The CAOS Group at the Barcelona Supercomputing Center

Designing Trustworthy Artificial Intelligence-Based Critical Systems

Axel Brando[1], Isabel Serra[1], Enrico Mezzetti[1], Francisco J.Cazorla[1] and Jaume Abella[1]

[1]The CAOS research group of the Barcelona Supercomputing Center (www.bsc.es/caos).

Emails (@bsc.es): axel.brando; enrico.mezzetti; isabel.serra; jaume.abella; francisco.cazorla
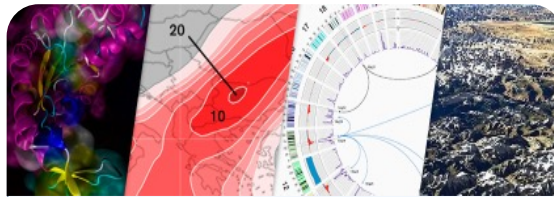
The Barcelona Supercomputing Center
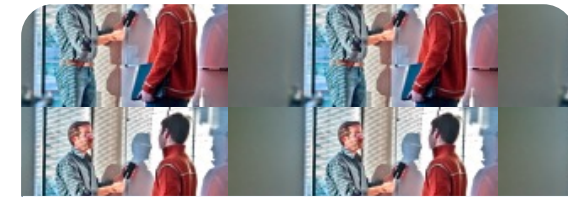
BSC-CNS objectives

Supercomputing services
to Spanish and EU researchers

R&D in Computer, Life, Earth and
Engineering Sciences

PhD programme, technology
transfer, public engagement

**BSC-CNS is
a consortium
that includes**

Spanish Government	**60%**	GOBIERNO DE ESPAÑA / MINISTERIO DE CIENCIA, INNOVACIÓN Y UNIVERSIDADES
Catalan Government	**30%**	Generalitat de Catalunya **Departament d'Empresa i Coneixement**
Univ. Politècnica de Catalunya (UPC)	**10%**	UPC UNIVERSITAT POLITÈCNICA DE CATALUNYA BARCELONATECH

axel.brando@bsc.es
jaume.abella@bsc.es
francisco.cazorla@bsc.es

BSC-CNS organization
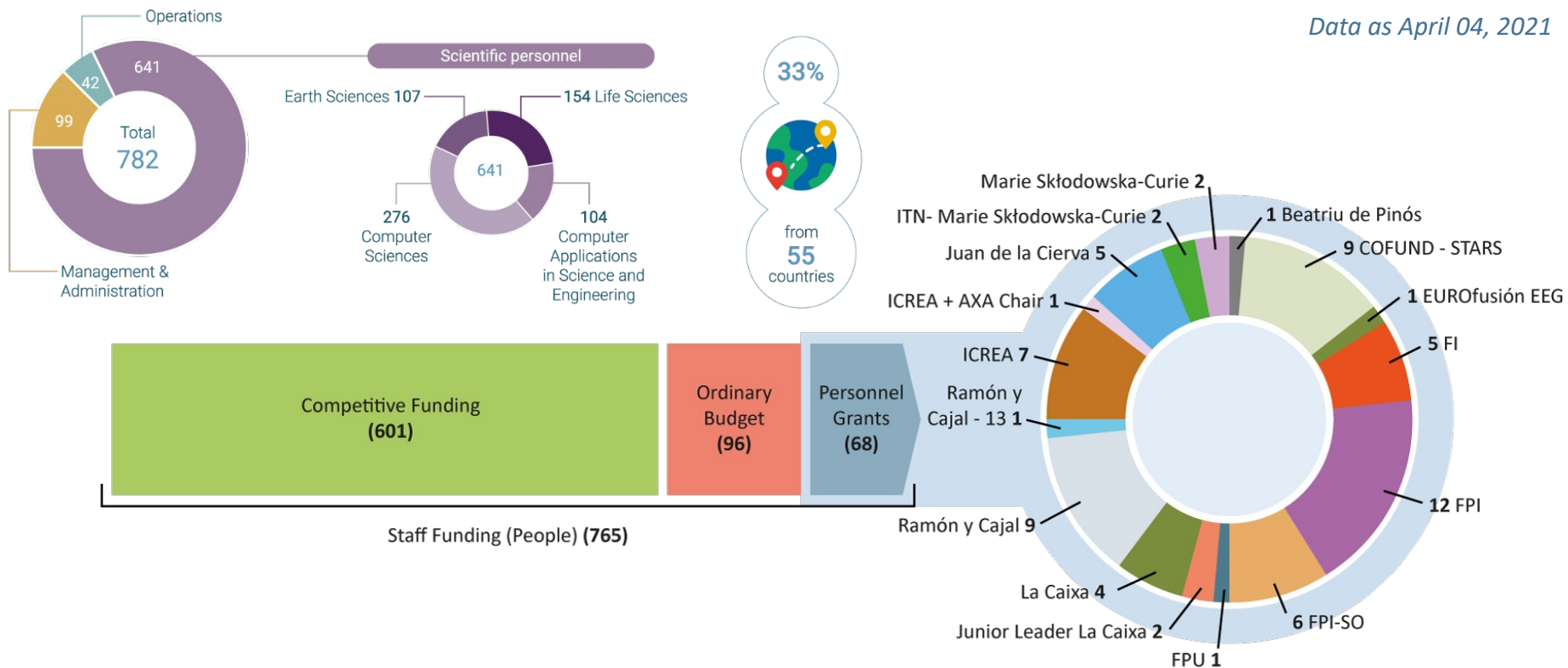
axel.brando@bsc.es
jaume.abella@bsc.es
francisco.cazorla@bsc.es
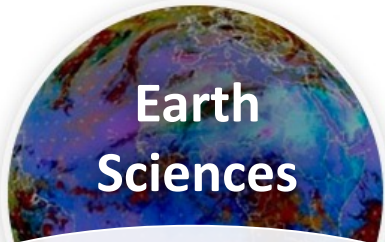
The Barcelona Supercomputing Center

Mission of each Scientific department

Computer Sciences

To influence the way machines are built, programmed and used: programming models, performance tools, Big Data, Artificial Intelligence , computer architecture, energy efficiency

Earth Sciences

To develop and implement global and regional state-of-the-art models for short-term air quality forecast and long-term climate applications

Life Sciences

To understand living organisms by means of theoretical and computational methods (molecular modeling, genomics, proteomics)

CASE

To develop scientific and engineering software to efficiently exploit super-computing capabilities (biomedical, geophysics, atmospheric, energy, social and economic simulations)

axel.brando@bsc.es
jaume.abella@bsc.es
francisco.cazorla@bsc.es

The CAOS group - BSC

- International and multidisciplinary (~50 members, 15+ years of experience)
  - Multiple (**EU projects** (+7) coordinated, (+10) participated, bilateral **contracts with industry**, spin-off creation
  - +200 publications in **top conferences and journals**, several **best paper awards**

axel.brando@bsc.es
jaume.abella@bsc.es
francisco.cazorla@bsc.es

The CAOS – BSC group

- International and multidisciplinary (~50 members, 15+ years of experience)
  - Multiple (**EU projects** (+7) coordinated, (+10) participated, bilateral **contracts with industry**, spin-off creation
  - +200 publications in **top conferences and journals**, several **best paper awards**
- Design and validation **safe, time predictable, and high-performance** hardware and software solutions for Embedded and safety-Critical Systems (ECS)
  - **AI is central due to increasing autonomy of systems** (cars, planes, space missions, robots, etc.)

DO178C

ISO26262

EN50126/8

axel.brando@bsc.es
jaume.abella@bsc.es
francisco.cazorla@bsc.es

Key AI-related research topics in CAOS

- **Safe deployment of AI software in high-performance processors**
  - **<u>Functional safety</u>**: Models with diverse redundancy
  - **<u>Software Timing</u>**: Techniques to derive bounds to the execution of AI software on complex processors

- Designing **probabilistic machine learning models** and AI-based systems:
  - Uncertainty quantification
  - Improve explainability and traceability
  - Combine with Extreme Value Theory (EVT) to enhance AI forecasts

- Creating a generic **Trustworthy Artificial Intelligence** framework
  - Applicable **<u>across multiple domains</u>** (e.g., from autonomous cars, financial, pharmacological to language (LLM) domain).
  - Increase the reliability, robustness and trustworthy of any AI-based forecasting system.
  - Identify probabilistic sources of uncertainty, apply risk mitigation measures and causal analysis, increase interpretability.

axel.brando@bsc.es
jaume.abella@bsc.es
francisco.cazorla@bsc.es

Example of topics of interest in calls 2024

- **Example call:**
  - HORIZON-CL5-2024-D3-02-04: Critical technologies for the future ocean energy farms.
- **Experience:**
  - Designing safety systems and generating probabilistic trustworthy AI models in a wide-range of critical domains.
- **Contribution:**
  - Design and validate embedded systems tailored for ocean energy apps.
  - Develop real-time, safety-critical software for controlling and monitoring ocean energy devices.
  - Identify potential biases, vulnerabilities and propose mitigation strategies.
  - Create and certify robust and reliable (probabilistic) AI-based systems that be used in the development of these critical technologies.
  - Evaluate the resilience of AI systems to unexpected conditions.
  - Implement fault-tolerant mechanisms to handle system failures.

Other examples of calls:
HORIZON-CL5-2024-D3-01-11; HORIZON-CL5-2024-D3-01-10; HORIZON-CL5-2024-D4-02-02; HORIZON-CL5-2024-D4-02-03.

axel.brando@bsc.es
jaume.abella@bsc.es
francisco.cazorla@bsc.es

Summary and contact information

[1]The CAOS research group of the Barcelona Supercomputing Center (BSC-CNS).

- Interdisciplinary group with 15+ years of experience in EU and industrial projects.

- Our keywords include safe, time predictable, high-performance embedded and safety-critical systems, Probabilistic Machine Learning, Extreme Value Theory, Trustworthy AI.

- Regarding cluster 5 areas, our holistic framework can be applied in the different areas when (1) embedded systems are considered, (2) or probabilistic modelling is required, (3) or high-risk AI systems are used:

  Climate sciences and responses
  Cross-sectoral solutions for the climate transition
  Sustainable, secure and competitive energy supply
  Efficient, sustainable and inclusive energy use
  Clean and competitive solutions for all transport modes
  Safe, Resilient Transport and Smart Mobility services for passengers and goods

Axel Brando[1], Isabel Serra[1], Enrico Mezzetti[1], Francisco J.Cazorla[1] and Jaume Abella[1]

Emails (@bsc.es): axel.brando; Isabel.serra; enrico.mezzetti; Jaume.abella; francisco.cazorla

The CAOS Group at the Barcelona Supercomputing Center

Designing Trustworthy Artificial Intelligence-Based Critical Systems

Axel Brando[1], Isabel Serra[1], Enrico Mezzetti[1], Francisco J.Cazorla[1] and Jaume Abella[1]

[1]The CAOS research group of the Barcelona Supercomputing Center (www.bsc.es/caos).

Emails (@bsc.es): axel.brando; enrico.mezzetti; isabel.serra; jaume.abella; francisco.cazorla