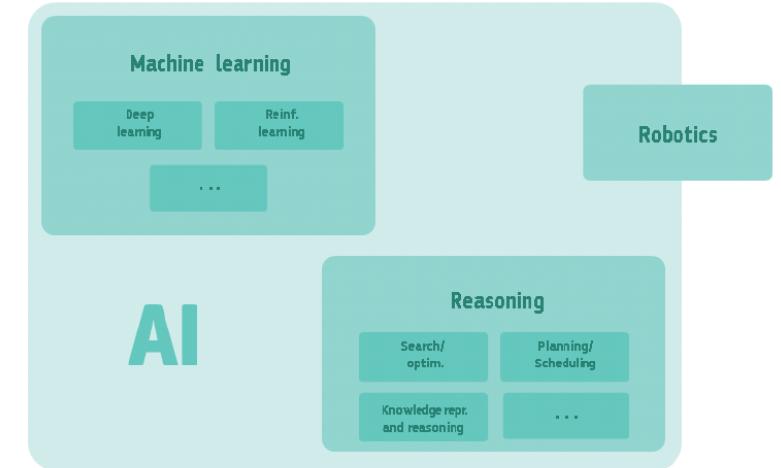


Figure 2: A simplified overview of AI's sub-disciplines and their relationship. Both machine learning and reasoning include many other techniques, and robotics includes techniques that are outside AI. The whole of AI falls within the computer science discipline.

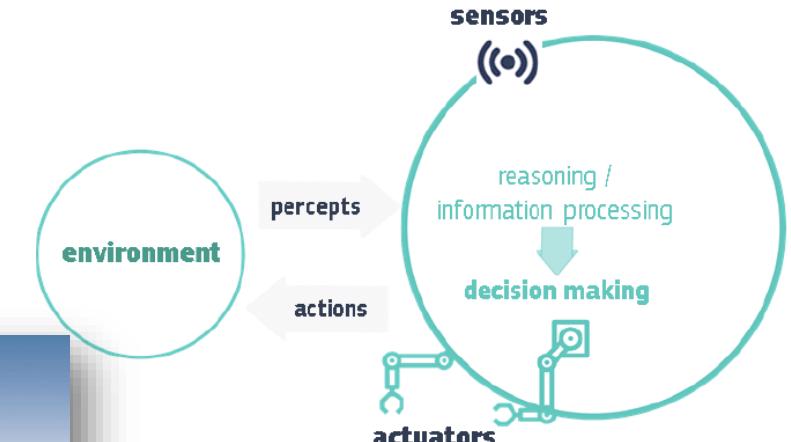


Figure 1: A schematic depiction of an AI system.

Artificial Intelligence: Final Comments

(Japan) Towards Society 5.0

