Horizon Europe project EXTRACT kicks off with a holistic approach to extreme data across the compute continuum

17 March 2023- The EU-funded EXTRACT project began on 1 January 2023 bringing together a 10-partner consortium from Spain, France, Italy, Finland, Israel and Switzerland. This three-year project will work to provide a distributed data-mining software platform for extreme data across the compute continuum. It pursues an innovative and holistic approach to data mining workflows across edge, cloud and high-performance computing (HPC) environments and will be validated through two use cases that require extreme data: crisis management in the City of Venice and an astrophysics use case.

Photo 1: EXTRACT consortium meets in Barcelona for its kick-off meeting hosted at the home of the Barcelona Supercomputing Center´s famous MareNostrum supercomputer

Data has become one of the most valuable assets worldwide due to its ubiquity in the thriving technologies of Cyber-Physical Systems (CPS), Internet of Things (IoT) and Artificial Intelligence (AI). While these technologies provide vast data for a variety of applications, deriving value from this raw data requires the ability to extract relevant and secure knowledge that can be used to form advanced decision-making strategies.

Current practices and technologies are only able to cope with some data characteristics independently and uniformly. EXTRACT will create a complete edge-cloud-HPC continuum by integrating multiple computing technologies into a unified secure compute-continuum. It will do so by considering the entire data lifecycle, including the collection of data across sources, the mining of accurate and useful knowledge and its consumption.

The EXTRACT Project has received funding from the European Union's Horizon Europe programme under grant agreement number 101093110

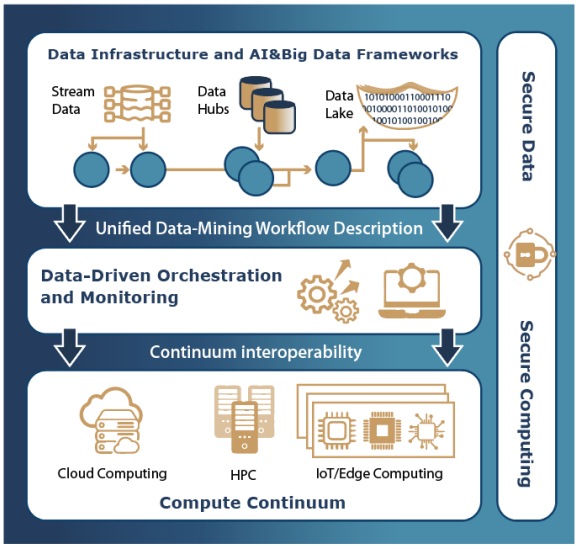

Figure 1. Main EXTRACT platform components . BSC©

The EXTRACT platform will be validated in two real-world use-cases, each having distinct extreme data and computing requirements.

A **Personalized Evacuation Routing (PER) System** will serve to guide citizens in an urban environment (the city of Venice) through a safe route in real time. The EXTRACT platform will be used to develop, deploy and execute a *data-mining workflow* to generate **personalized evacuation routes for each citizen**, displayed in a mobile phone app, by processing and analysing extreme data composed of Copernicus and Galileo satellite data, IoT sensors installed across the city, 5G mobile signal, and a semantic data lake fusing all this information.

The **Transient Astrophysics with a Square Kilometer Array Pathfinder (TASKA)** case will use EXTRACT technology to develop data mining workflows that effectively reduce the huge amount of raw data produced by NenuFAR radio-telescopes by a factor of 100. This will allow the populating of high-quality datasets that will be openly accessible to the astronomy community (through the EOSC portal) to be leveraged for multiple research activities.

Eduardo Quiñones, established researcher at the Barcelona Supercomputing Center and EXTRACT coordinator, is confident that:

> "By seamlessly integrating major open-source AI and Big Data frameworks, EXTRACT technology will contribute to providing the technological solutions Europe needs to effectively deal with extreme data. It will go beyond facilitating the wider and more effective use of data to reinforce Europe's ability to manage urgent societal challenges."

About EXTRACT

The EXTRACT project (A distributed data-mining software platform for extreme data across the compute continuum) is funded under Horizon Research and Innovation Action number 101093110. The project began on 1 January 2023 and will end 31 December 2025. The consortium, formed of 10 partners, is coordinated by the Barcelona Supercomputing Center (BSC). Consortium members include: Ikerlan (Spain), Universitat Rovira I Virgili (Spain), Observatoire de Paris (France), the Centre National de la Recherce Scientifique (France), Université Paris Cité (France), Logos Ricerca e Innovazione (Italy), City of Venice (Italy), Binare (Finland), Mathema srl (Italy), IBM Israel (Israel), sixsq (Switzerland).

The EXTRACT Project has received funding from the European Union's Horizon Europe programme under grant agreement number 101093110