



**Barcelona
Supercomputing
Center**

Centro Nacional de Supercomputación

Life Sciences Department

Barcelona Supercomputing Center

By Alfonso Valencia. Ph.D.

ICREA Prof.

Director Life Sciences Dept BSC

Director INB-ISCIH / ELIXIR-ES

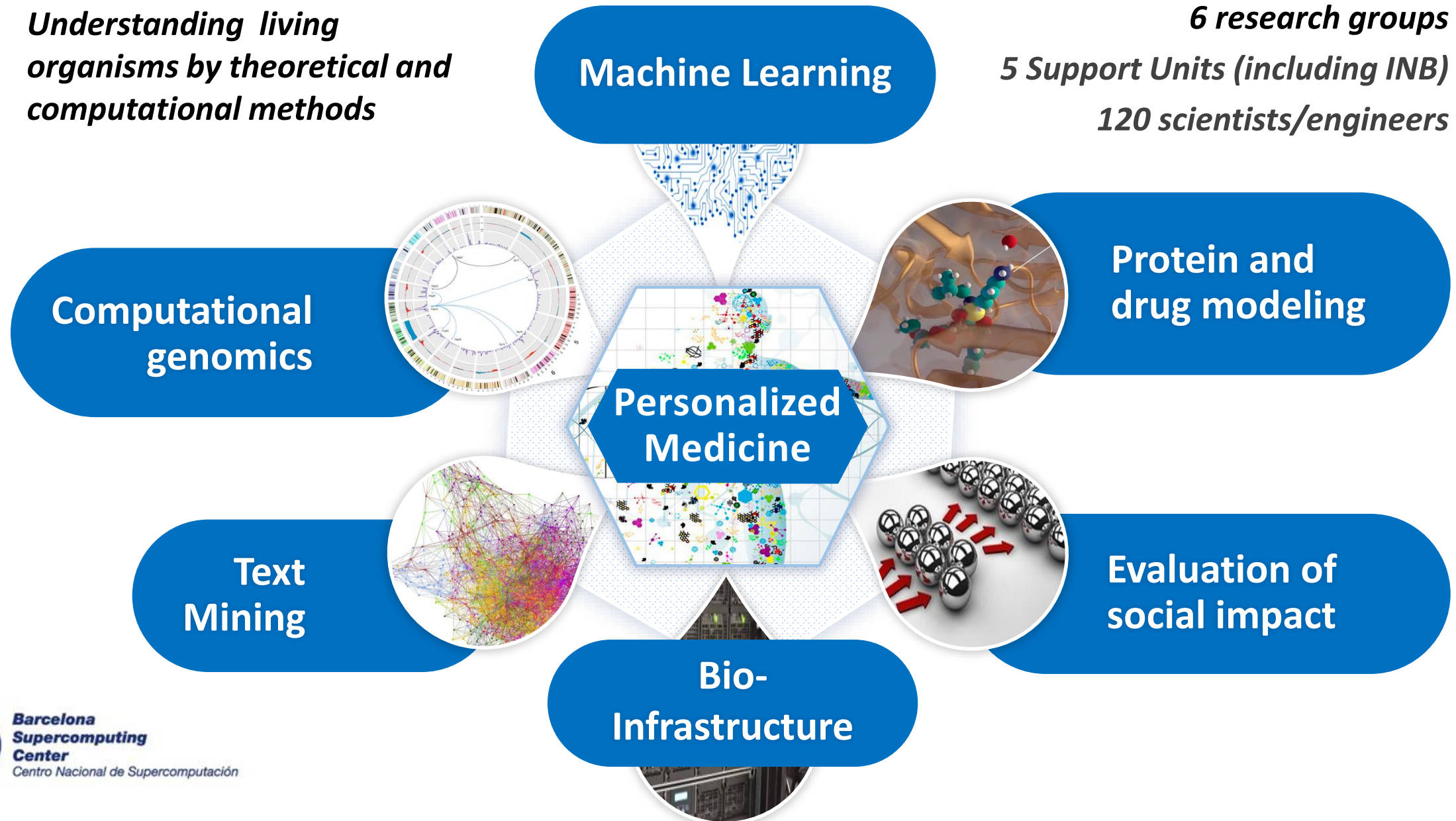
Life Sciences objectives and mission

Understanding living organisms by theoretical and computational methods

6 research groups

5 Support Units (including INB)

120 scientists/engineers



Genomic
Regulation
JRG



Computational
Genomics



Coordination
Node



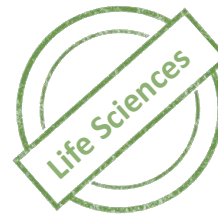
Microbial
Genomics



Genome
Informatics



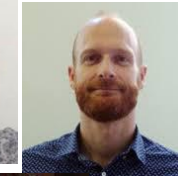
Life Sciences 2020



Computational
Bioinformatics



Text
Mining



Social
Impact



Protein
Modeling



Computational
Biology



Biological
Networks



Associated
Groups



Barcelona
Supercomputing
Center
Centro Nacional de Supercomputación

Milana
Morgenstern
Bar-Ilan

Vera
Pancaldi
Toulouse

Joaquín
Dopazo
Seville

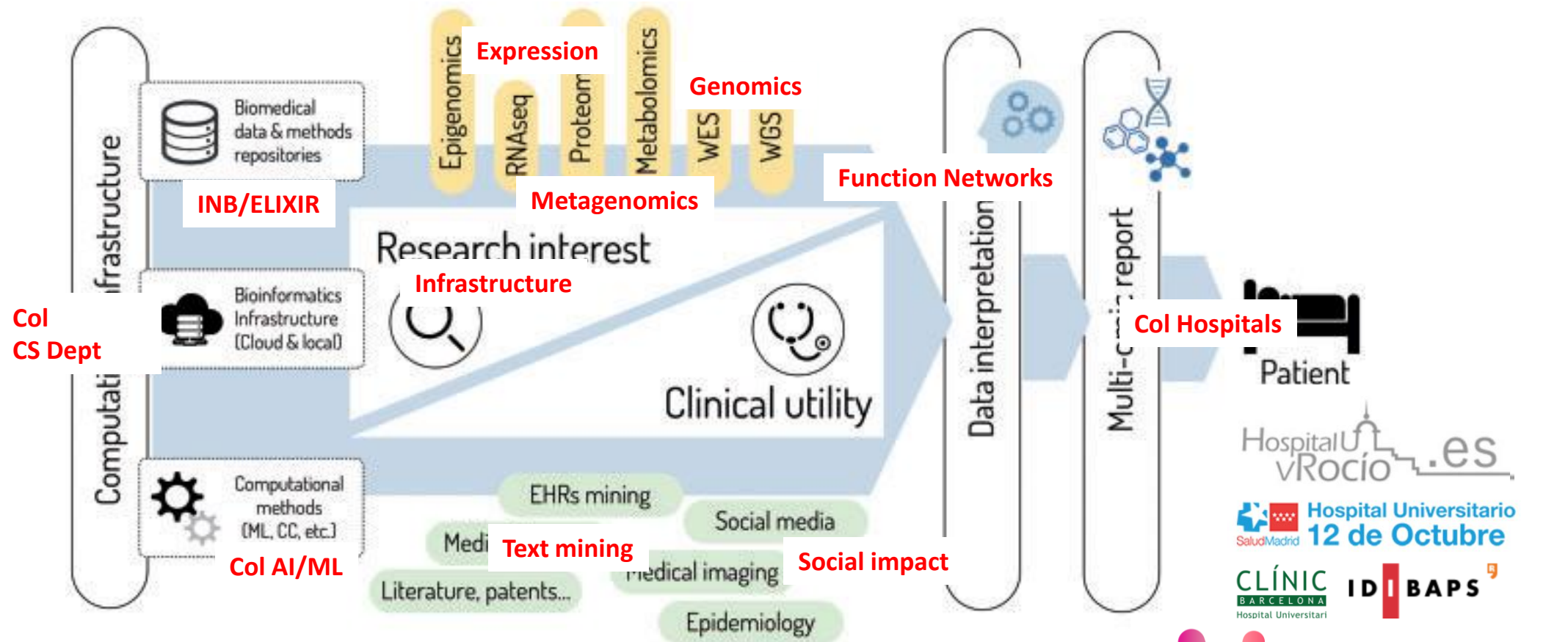
Biola
Javierre
F Carreras

Anïs
Baudot
Marseille

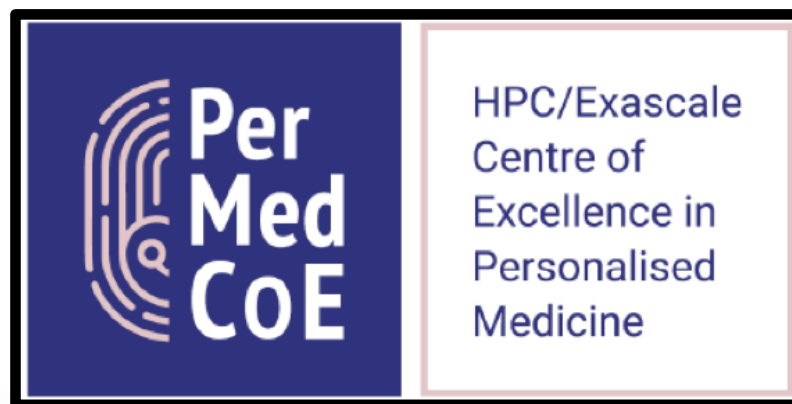
Osnat
Hakimi
UIC

Eudard
Porta
F. Carreras

Impact of the Life Sciences Dept in Personalise Medicine



The Dept leads the new HPC Center of Excellence for cell level simulations



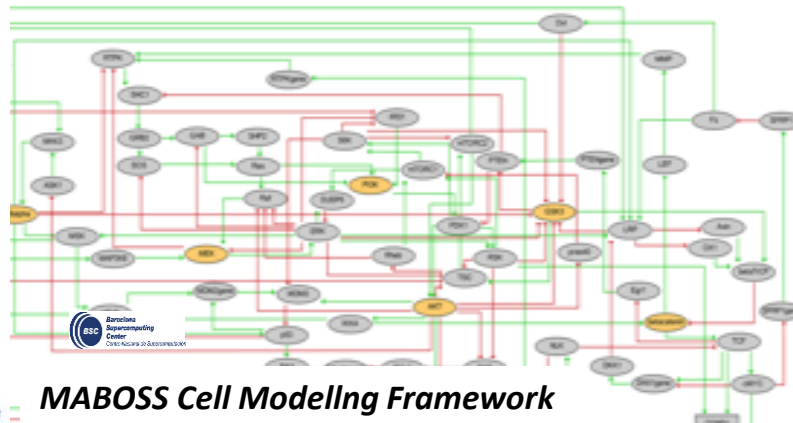
Simulations of intra and intercellular processes to fill the gap between molecular and organismic simulations



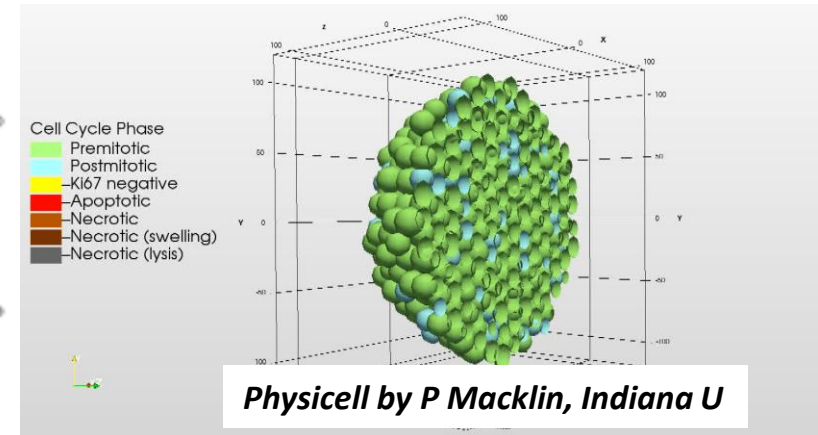
By Victor Guallar ICREA & BSC



By Mariano Vazquez, CASE - BSC

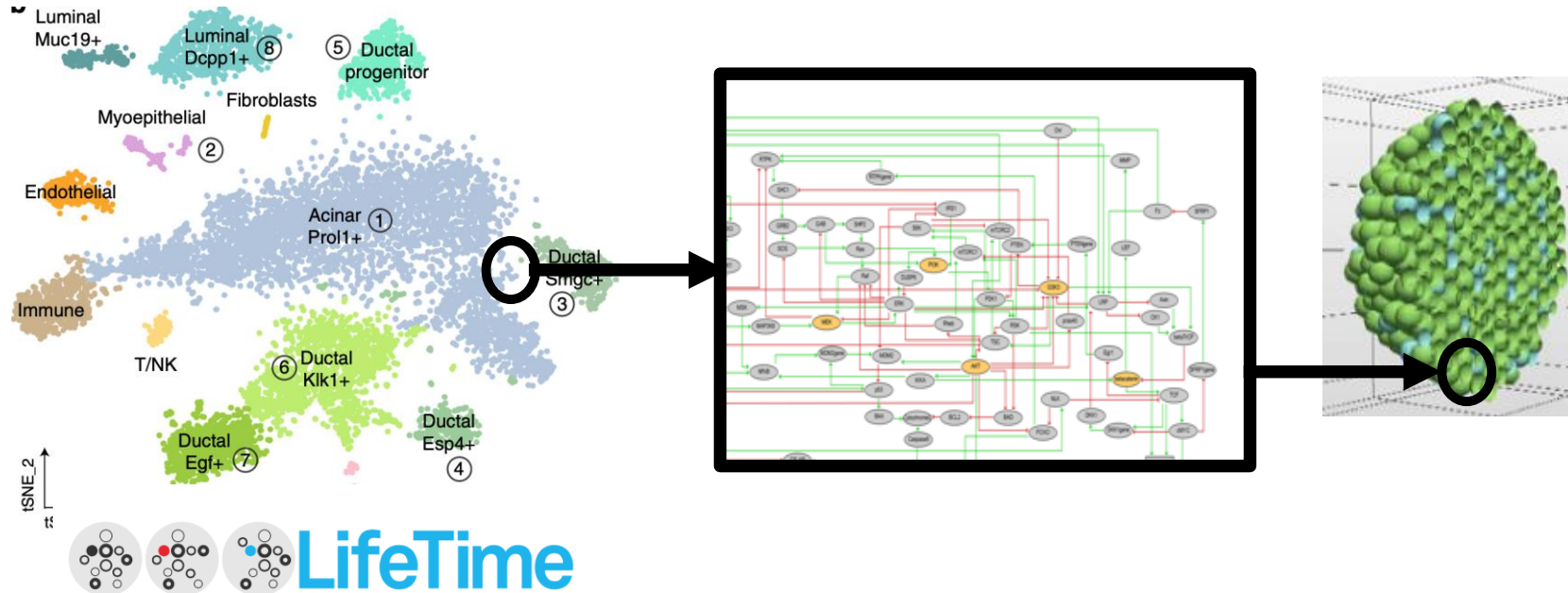


MABOSS Cell Modelling Framework



Physicell by P Macklin, Indiana U

From Single Cell Data to Models of Tumors

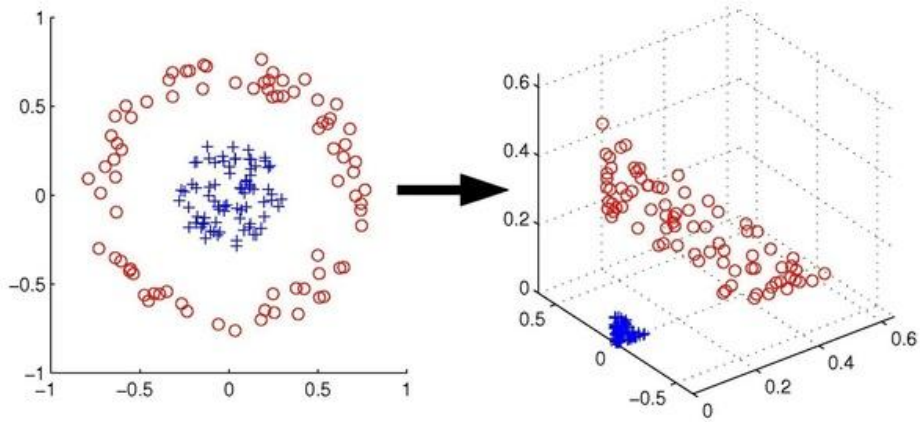



End Users Communities

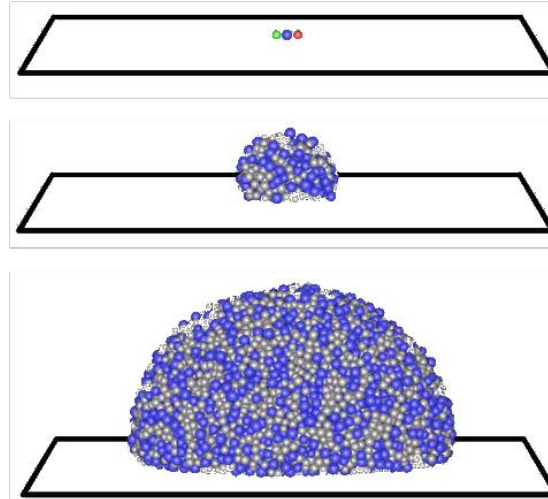


GOAL : “Replace the current generation of bioinformatics methods with **cellular models, providing mechanistic descriptions and testable hypotheses**, instead of current statistical approximations and intuitive descriptions”

Life Sciences in Simulations AI and NLP



Machine Learning



Simulations

Barack Hussein Obama II * (born August 4, 1961 *) is an American * attorney and politician who served as the 44th President of the United States * from January 20, 2009 *, to January 20, 2017 *. A member of the Democratic Party *, he was the first African American * to serve as president. He was previously a United States Senator * from Illinois * and a member of the Illinois State Senate *.

Natural Language Processing

Minimal data sets for COVID-19 modelling

	A	B	C	D	E	F	G	H	I	J
1			Visor COVID-19 H120							
2			Observaciones							
3										
4		Fecha-hora:	3/29/2020 12:36							
5		+ Enlaces:	Recursos de referencia [+]							
14										
15										
16										
17	Arbol	Componente	Descripción	Estandarización [+]						
18	c1	c4	c5	Enlace terminológico	Tipo de enlace	Cardinalidad	Estándar	Contexto	Código	Término
19	-	COMPOSITION		Componentes	SNOMED CT	artefacto de registro	-1			
20	-	SECTION		Componentes	SNOMED CT	artefacto de registro	-1			
21	-	ENTRY		Componentes	SNOMED CT	artefacto de registro	-1			
22	+ Cluster		Observaciones epidemiológicas en relación con COVID-19	Componentes	SNOMED CT	artefacto de registro	63261000122108	Observaciones epidemiológicas en relación con COVID-19		
49	+ Cluster		Observaciones sobre síntomas en relación con COVID-19	Componentes	SNOMED CT	artefacto de registro	63271000122103	Observaciones sobre síntomas en relación con COVID-19		
74	+ Cluster		Observaciones clínicas en relación con COVID-19	Componentes	SNOMED CT	artefacto de registro	63281000122100	Observaciones clínicas en relación con COVID-19		
127	+ Cluster		Observaciones de laboratorio en relación con COVID-19	Componentes	SNOMED CT	artefacto de registro	63291000122102	Observaciones de laboratorio en relación con COVID-19		

Observaciones epidemiológicas en relación con COVID-19
Observaciones sobre síntomas en relación con COVID-19
Observaciones clínicas en relación con COVID-19
Observaciones de laboratorio en relación con COVID-19

Natural Language Processing to extract COVID-19 variables from EHRs

Pre-
processing

Term
extraction

Doc
similarity

Annotation &
guidelines

Model
training

Automatic
labelling

Output



0

1

2

3

4

5



Input: emergency
discharge reports,
radiology reports (Chest
X/CT-scans (H12O,
Clínic)

Document
selection for
annotation

Expert knowledge
(human in the loop)
adaptation/fine tuning

Deep learning : procedures,
symptoms, drugs, diseases,
negation, pathogens, socio-
economic...

SPECIES **SINTOMA**
Al ingreso, la paciente reportó cuatro defecaciones al día con sangre

ENFERMEDAD **ENTIDAD-OBSERVABLE**
Las pruebas analíticas indicaron anemia grave (hemoglobina 7

PROCEDIMIENTO **ENFERMEDAD**
Una sigmoidoscopia mostró úlceras grandes y profundas en el colon:

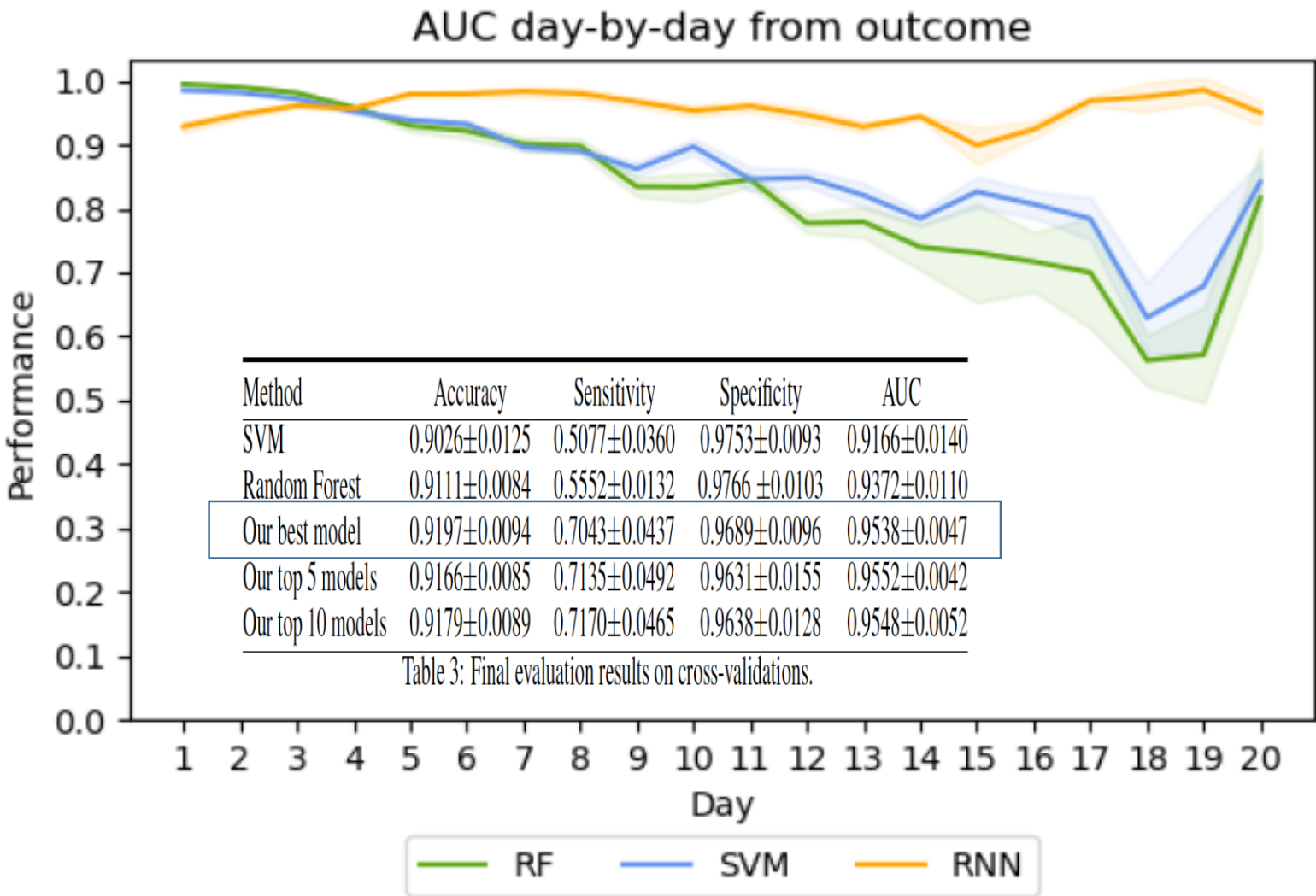
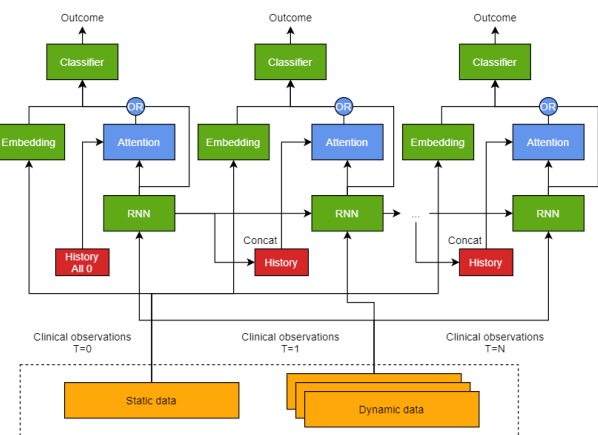
Structured
clinical data

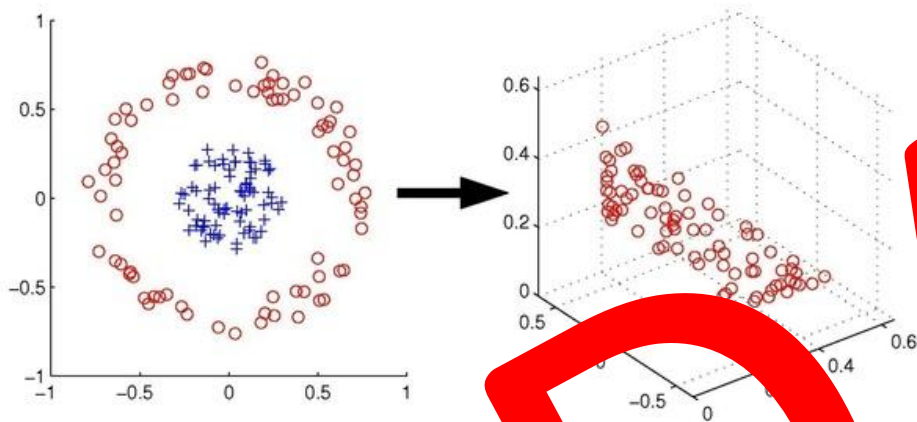
- Detection of COVID-19 cohorts
- Extraction of: other diagnoses, symptoms/ severity, and outcomes
- Early warning: predict the clinical course of the disease (increase in FIO2, admission to intensive care)

Predicting Disease Outcome with Recurrent Neural Networks and Attention mechanisms

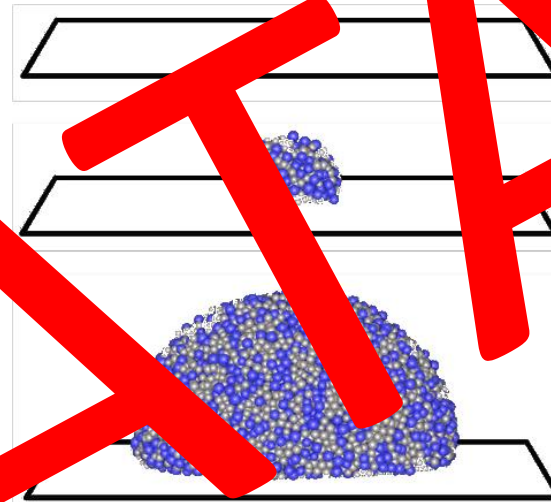


TOWARDS PREDICTING THE EVOLUTION OF COVID-19 MORTALITY RISK:
A RECURRENT NEURAL NETWORK APPROACH
CP Carrino, A Gonzalez-Agirre, J Armengol-Estapé,
A Gutiérrez-Fandiño, D Pérez Fernández, M Villegas. A Valencia





Machine Learning

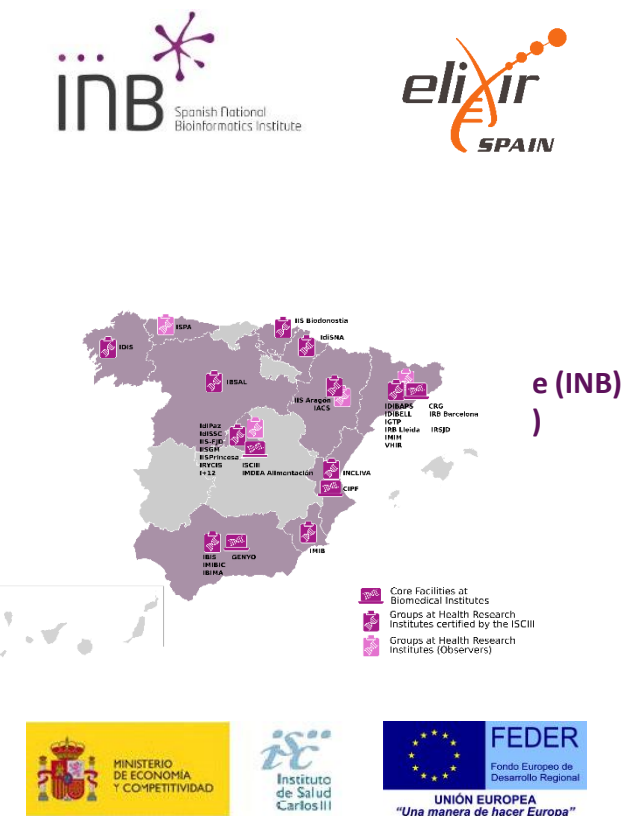


Simulations

Barack Hussein Obama II (born August 4, 1961) is an African attorney and politician who served as the 44th President of the United States from January 20, 2009, to January 20, 2017. A member of the Democratic Party, he was the first African American to serve as president. He was previously a United States Senator from Illinois and a member of the Illinois State Senate.

Natural Language Processing

Life Sciences Dept collaboration with Hospitals as part of the European Bioinformatics Infrastructure ELIIXIR



TransBioNet (mainly Institutes in Research Hospitals)

Acronym	TransBioNet member	
GENYO	Centre for Genomics and Oncological Research	H V del Rocio Sevilla
IBIMA	Instituto de Investigación Biomédica de Málaga	
IBiS	Instituto de Biomedicina de Sevilla	
IMBIC	Instituto Maimónides de Investigación Biomédica de Córdoba	
IACS	Instituto Aragonés de Ciencias de la Salud	
IIS Aragón	Instituto de Investigación Sanitaria Aragón	H Clinic Barcelona
ISPA	Instituto de Investigación Sanitaria del Principado de Asturias	
IBSAL	Instituto de Investigación Biomédica de Salamanca	
CRG	Centre for Genomic Regulation	
IDIBAPS	Institut d'Investigacions Biomèdiques August Pi i Sunyer	
IDIBELL	Instituto de Investigaciones Biomédicas de Bellvitge	H 12Octubre Madrid
IGTP		
IRB Barcelona		
IRBLleida		
IRSJD		
PSMAR-IMIM		
VHIR	Vall d'Hebron Institut de Recerca	H 12Octubre Madrid
CIPF	Centro de Investigación Príncipe Felipe	
INCLIVA	Institut d'Investigació Sanitària INCLIVA	
IDIS / FPGMX	Instituto de Investigación Sanitaria de Santiago de Compostela / Fundación Pública	
IdISBa	Instituto de Investigación Sanitaria Islas Baleares	
i+12	Instituto de Investigación Hospital 12 de Octubre	H 12Octubre Madrid
IdiPaz	Instituto de Investigación Hospital Universitario La Paz	
IDIPHIM	IIS Instituto de Investigación Sanitaria Puerta de Hierro	
IdISSC	Instituto de Investigación Sanitaria Hospital Clínico San Carlos	
IIS-FJD	IIS-Fundación Jiménez Díaz	
IISGM	Gregorio Marañón Health Research Institute	H 12Octubre Madrid
IISPrincesa	Instituto de Investigación Sanitaria Hospital de la Princesa	
IMDEA	Instituto Madrileño de Estudios Avanzados	
IRYCIS	Instituto Ramón y Cajal de Investigación Sanitaria	
ISCIII	Instituto de Salud Carlos III	
IMIB	Instituto Murciano de Investigación Biosanitaria	H 12Octubre Madrid
IDISNA	Instituto de Investigación Sanitaria de Navarra	
IISBiodonostia	Instituto de Investigación Sanitaria BioDonostia	
ibs.GRANADA	Instituto de Investigación Biosanitaria de Granada	
IDIVAL	Instituto de Investigación Sanitaria Marqués de Valdecillas	
CNIC	Centro Nacional de Investigaciones Cardiovasculares	H 12Octubre Madrid
IIB Sant Pau	Instituto de Investigación Biomédica de Sant Pau	



TransBioNet
Network of Translational Bioinformatics
Groups in Research Institutes of
Spanish Hospitals.
(INB hosted)

Hospitals



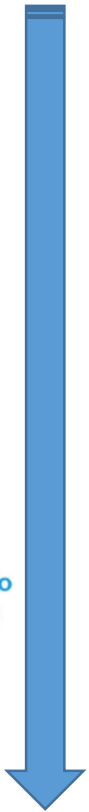
Hospital Universitario
12 de Octubre

Regional Health Systems

SAS

CatSalut
(AQUAS)

SERMAS

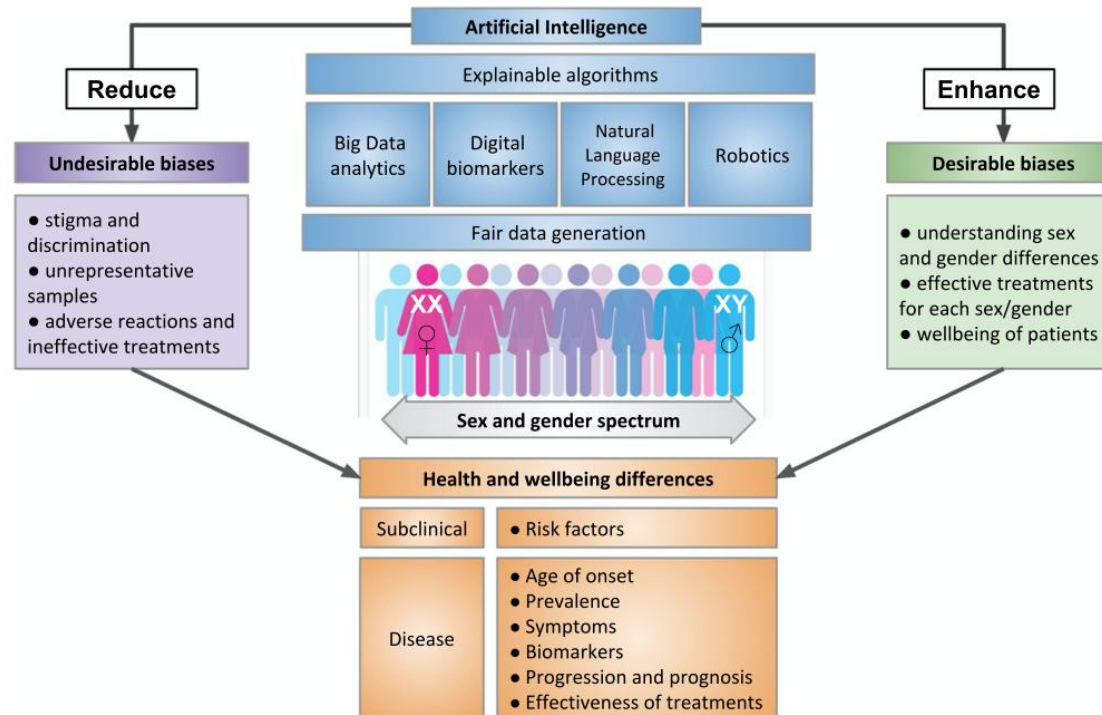


Research contracts as part of the
National NLP Plan



The Bioinfo4Women program of the Life Sciences Department.

(Gender bias as an area of research and activity (fellowships, seminars and conferences by young bright female bioinformaticians)



npj

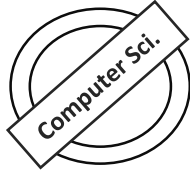
REVIEW ARTICLE OPEN

Check for updates

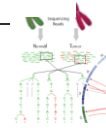
Sex and gender differences and biases in artificial intelligence for biomedicine and healthcare

Davide Cirillo^{1,10}, Silvana Catuara-Solarz^{2,3,10}, Czuee Morey^{3,4}, Emre Guney⁵, Laia Subirats^{6,7}, Simona Mellino³, Annalisa Gigante³, Alfonso Valencia^{1,8}, María José Rementería¹, Antonella Santucci Chadha³ and Nikolaos Mavridis^{3,9}

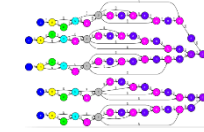
Internal Collaborations



David Carrera (*Data centric computing group*)

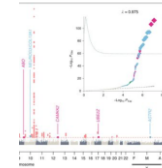


Rosa Badia (*Workflow and distributed computing group*)



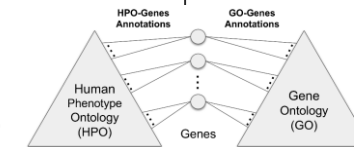
Marta García / Javier Teruel (*with operations*), Code adapt/ optimization

Ulises Cortes Dario Garcia (*HPC for Machine Learning*)

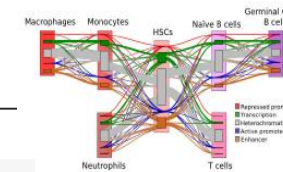


Deep learning version of PELE

Connecting biological ontologies for enrichment applications



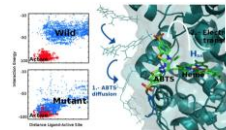
The epigenomic language project (with IBM Ziurich)



Quim Moret Text mining platform National NLP plan.



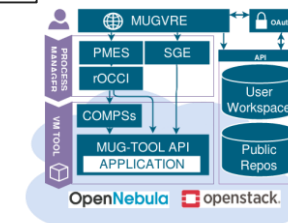
Fernando Cucchiatti, visualization



Mariano Vazquez, finite element simulations Digital Twins



Nadia Tonello, Collaboration in EOSC, data storage





**Barcelona
Supercomputing
Center**

Centro Nacional de Supercomputación