INNDIH

IFIC
INSTITUT DE FÍSICA CORPUSCULAR

VNIVERSITAT DE VALÈNCIA

CSIC
Consejo Superior de Investigaciones Científicas

Artemisa

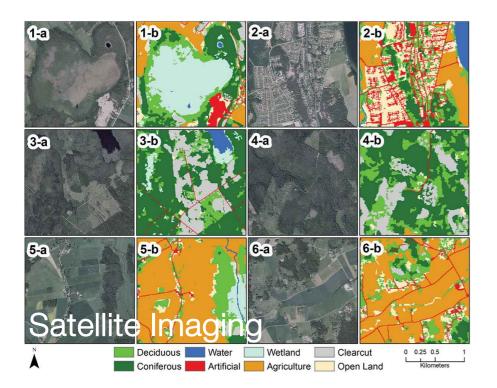## ARTificial Environment for ML and Innovation in Scientific Advanced Computing
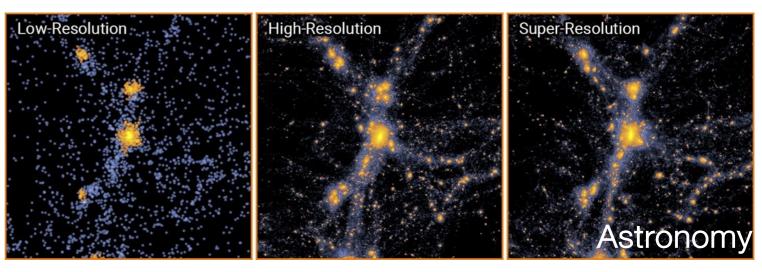
Instituto de Física Corpuscular

300 — Scientists and Technicians

65 — Annual Projects

400 — Annual Publications

Nuclear Physics

Theoretical Physics

Medical Physics

Astroparticles

Particle Physics

AI in Science

Satellite Imaging

Astronomy

Climate Change

Genetics

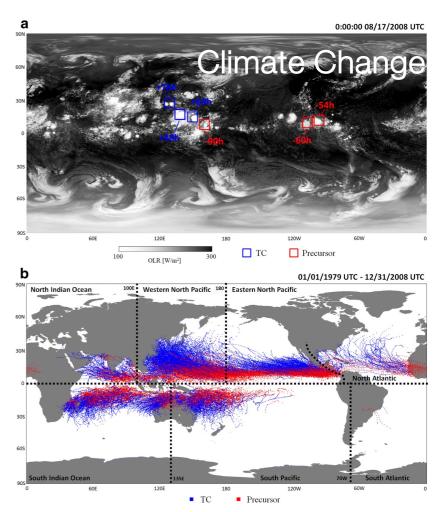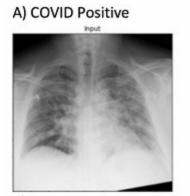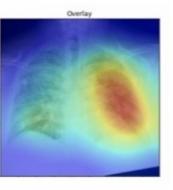image sources: wikipedia / medium.com
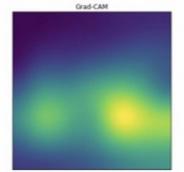
A) COVID Positive

Health

B) COVID Positive

Chemistry

Particle Physics
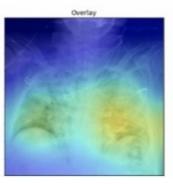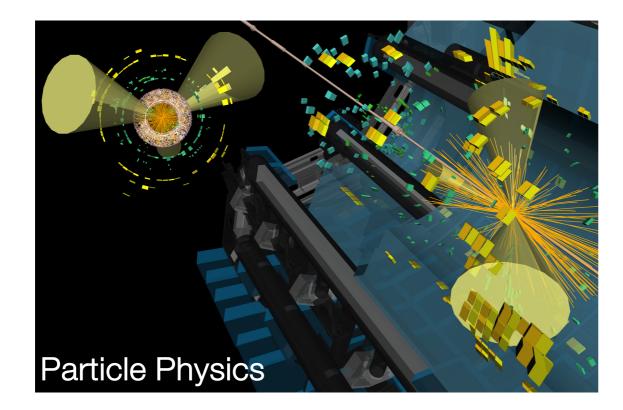
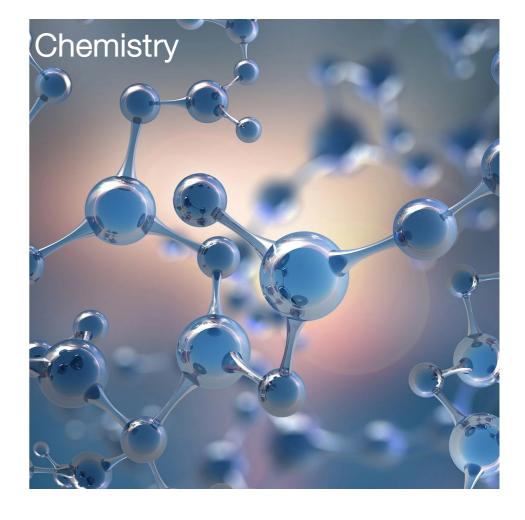image sources: wikipedia / medium.com

Artemisa created from using funds of the FEDER 2014-2020 Comunidad Valenciana (**IDEFEDER/2018/048**, bugdet: **1 M€**).

Granted a new project within the call "Recuperación y Resiliencia" (ASFAE/2022/024). Twofold objective:

- Improve support to users and develop further infrastructure

- Increase computing resources by adding new machines

Artemisa (IFIC) is part of the InnDIH (European Digital Innovation Hub) as part of CSIC and Universitat de Valencia.

Artemisa

Artemisa is a GPU-intensive computing infrastructure dedicated to artificial intelligence and machine learning located at IFIC´s data centre. Its advanced features and excellent performance make possible the accelerated development of projects involving artificial intelligence areas.

Artemisa has machine learning capacity for handling large amounts of data to produce empirical models in physics, chemistry, biology and social studies.

The facility is very well endowed with last generation GPUs plus ancillary CPU and disk space. It features some specific equipment such a modern 8-GPU (A100) Server for special applications.

Artemisa Facility

Who can profit from **Artemisa***?*

Researchers whose projects require computing and memory capabilities beyond those provided by conventional computers.

What AI tools are available?

There are open source libraries for AI tools available, such as Pytorch, TensorFlow, Scikit-learn or Keras. Additional tools can be installed if requested.

Which kind of projects is **Artemisa** *particularly suited for?*

- *Development of machine learning models that require large amounts of data.*

- *Projects which require resolution of optimization problems and/or applications that need distributed computing.*

- *General studies that would profit from high performance GPU usage.*

Artemisa Facility

- *2 user interfaces*

- *22 servers equipped with NVIDIA GPU Volta V100*

- *11 servers with two processors and one NVIDIA GPU Ampere A100*

- *Two multi-GPU Servers: one 4-V100 GPU and one 8-A100 GPU*

- **Artemisa** *also boasts a last generation CPU and storage system*

Artemisa Facility <u>Details</u>

#	Usage	General Characteristics	GPU
2	User interface	2 Intel Xeon Gold 6130 (16c), 192 GB RAM	2 GPU NVIDIA
2	Batch	2 Intel Xeon Gold 8160 (24c), 384 GB RAM	1 GPU NVIDIA Tesla V100 32GB
20	Batch	2 Intel Xeon Gold 6248 (20c), 384 GB RAM	1 GPU NVIDIA Tesla V100 32GB
11	Batch	2 AMD EPYC 7532 (32c), 384 GB RAM	1 GPU NVIDIA Ampere A100 40GB
1	Batch	2 Intel Xeon Platinum 8180 (28c), 768 GB RAM	4 CPUs NVIDIA Tesla V100 32GB SMX2 with NVLink
1	Batch	2 AMD EPYC 7642 (48c), 512 GB RAM	8 GPUs NVIDIA Ampere A100 40GB SMX2 with NVLink
5	Disk Servers	387 TB Lustre	-
3	Disk Servers	150 TB Lustre (SSD)	-

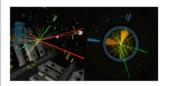## CHEST SCREENING EVALUATION FOR COVID-19 PATIENTS

A project funded by Instituto de Salud Carlos III and led by IFIC researchers is carrying out radiological imaging analysis using machine learning techniques with the aim of enhancing patient diagnose and evolution assessment.
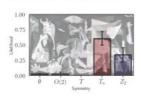
MACHINE LEARNING @ ATLAS EXPERIMENT

The ATLAS experiment at the Large Hadron Collider (LHC) is looking beyond the Standard Model of particle physics, searching for signs of unknown new physics. An important aspect to be able to find this new physics is the identification of the interesting events within all the events available. Interesting events are called "signal", while others are "background". Individuating these signal events, which are indeed extremely rare, is a really challenging task. The LHC has delivered billions of collisions which have been recorded by the ATLAS detector.

DARK MACHINES

Dark Machines (www.darkmachines.org) is an initiative to develop and apply machine learning methods to accelerate dark matter searches. It is composed of more than 300 high-energy physicists, astroparticle physicists and astrophysicists, from theory and experiment, as well as computer scientists.

IDENTIFYING SYMMETRIES THROUGH AI

Since the dawn of humanity, our species has tried to decipher the world around us, through art, literature, music or science. The skills developed and the tools used are different, as different as the audiences targeted. But the goals are basically the same: to dealing with complexity using the tools at hand.

MACHINE LEARNING IN MAGNETIC RESONANCE

Low back pain (LBP) is a very prevalent pathology and a frequent cause of disability. It is associated with rising costs for the health system and for society in developed countries, affecting 70% of the general population at some time in their lives, with an annual incidence of 40%

The multidisciplinary group lead by María de la Iglesia-Vayá from the Prince Felipe Research Center (CIPF) uses Artemisa to develop the first massive and open-access data repository of lumbar MRI for International collaborative research.

CUSTOM-DHM: DIGITAL HUMAN MODELING APPLICATIONS

The objective of the project is to advance in the integration of 3D models of the body in the development of digital products and applications, developing innovative tools that allow their 3D and / or 4D analysis for the clothing, health and wellness, audiovisual and orthopedic sectors, or any other sector that is interested in incorporating digital information from users.

IDENTIFYING WEED SPECIES IN CROP FIELDS

Convolutional Neural Networks (CNN) are currently being implemented in a wide variety of applications. This subdomain of Artificial Intelligence shows a powerful performance in machine vision applications and may be used to categorise and classify objects, amongst other image processing tasks.
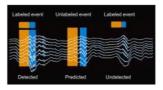
In the Artificial Perception Group of the Centre for Automation and Robotics (CAR) we are interested in identifying and classifying weed species within crop fields, which is a very specific problem, as the system will only need to process images of soil and plants.

NEUROCONVO: BRAIN WAVES WITH NEURAL NETWORKS
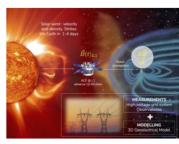
The brain generates activity in the form of oscillations. Brain waves span from very slow rhythms, typical of sleep, to faster oscillations during attention and cognitive processing. Moreover, changes of brain oscillations are markers of some neurological diseases. Given the dynamism of brain activity, these events are far from stationary and thus their identification in real time is a daunting task.

SENTIFLEX

Machine learning is one of the keys in the development of modern Earth Observation satellite missions. Model training requires precise Earth simulation as basis of applications such as vegetation monitoring, prospecting for minerals, soil use and climate change studies, among others.

NON-INVASIVE SCREENING FOR MENINGITIS

165 newborns die every day of Bacterial Meningitis (BM), an aggressive infection that leaves severe sequelae among 30% of survivors. Rapid detection, particularly in this age group, is difficult due to the little specificity and overlap of its symptoms with those of more common and less severe diseases. Current strategy to improve prognosis is the prompt antibiotic treatment after an early diagnosis by means of a lumbar puncture (LP), invasive and potentially harmful procedure.
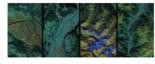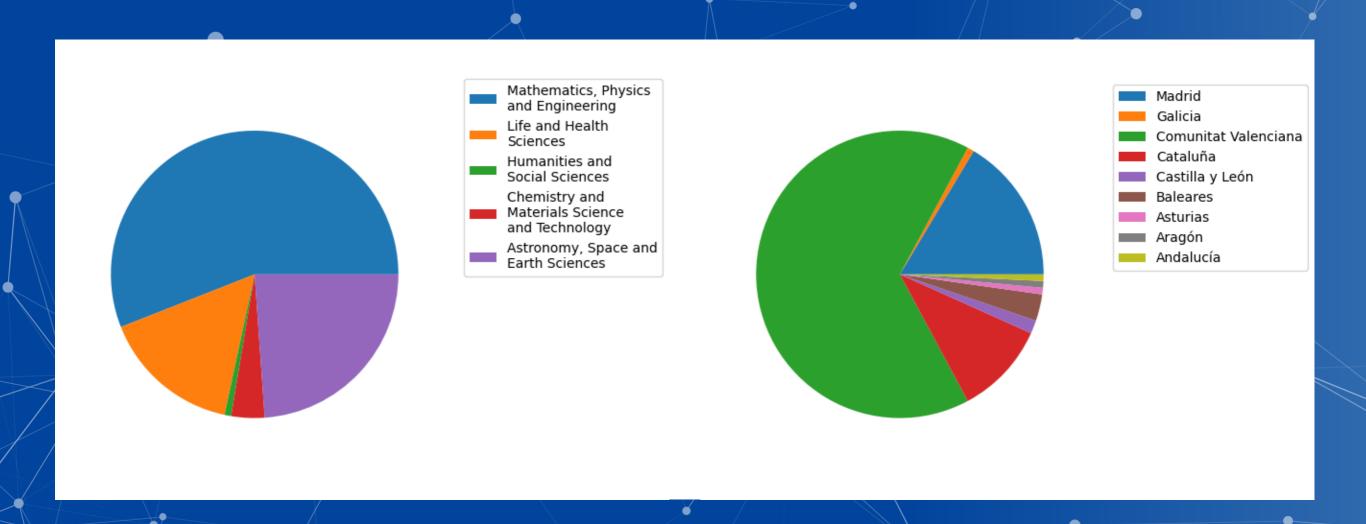
SOLAR STORMS AND THE SPANISH CRITICAL INFRASTRUCTURES

In the last decades, our society has become more interdependent and complex than ever before. Local impacts can cause global issues, as the current pandemic clearly shows, affecting the health of millions of human beings. It is also highly dependent on relevant technological structures, such as communications, transport, or power distribution networks, which can be very vulnerable to the effects of Space Weather. The latter has its origin in solar activity and their associated events, such as solar flares and...
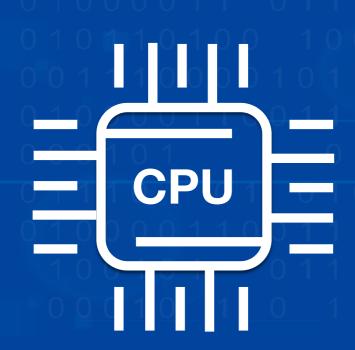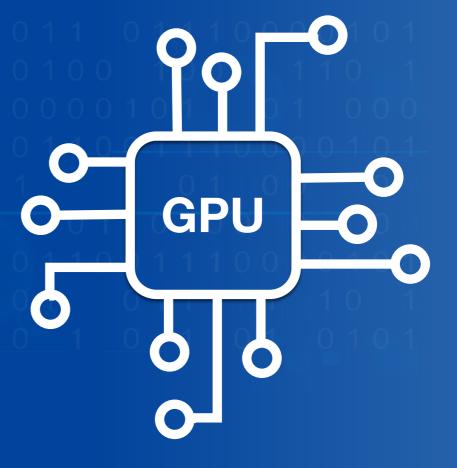
Artemisa Projects

Breakdown by **region** and **research area** for the projects in **Artemisa**. Most of the projects are from "Comunidad Valenciana" and "*Mathematics, Physics and Engineering*" but other regions and fields growing.

3400 CPUs

570.000 Hours Delivered

44 GPUs

290.000 Hours Delivered

- Project applications :

  - **3 Calls per year, 4 months each**

  - **Call is open for a month, accepting projects during that period.**

- Projects are discussed and evaluated by a committee.

- Access through **Artemisa** Intranet.

Application Calls

Project application form

Improving documentation

- Data Security is taken seriously …

What is ISO 27001?

ISO 27001 is the standard created by the International Organisation for Standardization (ISO) which deals with Information Security Management. It is a way of making sure that an entity is managing information security risks and data effectively.

"The ISO 27001 standard helps organisations to establish and maintain an effective Information Security Management System (ISMS), using a continual improvement approach. **You will systematically assess any risks to the organisation's information security and put in place policies and procedures to manage those risks."**

ISO 27001
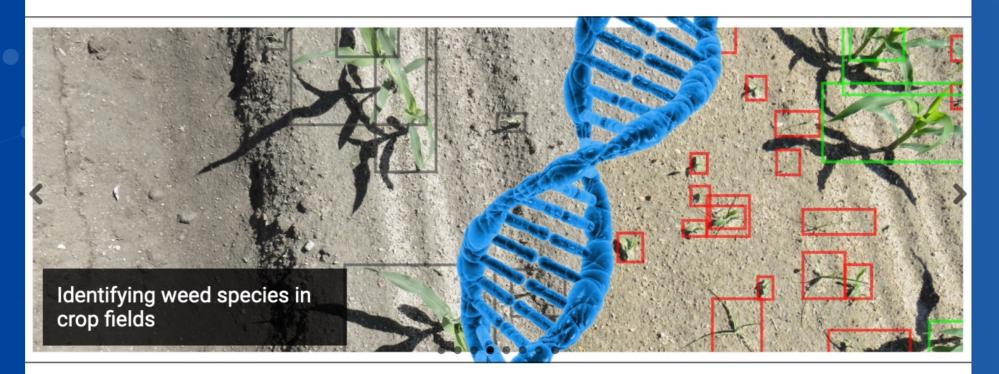
- Deep Learning studies and **Artemisa** usage keeps growing.

- Infrastructure is consolidated after 4 years of successful running,

- Further upgrades are expected for **Artemisa** facility during the next years. Looking for new resources to improve and enhance the infrastructure.

http://artemisa.ific.uv.es

Artemisa completely certified. Infrastructure dedicated to general AI learning.

ISO 27001 is an international security certification. At IFIC includes the full computing facility, personnel and the scientific committee, and a large number of IFIC components dedicated to perform daily operations. It takes care of 114 controls in 14 groups of interest:

A.5: Information security policies (2 controls)
A.6: Organization of information security (7 controls)
A.7: Human resource security - 6 controls that are applied **before, during, or after** employment
A.8: Asset management (10 controls)
A.9: Access control (14 controls)
A.10: Cryptography (2 controls)
A.11: Physical and environmental security (15 controls)
A.12: Operations security (14 controls)
A.13: Communications security (7 controls)
A.14: System acquisition, development and maintenance (13 controls)
A.15: Supplier relationships (5 controls)
A.16: Information security incident management (7 controls)
A.17: Information security aspects of business continuity management (4 controls)
A.18: Compliance; with internal requirements, such as policies, and with external requirements, such as laws (8 controls)

Convolutional neural networks

Multi-layered 'deep' neural networks, that are particularly adapted to image classification tasks by being able to identify the relevant features required to solve the problem.

Transfer learning

Old idea of using concepts learned in one domain on a new unknown one, this idea has enabled the use of deep convolutional nets trained on labelled data to transfer already-discovered visual features to classify images from different domains with no labels.

Generative adversarial networks

Pitching the computer against itself by co-evolving the neural network classifier with the difficulty of the training data set.
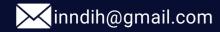
Reinforcement learning

A method for finding optimal strategies for an environment by exploring many possible scenarios and assigning credit to different moves based on performance.

sources: royalsociety.org

¿Qué es InnDIH?

InnDIH es el **European Digital Innovation Hub (EDIH) de la Comunidad Valenciana** que, con el liderazgo de **ITI**, representa la gran apuesta centralizada de colaboración público-privada donde el ecosistema empresarial, las universidades, los centros tecnológicos y de investigación, los institutos de investigación biotecnológica-sanitaria y las Administraciones Públicas aúnan esfuerzos para **contribuir a la digitalización de las pymes y la Administración Pública e impulsar el desarrollo económico de la Comunitat Valenciana**.

Contacto:

✉ inndih@gmail.com

in inndih

🐦 @INNDIHVlcRegion

Focos Tecnológicos

Inteligencia Artificial

Big Data

HPC| Computación de alto rendimiento

Ciberseguridad

Robótica y fabricación avanzada