Barcelona Supercomputing Center
Centro Nacional de Supercomputación

EXCELENCIA SEVERO OCHOA

Big data e inteligencia artificial y HPC en salud y biomedicina

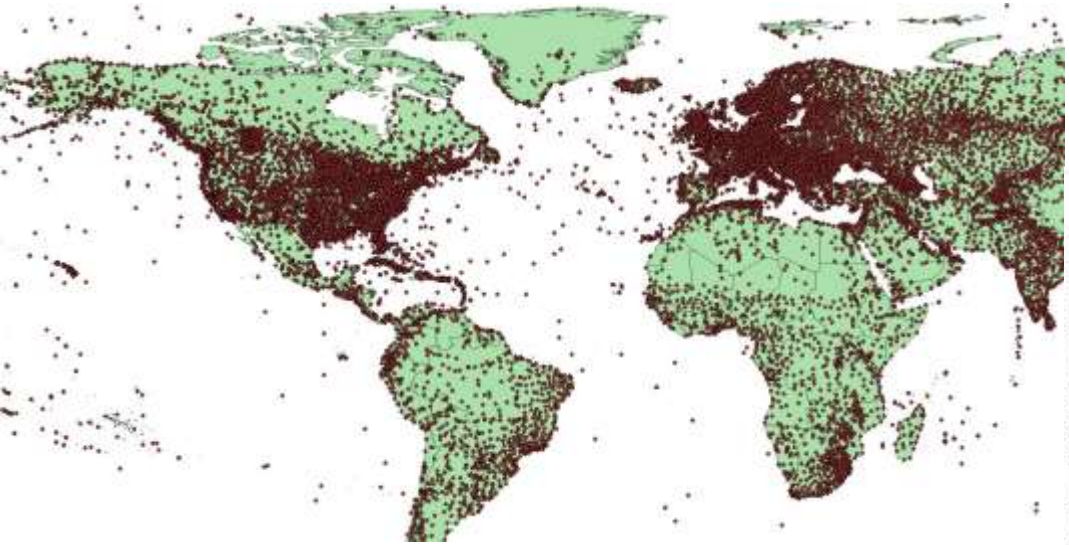
Alfonso Valencia. Ph.D.
ICREA Professor
Director Life Sciences Dept. BSC
Director Spanish Bioinformatics Institute
INB-ISCIII ELIXIR-ES

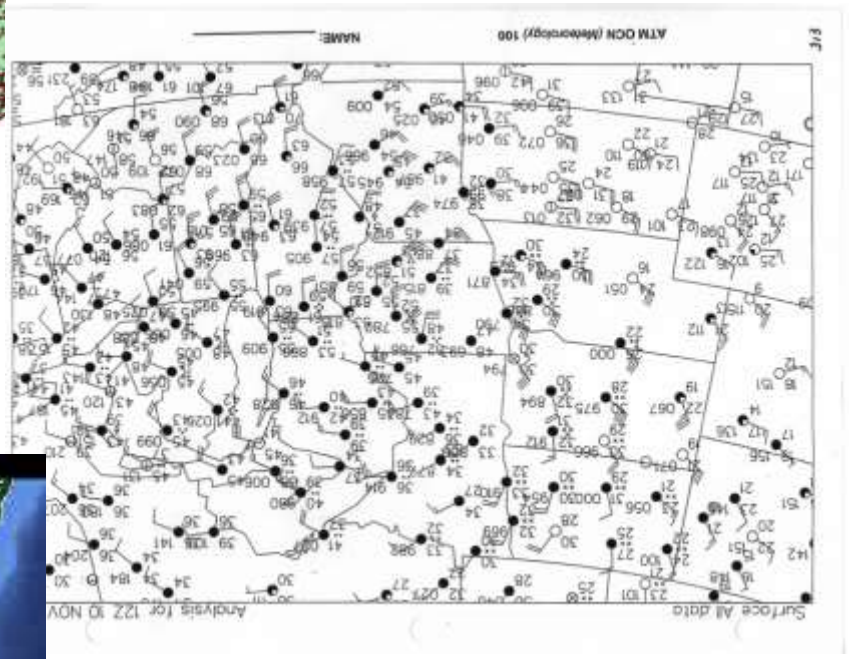
Plan TL InfoDay

21 October 19

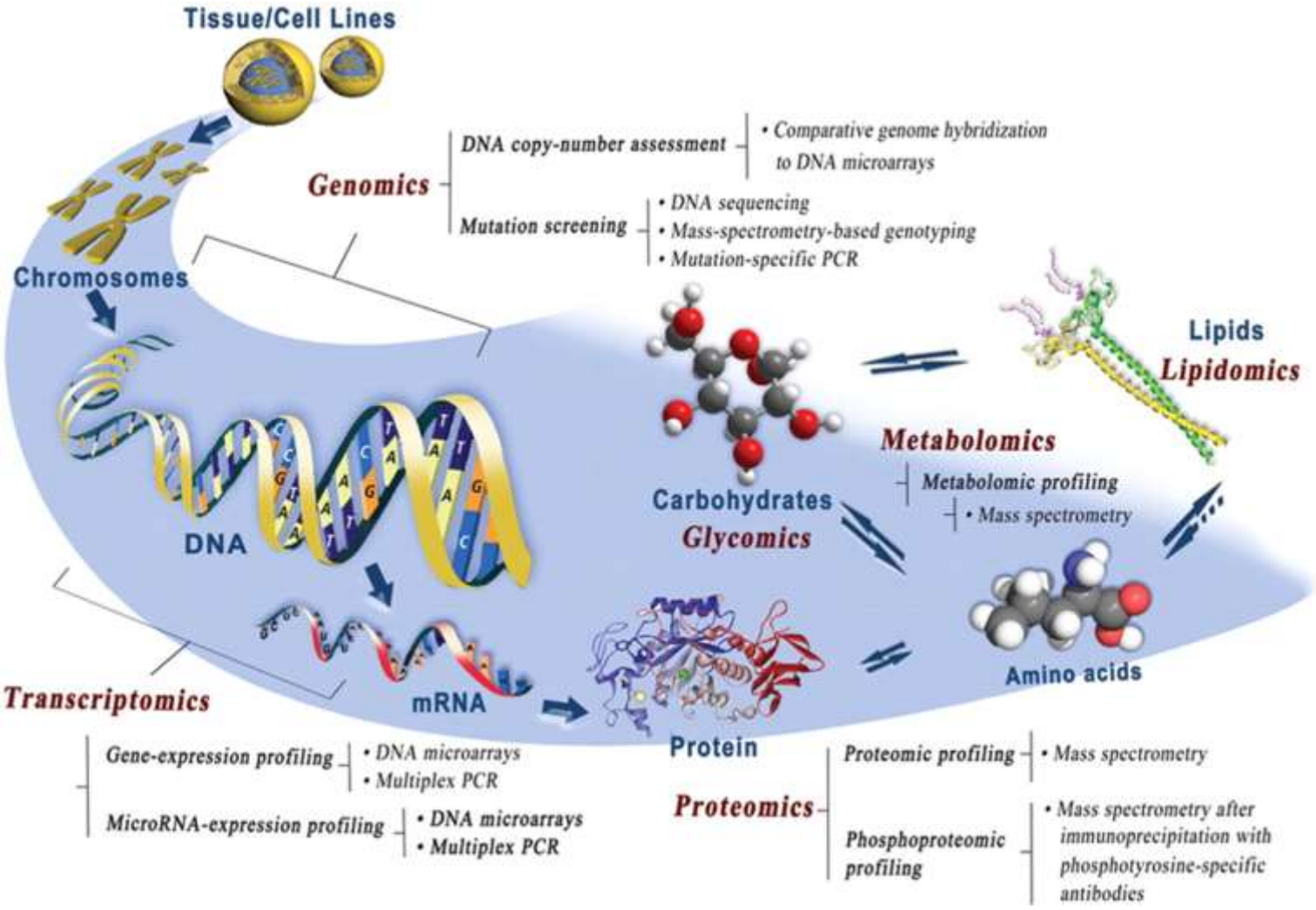
- **DATOS**

Datos sobre el tiempo





Datos sobre Biología: Genómica, Transcriptómica, Epigenómica, Proteómica, Metabolómica, Lipidómica, ….



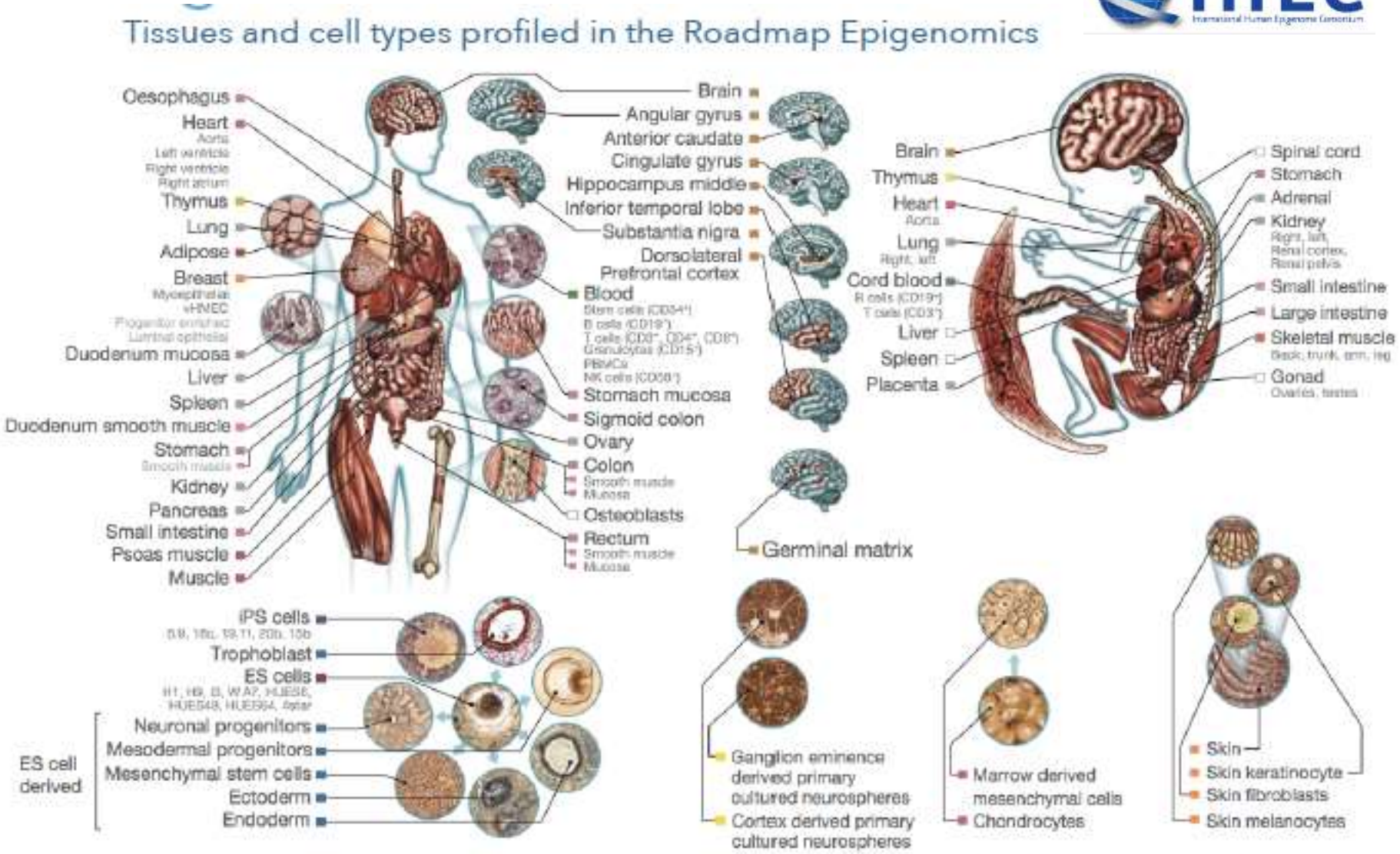
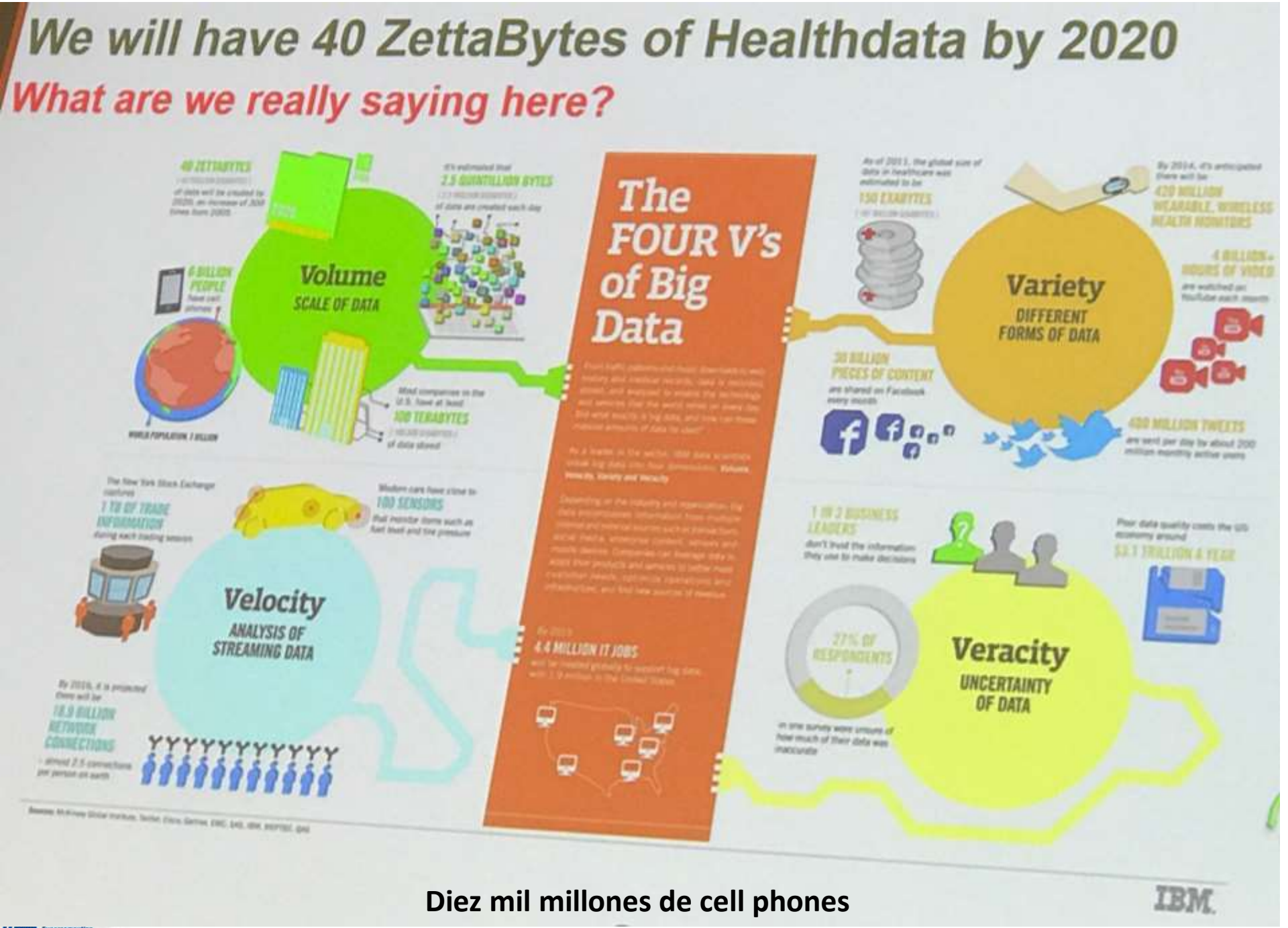


Figure 1 | Tissues and cell types profiled in the Roadmap Epigenomics Consortium. Primary tissues and cell types representative of all major lineages immune lineages, ES cells and iPS cells, and differentiated lineages derived fro ES cells. Box colours match groups shown in Fig. 2b. Epigenome identifiers

We will have 40 ZettaBytes of Healthdata by 2020

What are we really saying here?



Diez mil millones de cell phones

LifeTime

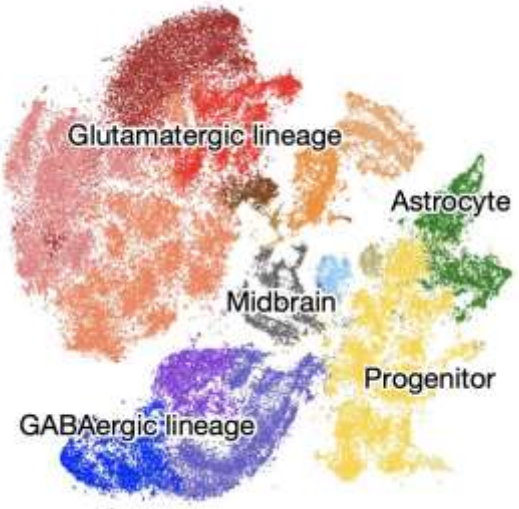
HUMAN CELL ATLAS

Single-cell multi-omics

Machine learning

Personalized organoid disease models

About 500 scientists, clinicians, patients, and stakeholders from 20 European countries and beyond have gathered in Berlin at the LifeTime Opening Conference. The pan-European initiative aims to revolutionize healthcare. It applies breakthrough technologies to the progression of human diseases and intends to find and implement new methods for personalized prevention, early diagnosis and treatment.



nature REVIEWS CANCER

Review Article | Published: 27 August 2019

Unravelling tumour heterogeneity by single-cell profiling of circulating tumour cells

Laura Keller & Klaus Pantel

Nature Reviews Cancer (2019) | Download Citation

ARTICLE

NATURE | www.nature.com/nature

https://doi.org/10.1038/s41586-019-1434-6

Resolving medulloblastoma cellular architecture by single-cell genomics

Received: 13 September 2018; Accepted: 21 June 2019;
Published online: 24 July 2019

Barcelona Supercomputing Center
Centro Nacional de Supercomputación

ELIXIR:
European infrastructure for biological information

EUROPEAN GENOME-PHENOME ARCHIVE

excelerate

Spanish National Bioinformatics Institute (INB)
Spanish Node of ELIXIR (ELIXIR-ES)

TransBioNet
network of Bioinformaticians in Research Institutes of Spanish Hospitals
(INB hosted)

elixir SPAIN

excelerate

- **DATOS**
- **INTELIGENCIA ARTIFICIAL**

Screenshot

High-performance medicine: the convergence of human and artificial intelligence

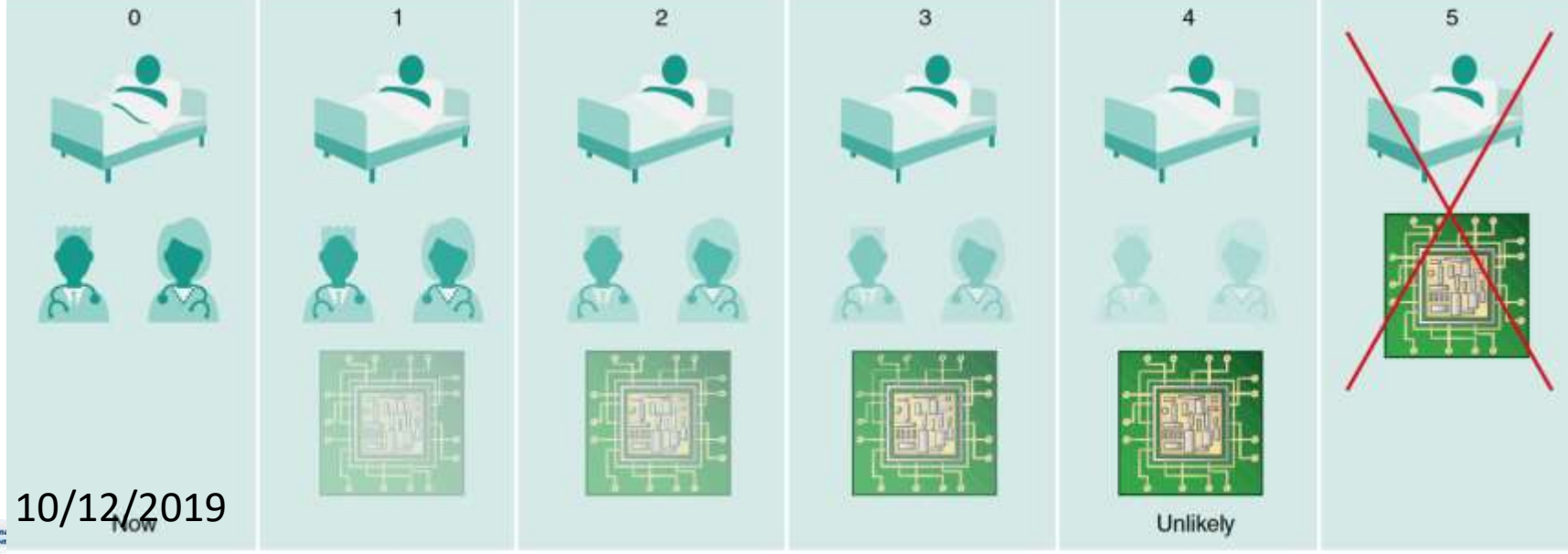
Eric J. Topol

NATURE MEDICINE | VOL 25 | JANUARY 2019 | 44-56 |

nature medicine

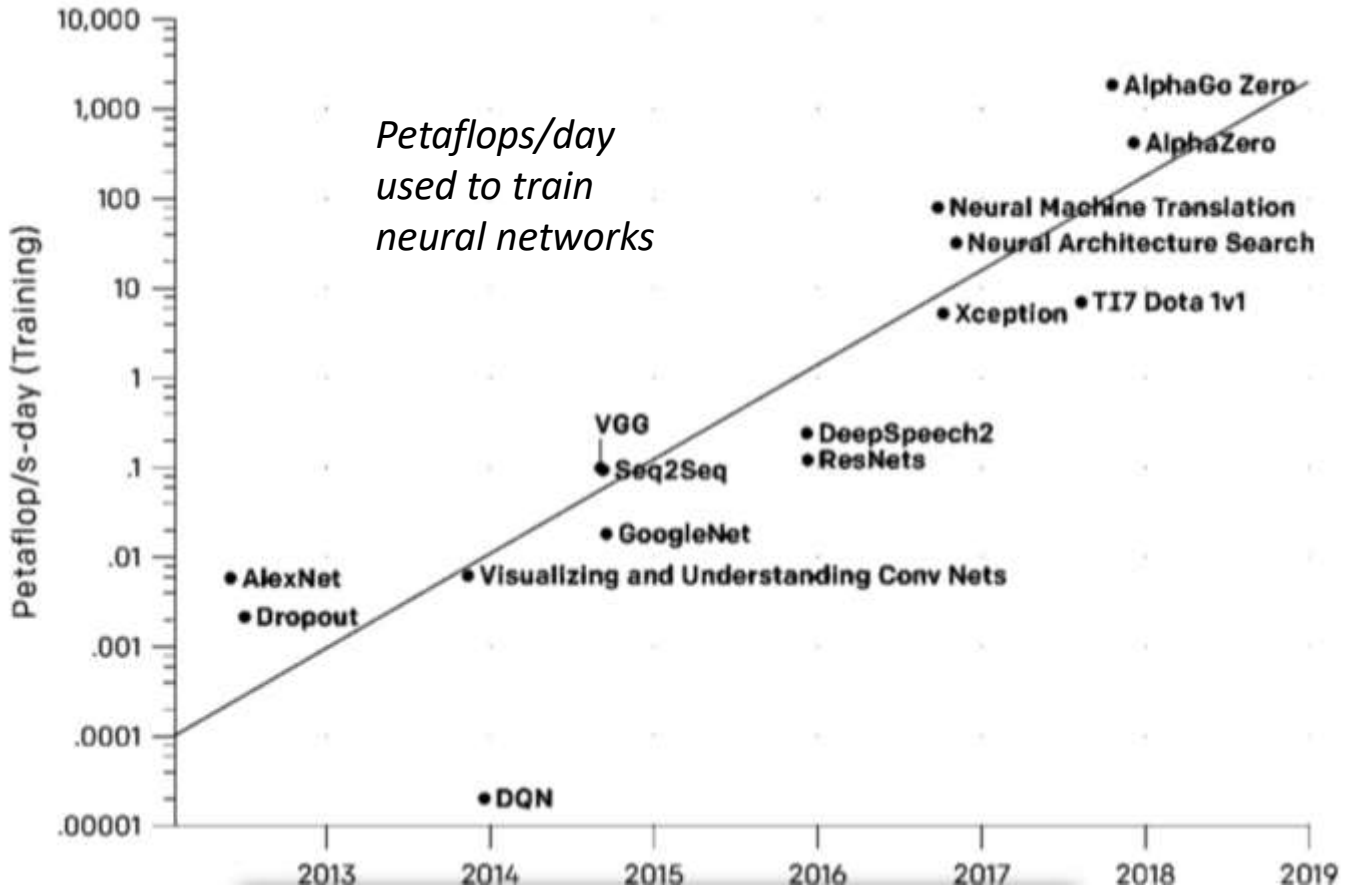
Human driver monitors environment			System monitors environment		
0	1	2	3	4	5
No automation	Driver assistance	Partial automation	Conditional automation	High automation	Full automation
The absence of any assistive features such as adaptive cruise control.	Systems that help drivers maintain speed or stay in lane but leave the driver in control.	The combination of automatic speed and steering control—for example, cruise control and lane keeping.	Automated systems that drive and monitor the environment but rely on a human driver for backup.	Automated systems that do everything—no human backup required—but only in limited circumstances.	The true electronic chauffeur: retains full vehicle control, needs no human backup, and drives in all conditions.

Humans and machine doctors



Advances in AI and HPC go hand by hand

Since GPUs were first used in AI (2012), **computing power** available to generate AI models has increased exponentially – and improvements in computing power has been key for **AI progress**.

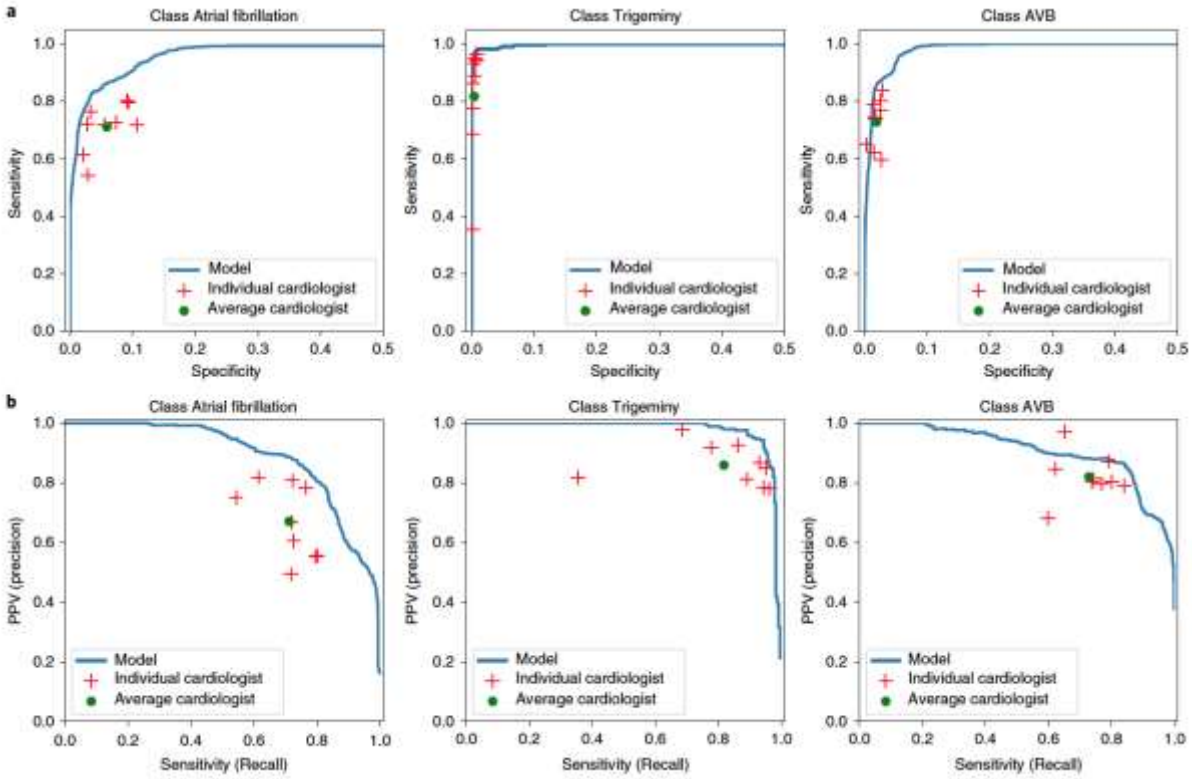


Cardiologist-level arrhythmia detection and classification in ambulatory electrocardiograms using a deep neural network

Awni Y. Hannun[1,6*], Pranav Rajpurkar[1,6], Masoumeh Haghpanahi[2,6], Geoffrey H. Tison[3,6], Codie Bourn[2], Mintu P. Turakhia[4,5] and Andrew Y. Ng[1]

nature medicine

NATURE MEDICINE | VOL 25 | JANUARY 2019 | 65–69 |



Letter | Published: 25 January 2017

Dermatologist-level classification of skin cancer with deep neural networks

Andre Esteva, Brett Kuprel, Roberto A. Novoa, Justin Ko, Susan M. Swetter, Helen M. Blau & Sebastian Thrun

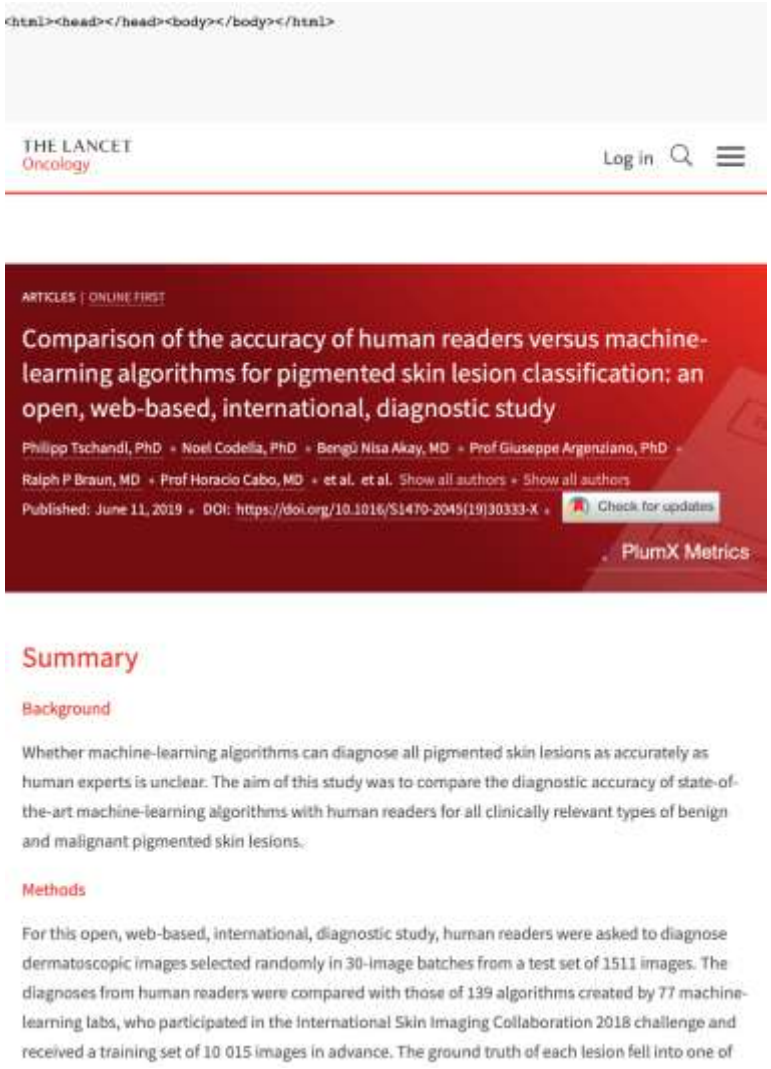
Nature **volume 542**, pages 115–118 (02 February 2017) | Download Citation

A Corrigendum to this article was published on 28 June 2017

AbstractAbstract

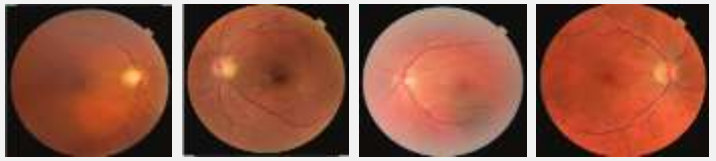
Skin cancer, the most common human malignancy[1,2,3], is primarily diagnosed visually, beginning with an initial clinical screening and followed potentially by dermoscopic analysis, a biopsy and histopathological examination. Automated classification of skin lesions using images is a challenging task owing to the fine-grained variability in the appearance of skin lesions. Deep convolutional neural networks (CNNs)[4,5] show potential for general and highly variable tasks across many fine-grained object categories[6,7,8,9,10,11]. Here we demonstrate classification of skin lesions using a single CNN, trained end-to-end from images directly, using only pixels and disease labels as inputs. We train a CNN using a dataset of 129,450 clinical images—two orders of magnitude larger than previous datasets[12]—consisting of 2,032 different diseases. We test its performance against 21 board-certified dermatologists on biopsy-proven clinical images with two critical binary classification use cases: keratinocyte carcinomas versus benign seborrheic keratoses; and malignant melanomas versus benign nevi. The first case represents the identification of the most common cancers, the second represents the identification of the deadliest skin cancer. The CNN achieves performance on par with all tested experts across both tasks, demonstrating an

Access options



BSC works on Medical Imaging

- Detecting retina pathologies
  - Trained models competitive with ophthalmologists
  - With Lenovo & Hospital Vall Hebron
- Learning from liver conditions
  - Learning about rare diseases
  - With Hospital Clinic
- Predicting and guiding in-vitro success
  - Finding the best embryo ASAP
  - With Hospital Clinic
- Supporting medical doctors on Rx review
  - Aid for Dr. in rural areas
  - With Asepeyo and ICS

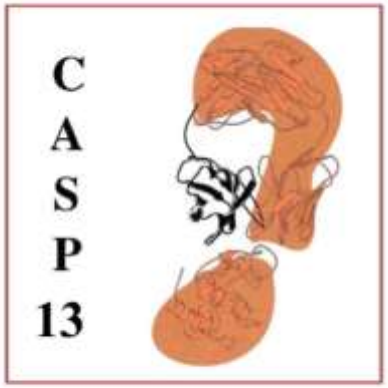


Glaucoma | Epiretinal Membrane | Nevus | Macular Degeneration

Barcelona Supercomputing Center
Centro Nacional de Supercomputación

By Ulises Cortes and Dario Garcia UPF & BSC

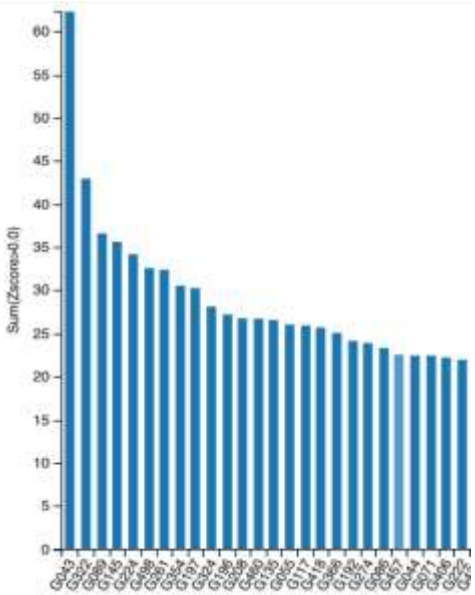
CRITICAL ASSESSMENT OF TECHNIQUES
FOR PROTEIN STRUCTURE PREDICTION



Thirteenth meeting
Riviera Maya, Mexico
DECEMBER 1-4, 2018

Predicting protein structure from the sequence is one of the fundamental problems in molecular biology.

It is the key to the prediction of the consequences of mutations in human diseases and to drug design



Nico Callewaert @NicoCallewaert · 11h
Probably my nomination for basic molecular science advance of 2018, need to see a bit more methods details but results in blinded **CASP13** test clearly impressive. deepmind.com/blog/alphafold/

On its first foray into the competition, AlphaFold topped a table of 98 entrants, predicting the most accurate structure for 25 out of 43 proteins, compared with three out of 43 for the second placed team in the same category.

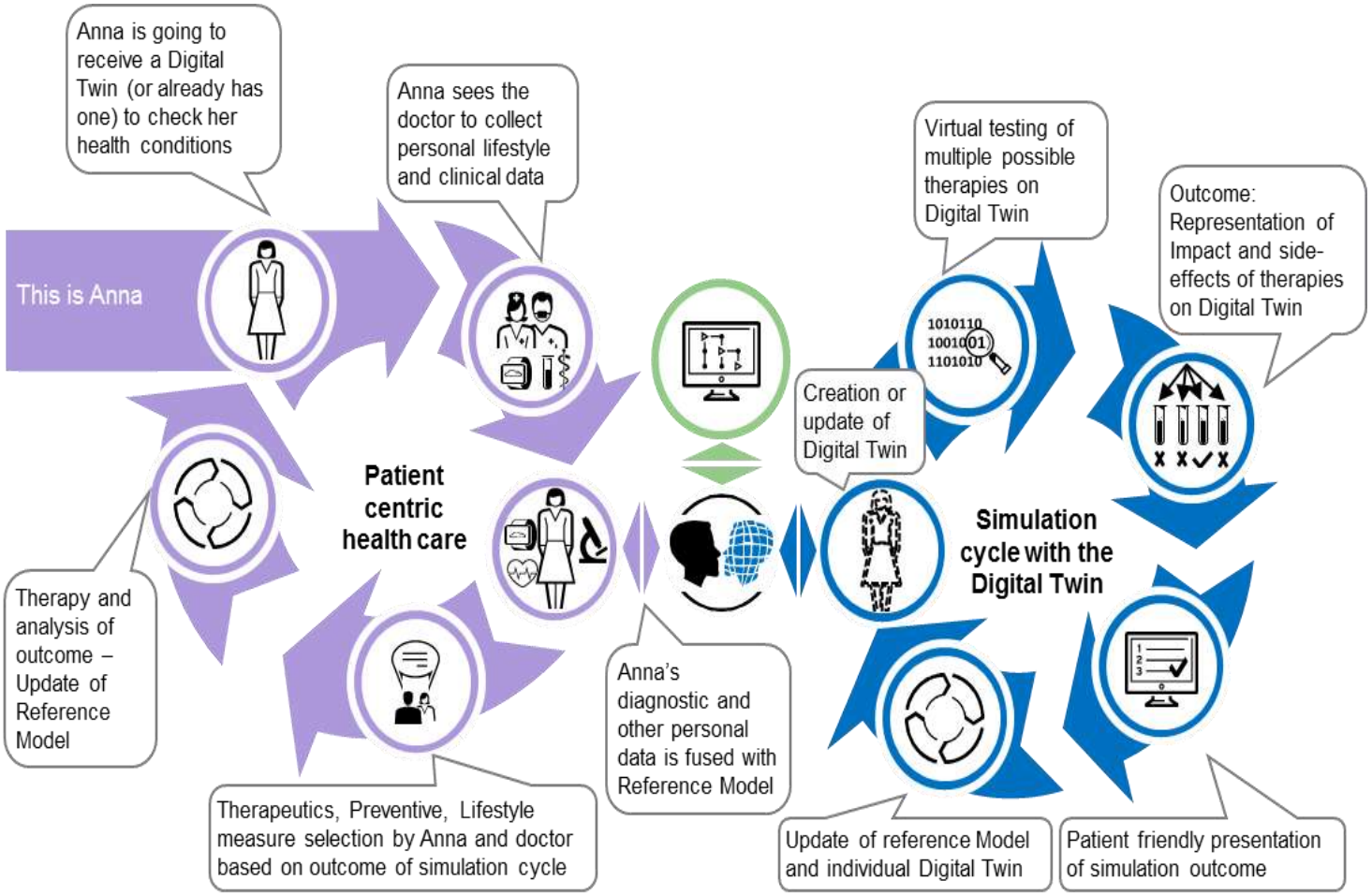
- **DATOS**
- **INTELIGENCIA ARTIFICIAL**
- **HPC**

The Evolution of the Research Paradigm

Numerical Simulation and Big Data Analysis

- Reduce expense
- Avoid suffering
- Help to build knowledge where experiments are impossible or not affordable

Digital Twin for Future Medicine



Is this scenario possible?
When?

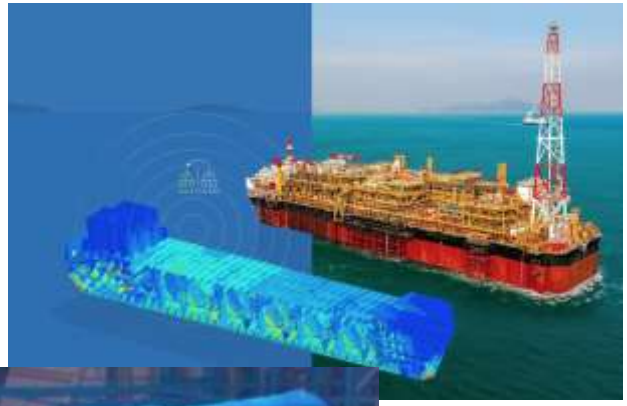
HORIZON

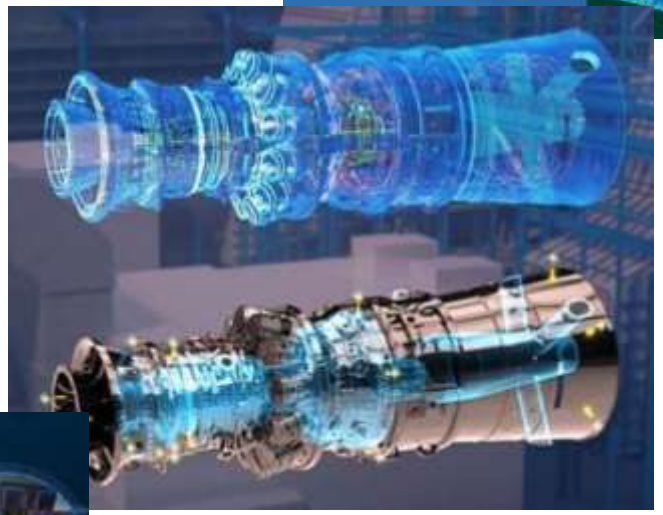
The EU Research & Innovation Magazine

INDUSTRY SCIENCE IN SOCIETY ICT

How digital 'twins' are guiding the future of maintenance and manufacturing

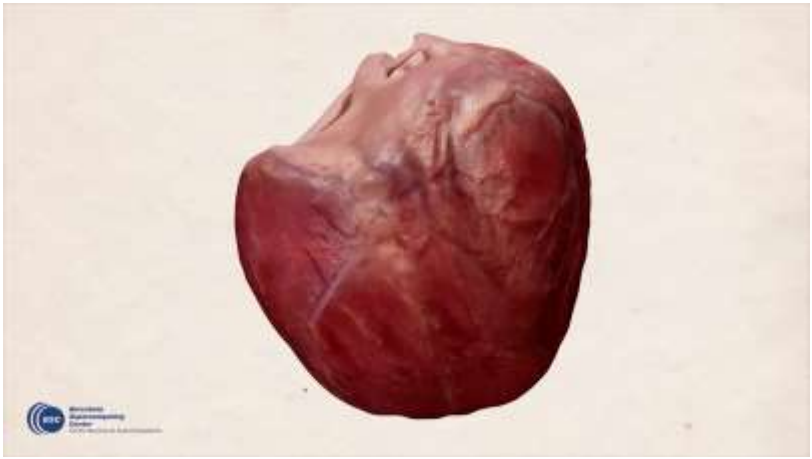
15 November 2019



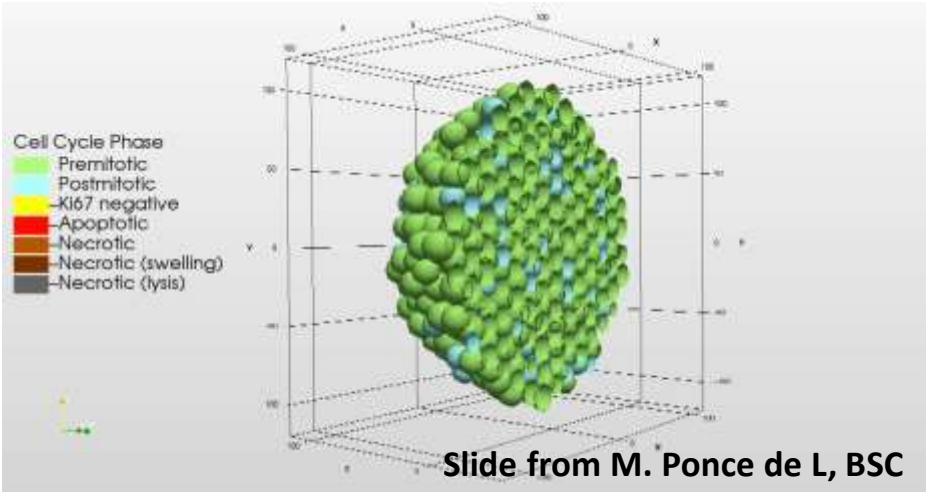


GE Healthcare is building digital twins of hospitals, mining data to improve performance at the individual asset and entire network level

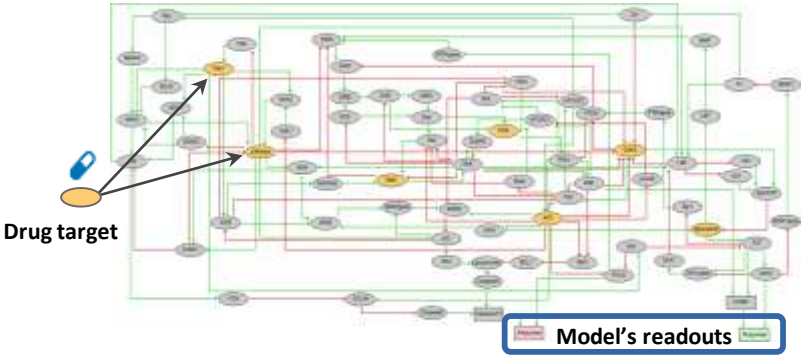
Simulations of biological systems at different levels



By Mariano Vazquez, CASE - BSC

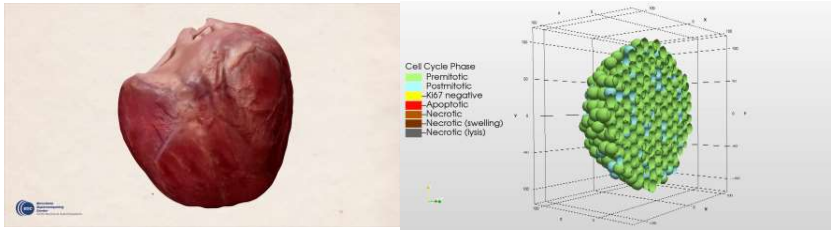


~48 h simulation time, 30 min wall time
~2500 cells



By Victor Guallar ICREA & BSC

Large Scale Simulations

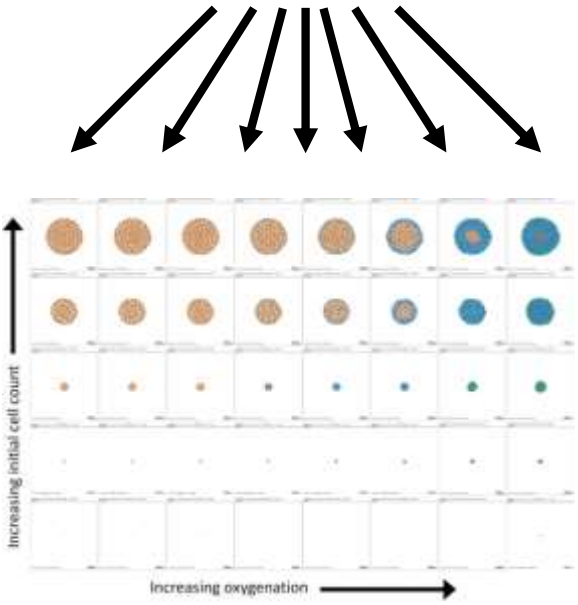


AI / ML systems

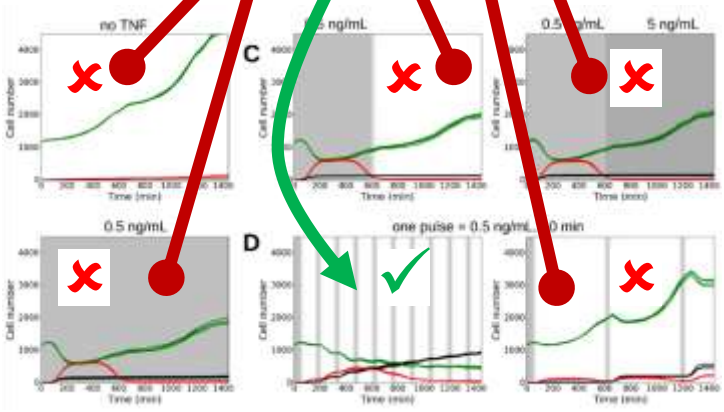
Exploration of the parameter space

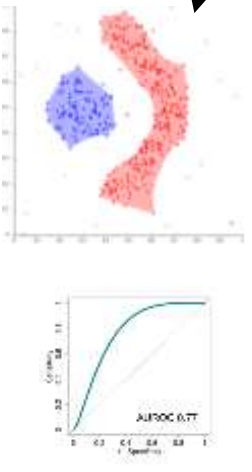
Monitoring and tailoring simulations during execution time

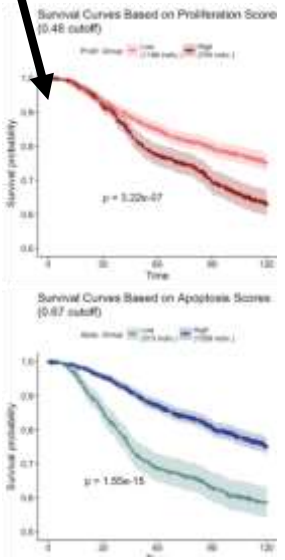
Analysing the results of the simulations



Event Recognition System

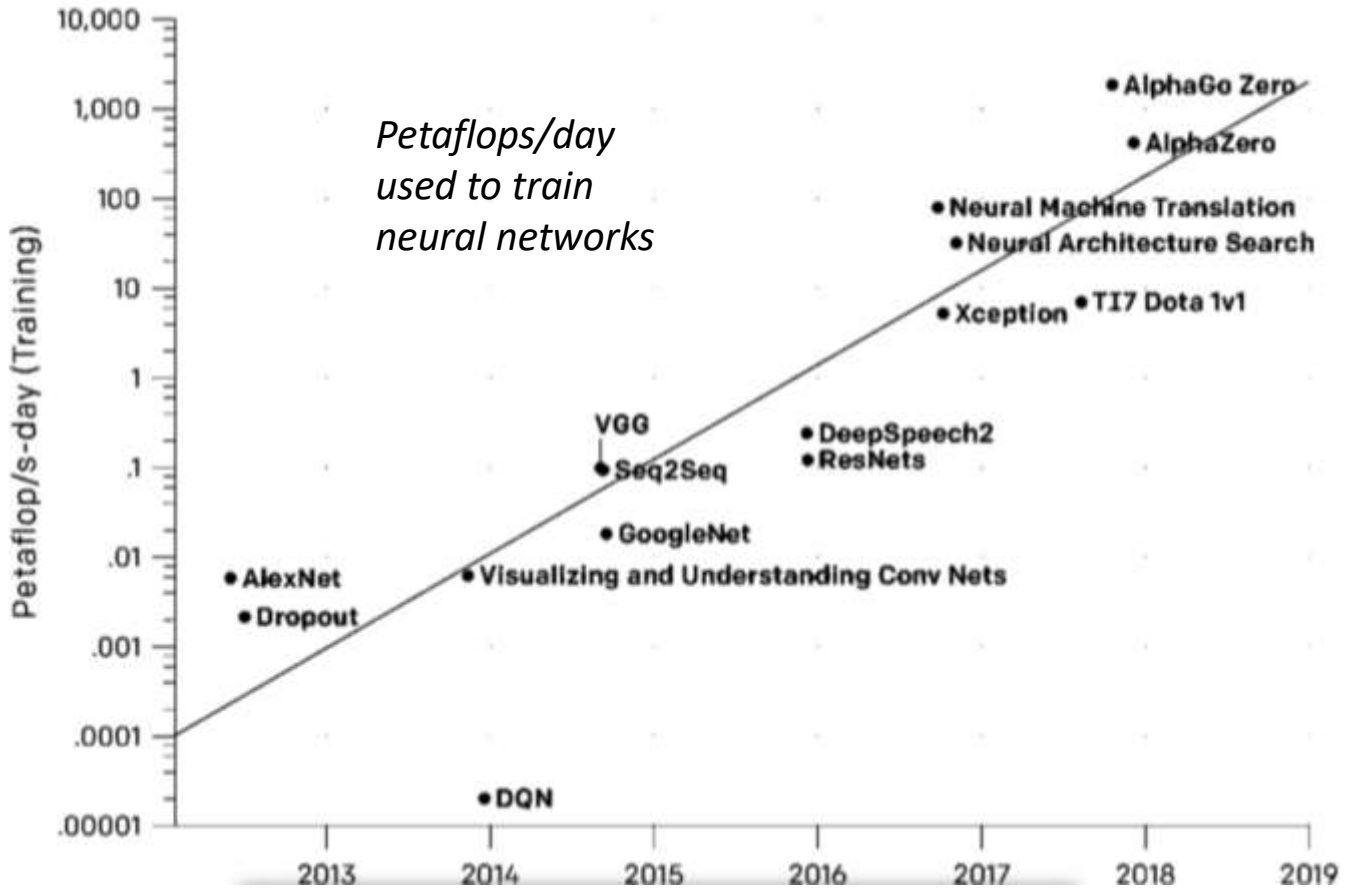






Advances in AI and HPC go hand by hand

Since GPUs were first used in AI (2012), **computing power** available to generate AI models has increased exponentially – and improvements in computing power has been key for **AI progress**.



Barcelona Supercomputing Center
Centro Nacional de Supercomputación

From Jordi Torres

Text mining & Cognitive computing for melanoma research

Barcelona Supercomputing Center
Centro Nacional de Supercomputación

ecialized info

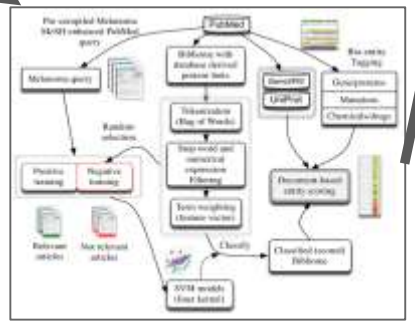
Watson General info

IBM Watson

PubMed

16M abstra

BSC Text mining



Bio-Knowledge corpus

Healthcare dictionaries

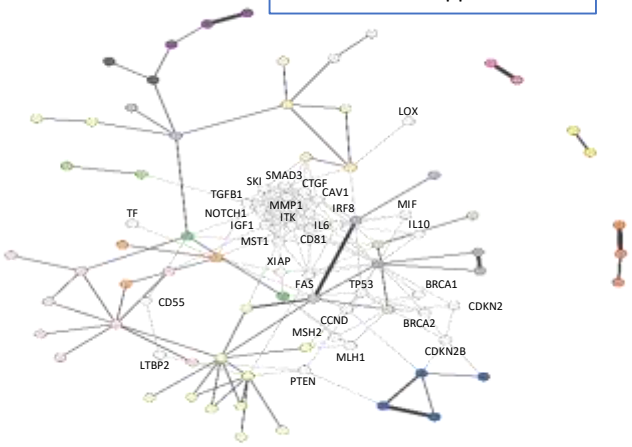
IBM Text Analytics Catalog

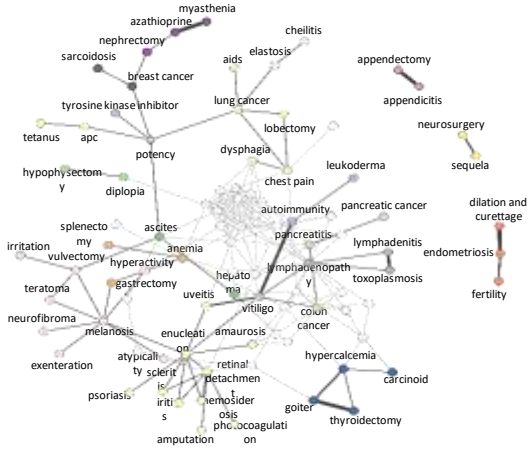
WATSON Cognitive computing

Content Analytics graph

Corpus/entities extraction
Network of entities
Inference of relationships

Bio-entities associations
Functional interpretation
Translational applications





Barcelona Supercomputing Center
Centro Nacional de Supercomputación

In collaboration with IBM Spain

Word embeddings

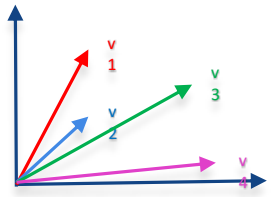
Word2Vec

fastText

|V|

v1	0,78	0,65	0,98
v2	0,23	0,12	0,32
v3	0,90	0,32	0,56
v4	0,08	0,43	0,65
v5	0,77	0,88	0,77

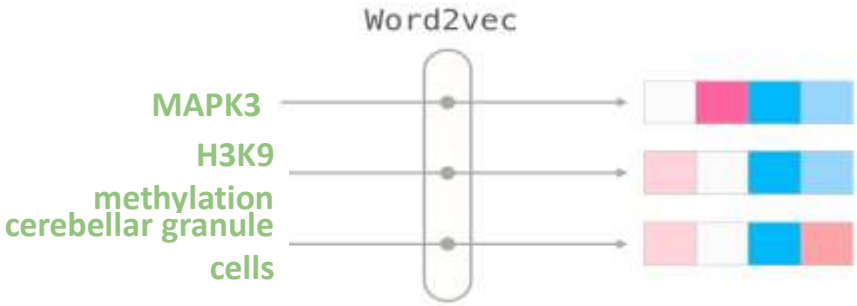
Word vectors



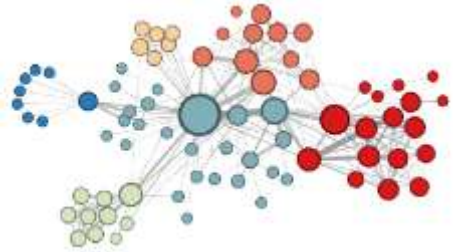
Evaluación intrínsica: cálculo similitud entre términos (sinónimos en SNOMED)

Evaluación extrínsica: comprobar su utilidad en otras tareas PLN (neuroNER)

Plan TL
Plan de Impulso de las Tecnologías del Lenguaje



Word embeddings



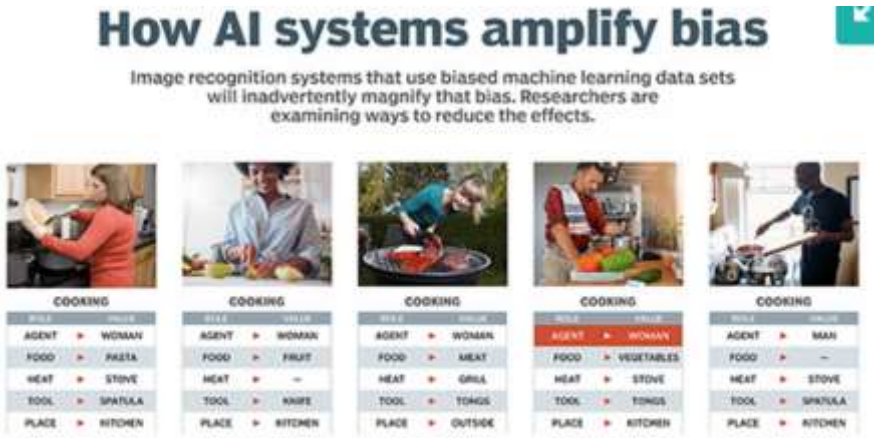
individualized Paediatric Cure (iPC)
Cloud-based virtual-patient models
for precision paediatric oncology

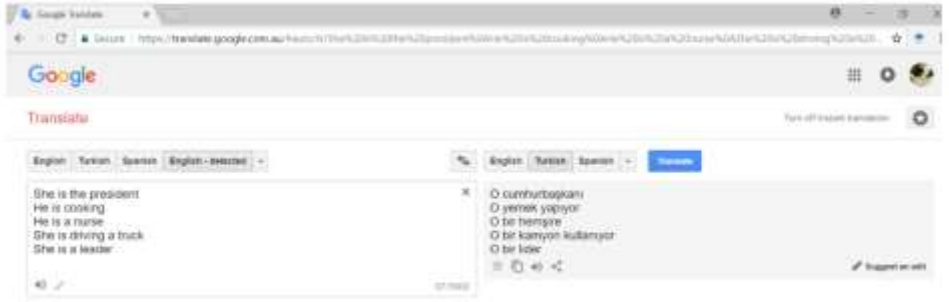
Barcelona Supercomputing Center
Centro Nacional de Supercomputación

European Commission
Horizon 2020
European Union funding
for Research & Innovation

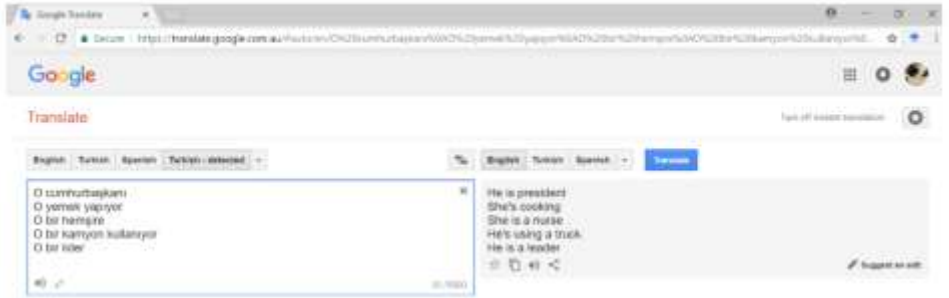
IBM

Gender and other biases …





Google image search



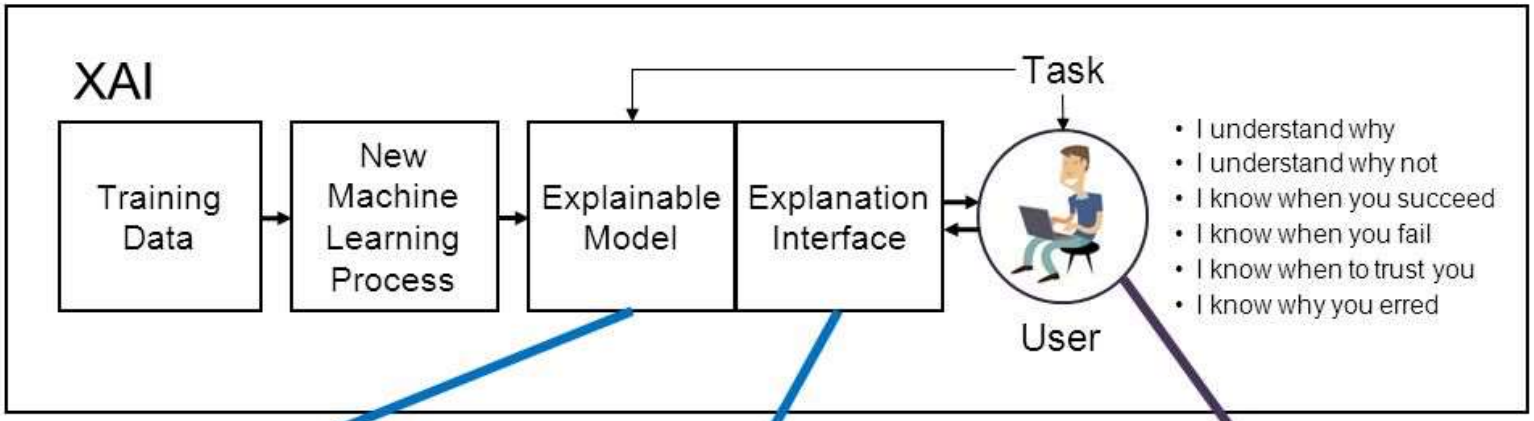
- Research activities in five key interconnected AI scientific areas (Explainable AI, Physical AI ,Verifiable AI, Collaborative AI, Integrative AI), which arise from the application of AI in real-world scenarios;
- The creation of a European Ethical Observatory to ensure that European AI projects adhere to high ethical, legal, and socio-economical standards;

AI4EU ETHICAL OBSERVATORY

The main goal of the AI4EU Observatory on Society and Artificial Intelligence (OSAI) is to support the distribution and the discussion of knowledge about the Ethical, Legal, Socio-Economic and Cultural issues of AI (ELSEC-AI) within Europe.

Explainable Artificial Inteligence

B. Program Scope – XAI Development Challenges



Explainable Models

- develop a range of new or modified machine learning techniques to produce more explainable models

Explanation Interface

- integrate state-of-the-art HCI with new principles, strategies, and techniques to generate effective explanations

Psychology of Explanation

- summarize, extend, and apply current psychological theories of explanation to develop a computational theory

EL PAÍS

BERNARDO DE MIGUEL | LLUÍS PELLICER

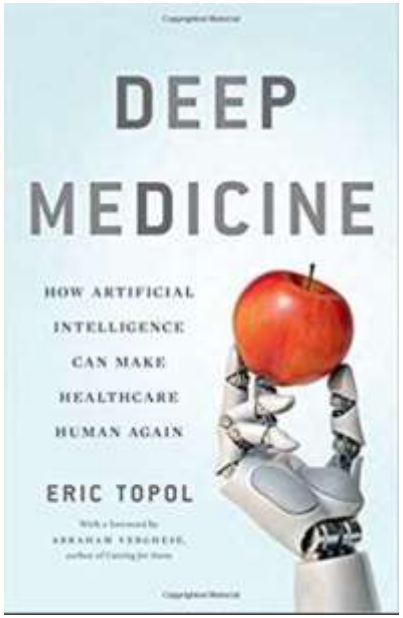
Bruselas · 17 MAR 2019 - 20:23 CET

El nuevo miedo de Europa se llama China

Mientras el Gobierno de Merkel ha decidido invertir 3.000 millones de euros en inteligencia artificial hasta 2025, China invertirá 130.000 millones hasta 2030", señala Müller-Markus. La reducción de

The United Kingdom is also betting big on AI's future and emphasizing healthcare. When the UK government issued four Grand Challenges, one centered on medicine, Theresa May declared, "The development of smart technologies to analyse great quantities of data quickly and with a higher degree of accuracy than is possible by human beings, opens up a whole new field of medical research and gives us a new weapon in our armory in the fight against disease."[66] In 2018, I was commissioned by the UK to work with the National Health Service on planning the future of its healthcare, particularly leading a review on the impact of AI and other medical technologies on its workforce over the next two decades.[67] The opportunity to

work with leaders of AI, digital medicine, genomics, and robotics, along with ethicists, economists, and educators was an extraordinary experience in the context of a single-payer healthcare system with a palpable will to change and adapt. The full report was issued in 2019, where we project major impacts at every level—the patient, the clinicians, and the health systems throughout the country.



10/12/2019