EXTRACT

A distributed data-mining software platform for extreme data across the compute continuum

D5.3 Final Dissemination, Exploitation, Community building and Sustainability report

Version 1.0

Documentation Information

Contract Number	**101093110**
Project Website	https://extract-project.eu/
Contractual Deadline	M39, 31th March 2026
Dissemination Level	Public (PU)
Nature	Report (R)
Author	Agathe Veillon (SIXSQ) (D5.3 Lead) Janine Marie Gehrig Lux (BSC) Mariia Gavriushenko (BINARE) Andrei Costin (BINARE)
Contributors	ALL
Reviewer	OBSPARIS
DOI	10.5281/zenodo.19208167
Keywords	Report, Dissemination, Exploitation, Sustainability, Community

EXTRACT
A distributed data-mining software platform for extreme data across the compute continuum

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

Change Log

Version	Date	Description Change
V0.1	09.03.2026	Document creation: table of contents, sections
V0.2	23.03.2026	Document review and formatting, ready for internal review
V1.0	27.03.2026	Document finalisation. Ready for submission
V1.1		

Table of Contents

1. Executive Summary

From the EXTRACT project's start on 1 January 2023, the Work Package (WP) 5 team responsible for Dissemination, Exploitation and Sustainability has actively sought to take steps towards raising awareness of the project and ensuring the project lives on beyond the project's end. Led by partners Binaré, Barcelona Supercomputing Center, and Sixq, and built on contributions from the entire consortium, the objectives pursued by this work package include:

- Fostering the adoption of results
- Maximizing project impact
- Creating a community of interest around project results
- Ensuring IP is duly managed
- Identifying key exploitable results
- Providing a long-term sustainability plan for EXTRACT technology.

This deliverable, *D5.3 Final Dissemination, Exploitation, Community Building and Sustainability Report*, provides a comprehensive account of the EXTRACT project's communication, dissemination, community building, exploitation, and sustainability activities from project inception through M39. Building on the initial plan established in M6 (*D5.1 Dissemination, Exploitation and Community Building Report*) and its subsequent update in M15 (*D5.2 Updated Dissemination, Exploitation and Community Building Report, First Sustainability Report*), this final report details the actions undertaken, evaluates outcomes, and outlines future strategies to ensure continued impact and uptake of EXTRACT technologies beyond the project's end.

Throughout the project, the consortium has actively promoted awareness and engagement with target audiences, leveraging branding, communication channels, and stakeholder outreach to maximize visibility and adoption. Early efforts focused on establishing a strong project identity and accessible tools, resulting in notable achievements such as increased website sessions and social media followers. The communication and dissemination team, led by the Barcelona Supercomputing Center (BSC), regularly assessed performance against KPIs, identifying successful activities and areas for improvement.

As the project advanced, dissemination activities intensified with highlights including 13 website sessions, ten project videos, 47 event participations and 14 publications. A positive review of dissemination activities by the Horizon Results Booster service in month 35 added credibility to project activities, as well as highlighted opportunities for improvement, which were subsequently incorporated.

Community building and training were also important activities for fostering adoption and practical use of EXTRACT technologies. The consortium engaged with relevant European AI, data, infrastructure, research, and industry communities, and adapted its training approach to better serve both internal partners and external stakeholders. Project activities, training and events were adapted to respond to opportunities and needs, including expanded internal training and a phased rollout of external training, emphasizing recorded materials and ongoing engagement.

Exploitation efforts have been guided by a systematic process for identifying Key Exploitable Results (KERs), managing intellectual property rights, and analyzing

market opportunities. The consortium established clear contact points and roadmaps to facilitate the transition of project outcomes into real-world applications.

The sustainability plan is anchored in two main strategies: leveraging EXTRACT's algorithms through a dedicated marketplace and building a robust ecosystem of stakeholders and collaborators. These efforts ensure that benefits from the project will continue to be delivered and enhanced after the conclusion of financial support.

This report documents the evaluation of performance and KPIs, summarizes lessons learned, and provides actionable recommendations for future dissemination, exploitation, and sustainability activities. The EXTRACT consortium's coordinated approach has positioned the project for lasting influence within the broader scientific and industrial communities.

This final report demonstrates the EXTRACT project's commitment to impactful communication, effective dissemination, strategic exploitation, sustainable outcomes, and community engagement, ensuring that its technologies and results remain accessible and valuable for years to come.

2. Introduction

The *"EXTRACT: A DISTRIBUTED DATA-MINING SOFTWARE PLATFORM FOR EXTREME DATA ACROSS THE COMPUTE CONTINUUM*" project addresses the topic of *HORIZON-CL4-2022-DATA-01-05*, extreme data mining, aggregation and analytics technologies and solutions. It integrates edge, cloud, and high-performance computing (HPC) technologies to enhance data mining methods for exceedingly large, continuously expanding, sparse, dispersed, and heterogeneous data. The project aims to deploy a distributed data-mining software platform, facilitating the efficient development of complex workflows to address extreme data and optimize decision-making. The project´s two use cases, crisis management in Venice and solar activity monitoring using radio telescopes, serve to validate the project technology and highlight project impacts for key societal challenges.

Throughout the project's duration, emphasis has been placed on showcasing the platform's key features and its application in the two use cases. This information has been shared with target audiences to raise awareness about the project and to foster a community around its goals and outcomes.

Deliverable *D5.3* presents a final dissemination, exploitation, sustainability, community building report and final strategies of the EXTRACT project. It provides a description of the activities performed until M39 and details dissemination, community building, exploitation and sustainability for the period M16-M39.

The report presents the outcomes of the strategy laid out in D5.1 *Dissemination, Exploitation and Community Building Report* and updated in D5.2 *Updated dissemination, exploitation and community building, and first sustainability report* to ensure that the project results have reached the different target audiences so they can be exploited and reused to benefit the wider community.

Structure of the document

The deliverable is divided into nine chapters. Chapter 2 contextualizes the final dissemination, exploitation, community building and sustainability plan for the EXTRACT project by providing a summary of the objectives of Work Package 5 (WP5, Dissemination, Exploitation and Sustainability) and the structure of the document. This is followed by Chapter 3, Final Report on Project Communication activities, which describes the project's five main communication channels and their effectiveness from M1-M39. Chapter 4, Final Report on Dissemination activities, details the dissemination activities pursued by the consortium. Chapter 5, Final Report on Community Building and Training Plan, reflects on the activities undertaken to identify and build a community around the project and how these activities raised awareness of this community about EXTRACT results through training, event participations, and collaboration in the European Artificial Intelligence (AI), data, research and industrial communities. Chapter 6, Final Exploitation Plan and Report, highlights key activities conducted after M15 to ensure the identification and proper exploitation of key results. Key Exploitable Results were identified and characterized, with individual roadmap defined, and Intellectual Property management was refined. Chapter 7, Final Sustainability Plan, defines sustainability in European Research projects and outlines EXTRACT's dual approach to ensure longevity. It focuses on activities conducted between M16 and M39 to raise awareness on sustainability issues and to prepare the end of the project. It tackles results availability and platform replicability. Chapter 8, Evaluation of Performance and KPIs (until M39), presents activities, KPIs, measures and status. The deliverable D5.3 concludes with Chapter 9, Conclusion.

Aim, Scope, Objectives

This deliverable reports on the status of the activities defined in *D5.1 Dissemination, Exploitation and Community Building Report* (D5.1) and incorporates insights from D5.2 *Updated Dissemination, Exploitation and Community building report. First Sustainability Report* to provide details on key exploitable results and plans for future sustainability. This deliverable reports on the status of activities from M16-M39.

D5.1 set the initial framework for communication, dissemination, training and exploitation, while D5.2 monitored the activities set out and reflected on key steps to ensure project KPIs. Key activities included the creation of the definition of target audiences and messages, the creation of a cohesive brand and corresponding communication and dissemination channels and tools to facilitate the consortium´s ability to to reach the defined target audiences appropriately. It has also identified initial actions needed to identify and build a committed community around project results and a training plan that would help ensure project results could be used by potential stakeholders.

The goal of the present deliverable D5.3 is to ensure that the project aligns with the initially defined goals. D5.3 objectives include:

- Describe the activities conducted in the period M16-M39;
- Reflecting on the success of the strategy by measuring project results against key project indicators from M1 to M39;

- Reflect on project outcomes and future usage.

3. Final Report on Project Communication Activities

EXTRACT project communication activities were carried from the onset of the project and were designed to effectively share project results with the target audiences and wider stakeholders. The potential impact of the project and its application of the extreme data mining technology in two use cases was emphasized to demonstrate the potential of the project to address key societal challenges identified by the European Commission.

An initial communication strategy was established part of *D5.1 Dissemination, exploitation and community building report,* submitted at M6*,* defining target groups, messaging, activities and success indicators across five main channels: 1) branding and project identity 2) project website, 3) social media, 4) press strategy and 5) promotional material. The strategy was reviewed and updated at M18 as part of D5.2 to ensure continued relevance and alignment with project objectives and to seize emerging opportunities.

The communication strategy was rolled out in three main overlapping stages. The first year of the project focused primarily on raising awareness of the EXTRACT project among the target audiences defined in *D5.1* (research, industry, policymakers, and the general public) by using different tools and channels and adapting key project messages tailored to each audience (research and industrial community, policymakers, general public). The second year of the project emphasized building a community with related projects and key EU initiatives, including the AI, Data and Robotics Association (ADRA), AIonDemandPlatform (AIoD) Big Data Value Association (BDVA), European Open Science Cloud (EOSC) and others, broadening the reach of results and laying the groundwork for lasting impact. During this time, the analysis of our activities in D5.2 helped identify opportunities and areas of improvement, which ensured that the third year of the project properly supported the dissemination and uptake of project results through presentations, workshops and publications.

The following sections report on the status of these activities at M39, reflecting the consortium's continued commitment to clear, impactful and audience appropriate communication through the project's lifetime.

3.1 Branding and Project Identity

A clear and distinct project identity was established from the beginning of the project to facilitate a clear narrative that would help raise visibility and recognition for the project. Project templates for presentations, posters, social media and a brand guide were set out at the beginning of the project and were implemented consistently throughout the project's lifetime. Branding materials and guidance were made available to all consortium members via the internal repository.

All materials created included the name of the project, the website, the EU acknowledgement and a version of the project logo, all in the project's signature colors and style. A detailed description of the graphic identity is available in *D5.1.* Project partners consistently applied the project brand in their communication and dissemination activities.

Success for the project brand identity was measured by the creation of key reusable communication materials, including one power point template and overview presentation, word and poster templates. Additionally, a brand guide helped define use of design elements and the dissemination officer oversaw their correct use. This KPI was met.

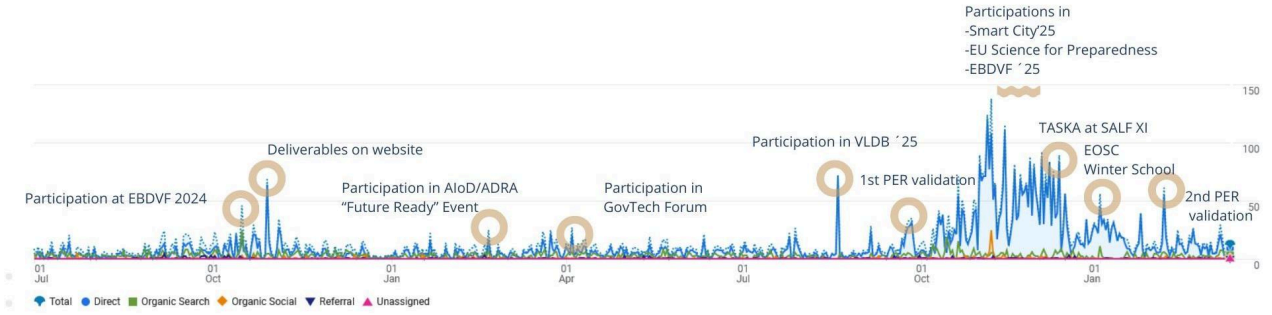
3.2 Project Website

The EXTRACT project website https://extract-project.eu served as the main hub for publication communication of the project, directed at all target audiences. The project website was maintained by partner Barcelona Supercomputing Center (BSC) with the support of the entire consortium. In addition to including important information about the project (objectives, partners, use cases, equality and diversity, branding), it also served to host project results (publications, deliverables and demos) and counted with updated news and events on the project's progress.

Content was regularly updated thanks to contributions from all partners, including monthly news pieces providing progress overviews on their different activities (73 published by M39). These insights into the work of each partner provided fresh content that could be shared across the other project communication channels, thus increasing the potential for engagement. Publication and event pages were also frequently updated and served as important sources of information for project target audiences. The publication section is the most visited page on the website, after the homepage, and it contains the open-access publications produced by the project in addition to the slides and posters presented at events. Sixty-four items are available in this section. The event section lists past and upcoming events to encourage visitors to learn about the types of events the consortium participates in and where there could be possible opportunities for collaboration.

Impact of the project website was monitored through the number of website sessions and active users measured using Google Analytics. A key performance indicator (KPI) of 3000 sessions by the end of the project was set as part of D5.1. At M18 there were 2,500 sessions and by M39, this number increased to 13,000 sessions, vastly surpassing the initial goal. Moreover, active users also increased markedly with 1,200 active users in M18, compared to 8,900 active users at M39. Analysis of peak website traffic periods (see Figure 1) shows a correlation between participation in events and updated content on the web suggesting that these activities drove audiences to the website to learn more, confirming its effectiveness as a main communication tool.

Figure 1: EXTRACT website sessions recorded by Google Analytics M19-M39, pin-pointed to some main events and achievements.

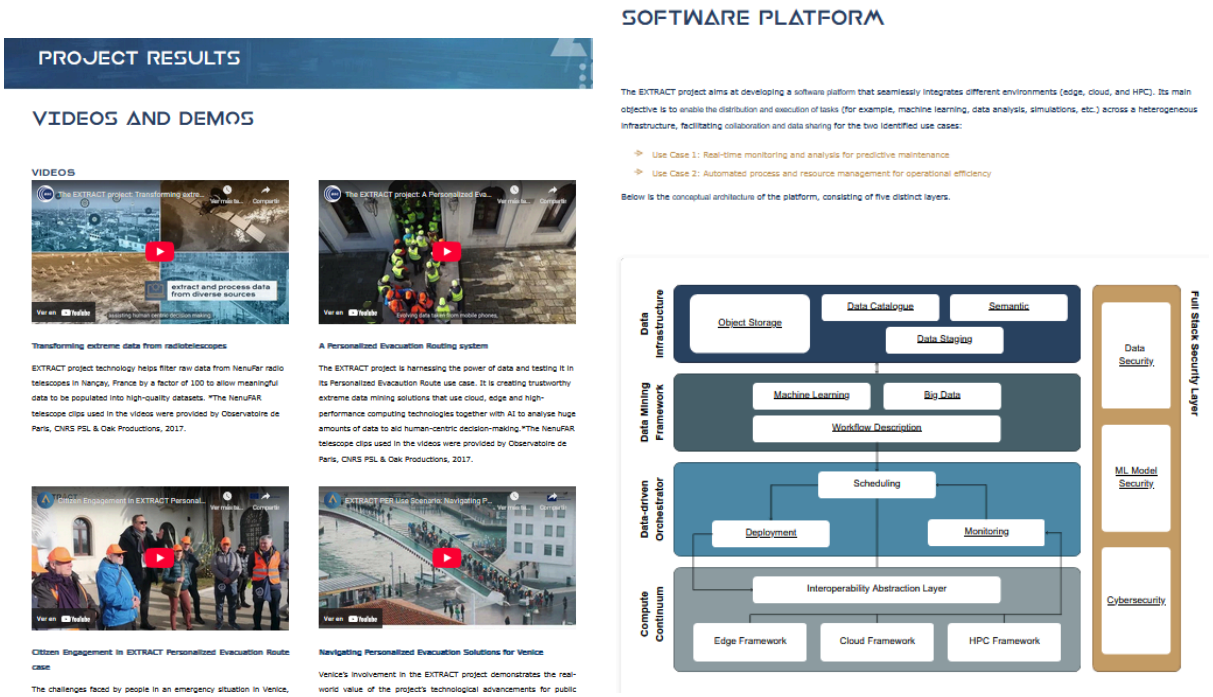


In addition to monitoring website sessions and users, Google Analytics was used to analyze the most visited web pages to better understand if website content was relevant for project audiences. Visitors primarily visited the project homepage, the publications page, and the news page, indicating interest in key project results, like publications and presentations, as well as project progress.

At M18, the project website expanded to include a section on project videos and demos and on the project's software stack. Given the popularity of multi-media and the growth in number of project videos, a new webpage was created to centralize the videos and demos in one place to facilitate access and encourage users to watch and explore. More information on the project videos is available in the section on promotional materials.

A dedicated software stack webpage was also created to show how the different EXTRACT technologies integrate into a unified platform spanning edge, cloud and HPC environments. Building on the software stack diagram, the page guides audiences through each architectural layer, linking diagram elements to detailed technology descriptions to convey how the platform manages the full lifecycle of a data-mining workflow. The projects Key Exploitable Results will be linked to this page, with details on how to access them and their target user groups.

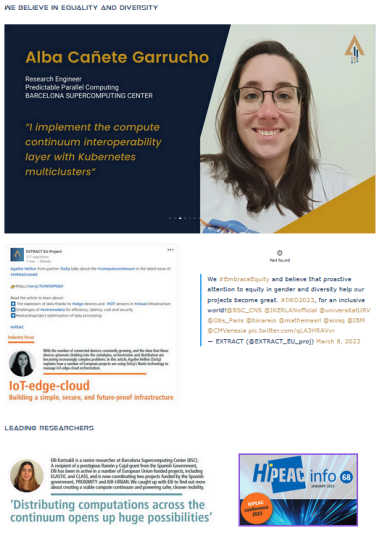
Figure 2: Screenshot of updated EXTRACT project website pages created after M18.



3.2.1 Diversity and equality

Promoting diversity and equality has been a consistent activity through the project. A dedicated Equality and Diversity page was developed to centralize and showcase the individual contributions of each partner in this area, providing a unified view of the consortium's collective commitment to this goal. The page also hosts a section highlighting the women driving EXTRACT research and development. Participation in awareness campaigns and spotlight features profiling the women in the project help bring visibility to the inclusive research culture the consortium strives for and reinforce the project's support for diversity in European research and innovation.

Figure 3: Excerpt from the EXTRACT Equality and Diversity web page



The website's initial success indicator was set at 3,000 sessions by project end. This target was raised to 6,000 sessions following the monitoring reported as part of D5.2, where the project already had 2,500 sessions. The project ultimately surpassed both targets, accumulating over 13,000 sessions by M39. This KPI surpassed its original target.

3.3 Social Media

Social media was used throughout the project's lifetime as a dynamic and responsive channel for reaching all target audiences (research, industrial, policymaker and general public communities). It complemented the project website by providing quicker and more varied communication. Social media accounts were used to highlight participation in events, updates and to raise visibility and connections with the broader EU research community and sibling projects. Posts were consistently designed to drive traffic back to the project website, so that audiences could access deeper technical details and project outputs.

The project initially launched two social media accounts, 1 LinkedIn and 1 X (ex-Twitter) account. However, following the analysis at M18 as part of D5.2 and in response to feedback from project partners, a BlueSky account was opened to diversify the project's social media presence in light of the evolving landscape of X.

EXTRACT EU Project @EXTRACT_EU_proj @extracteuproject.bsky.social

The X account was gradually faded out as the more general public and political audiences targeted with that account started to leave and EXTRACT X followers decreased massively. Bluesky was used to try and fill the gap left by X, however, by M30, it became clear that this channel was not very successful, with only ten followers.

Instead of adding resources to this channel, the dissemination team focused efforts on the LinkedIN account. This account, initially targeting industry and researchers, saw a large increase of followers (494 at M39, compared to 250 at M18).

3.3.1 LinkedIn

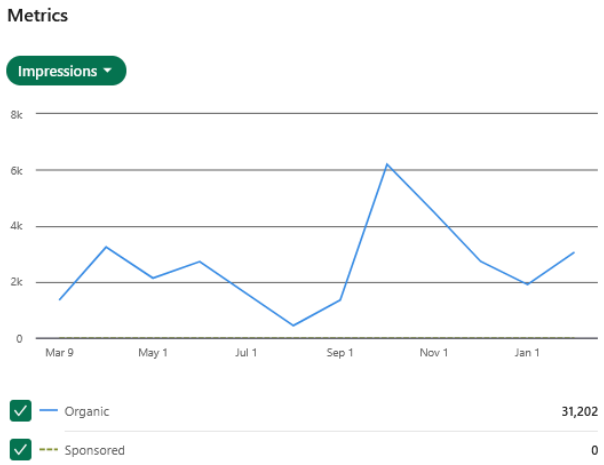
With 494 LinkedIn followers at M39, and an average engagement rate of 21% in the last year of the project, LinkedIn has been a very successful channel for engaging with our audiences. The LinkedIn account has been used to post 2-3 times a week on project highlights, engage with the EU AI and research ecosystem and industry and to bring attention to project results and activities, including participation in international awareness days and fun campaigns like Flashback Friday that keep the LinkedIn channel very active and offer an engaging way to highlight past presentations and event participations.

The EXTRACT account follows and builds connection using hashtags such as #extremedata #compute #continuum #urbandigitaltwin #datamining #HorizonEurope #HPC #IoTs #edge and we engage with and follow projects from our cluster (HORIZON-CL4-2022-DATA-01-05) and the European Artificial Intelligence (AI) Ecosystem AI, Data and Robotics Association ecosystem (ADRA),

AIonDemand, the Big Data Value Association (BDVA), EOSC Association among others.

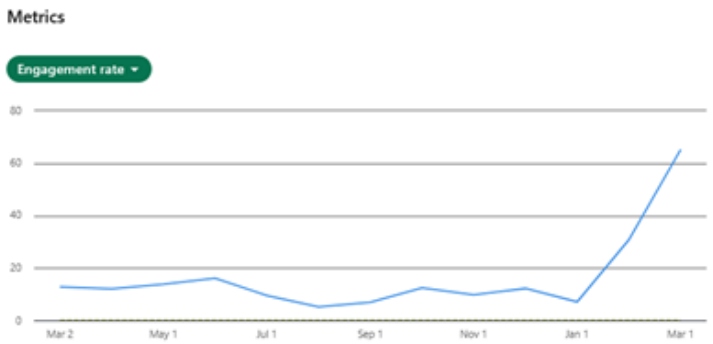
Activity is monitored via LinkedIn Analytics, which shows general information on followers, visitors, visitor demographics and more. Figure 4 demonstrates the evolution in the amount of impressions (number of views) in the last year of the project. At M18, the project reported 18,458 impressions on the channel, and in the last year of the project, there were more than 32,000. These numbers suggest that LinkedIn is working well to reach our target audiences and that the material we post was of interest to our audience.

Figure 4: LinkedIn impressions recorded by LinkedIn analytics from M27-M39. The free version of LinkedIn Analytics only measures a twelve month period.



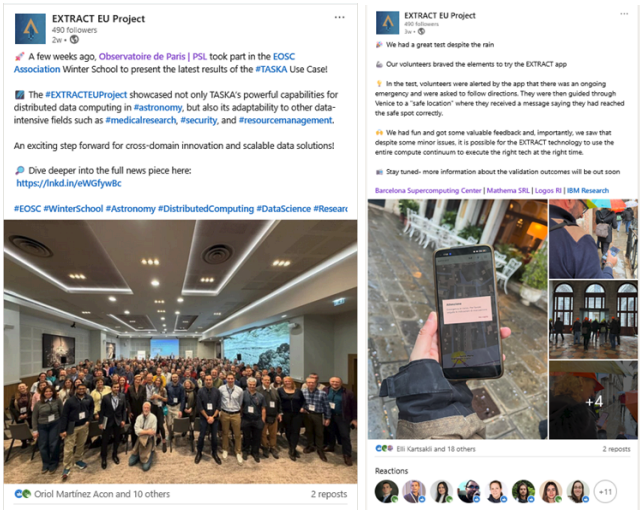
In addition to the high number of impressions, the engagement rate (interaction with content, e.g. shares, comments, likes, clicks) also reflects positively on the success of the LinkedIn strategy. The average engagement rate in the last year of the project has been 21%, with rates reaching as high as values now as high as 60% in the last months of the project. Given that a 2% engagement rate is considered to be good in the context of social media marketing, the social media strategy for LinkedIN was extremely successful.

Figure 5: LinkedIn engagement rate as recorded by LinkedIn analytics from M27-M39.



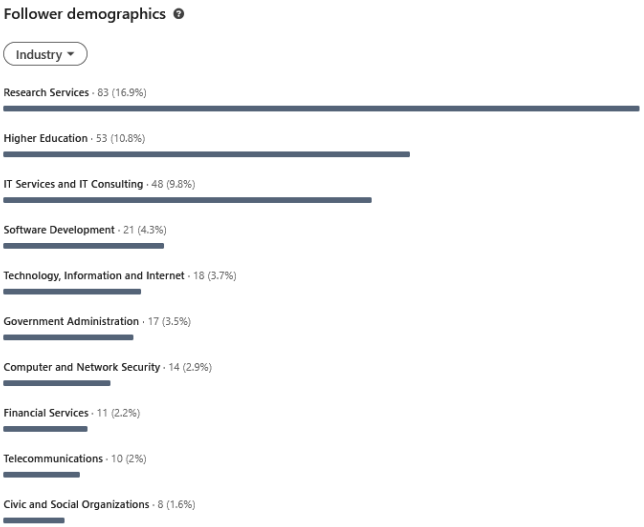
Some examples of especially successful content on LinkedIn include posts announcing links to participation in European events, like the TASKA training at the EOSC Winter school and the second PER validation in Venice (Figure 6). The EOSC post received 264 impressions and an engagement rate of 9,47%, while the PER post received 495 impressions and a 65,06% engagement rate.

Figure 6: Examples of posts with high impression and engagement rates.



The WP5 dissemination team also monitored professional and demographic information using LinkedIn analytics to develop a clearer understanding of the project's audiences and their interests. By M39, EXTRACT LinkedIn followers primarily came from research, higher education and IT services and consulting, with a meaningful presence from government and civic and social organizations (Figure 7). This distribution closely reflects the project's target audiences, suggesting that the project's communication efforts successfully reached the communities who could most benefit from project outputs and long-term impact.

Figure 7: LinkedIn follower percentage by type of industry.



Campaigns were created to show project support and involvement in impact-focused campaigns like women in science, diversity, AI, space, crisis management and HPC. The EXTRACT project has participated in International Day of Women and Girls in Science (11 February), International Women's Day (8 March), Space Week (4-10 October). Although these campaigns are not strictly connected to the project´s results or scientific objectives, they are important for emphasizing the values of the project and the consortium while gaining access to a larger platform of like-minded individuals and organizations.

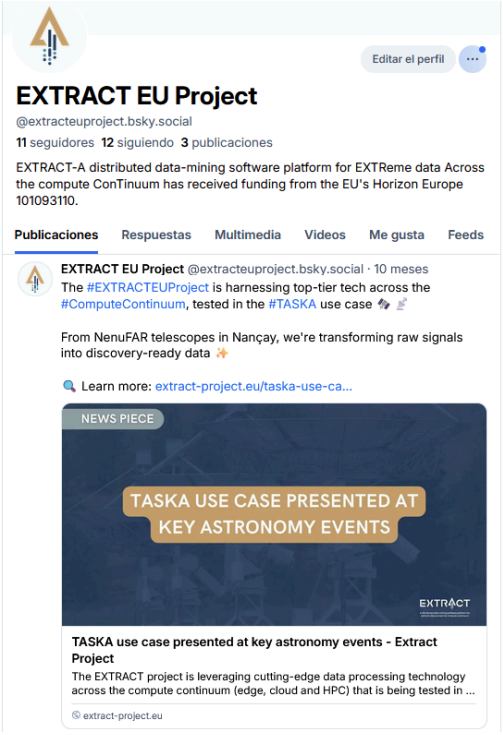
3.3.2 X (ex-Twitter)

Following the mid-term review of the project, it became increasingly apparent that X (ex-Twitter) was no longer a useful channel for the project and project partners had professional and personal reasons for no longer wanting to support this channel. The account was phased out and eventually closed by the end of 2024.

3.3.3 Bluesky

The project dissemination team briefly experimented with Bluesky to see if it could be an alternative channel to X. However, building a community in this channel did not prove effective. Initial attempts to engage other X converts in the European Community and piggy-backing on partner institutional content to increase visibility, did not yield any notable results. After roughly eight months, project updates on this EXTRACT project account were phased out. The account still exists but is not in use.

Figure 8: An excerpt from the BlueSky account and types of content.



The social media channel pursued by the WP5 dissemination team was moderately successful. **The project met its success criteria with 500 combined followers between LinkedIn (494) and Bluesky(10). X and Bluesky were not effective and were phased out, while the LinkedIn channel proved successful and a good fit for the project target audiences. The KPI of 400 followers was met.**

3.4 Press Strategy

EXTRACT complemented its online presence with a traditional press strategy to ensure project updates reached beyond specialist audiences and targeted key moments such as the project launch, cluster collaboration activities, and high-visibility events like the Smart City World Expo Conference and policy-focused events held as part of the personalised evacuation routing system case validation.

Two official press releases were issued over the course of the project, the first on 3 March 2023, which introduced the project's objectives, partners and use cases and received attention from prominent technical media. The second, launched on 02 July 2024, introduced the EU sibling cluster activities. The final project press release is planned for M40, following the submission of project deliverables.

Figure 9: Second EXTRACT press release, launched 2 July 2024 to the technical media


NEWS

Introducing DATANEXUS: the EU Cluster Harnessing Extreme Data using Advanced European Technologies

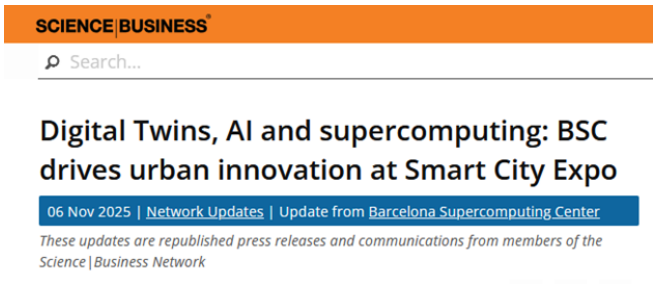
Date: July 01, 2024

DATA NEXUS

As part of the Horizon Europe funding programme, seven pioneering projects have received funding to address the challenges posed by extreme data generated by Internet of Things (IoT), industrial, business, administration, environmental, scientific, societal sources. These projects are pursuing innovative approaches that integrate cutting-edge technologies, including artificial intelligence (AI), high performance computing (HPC) and Edge/Fog/Cloud Computing.

Beyond formal press releases, partners cultivated relationships with technical media outlets, contributing tailored pieces or spaces for exchange for their channels. BSC and the Metropolitan City of Venice were featured in several press clippings, while other partners leveraged their institutional channels and networks to further amplify project news. External references to the project, counted as press clippings in this project, add up to 18. This is a satisfactory number of press clippings, and WP5 will ensure the final press release is informative and reaches a broad audience of interested stakeholders.

Figure 10: Examples of press clippings feature EXTRACT partners BSC and Metropolitan city of Venice



Success for this activity was set at three project press releases with 25 press clippings. This indicator is on track to be met with a third press release planned for the end of the project. It is expected that there will be several press impacts related to it. Additionally, BSC is preparing an updated text for the HiPEAC Magazine, any important reference in the field.

3.5 Promotional Material

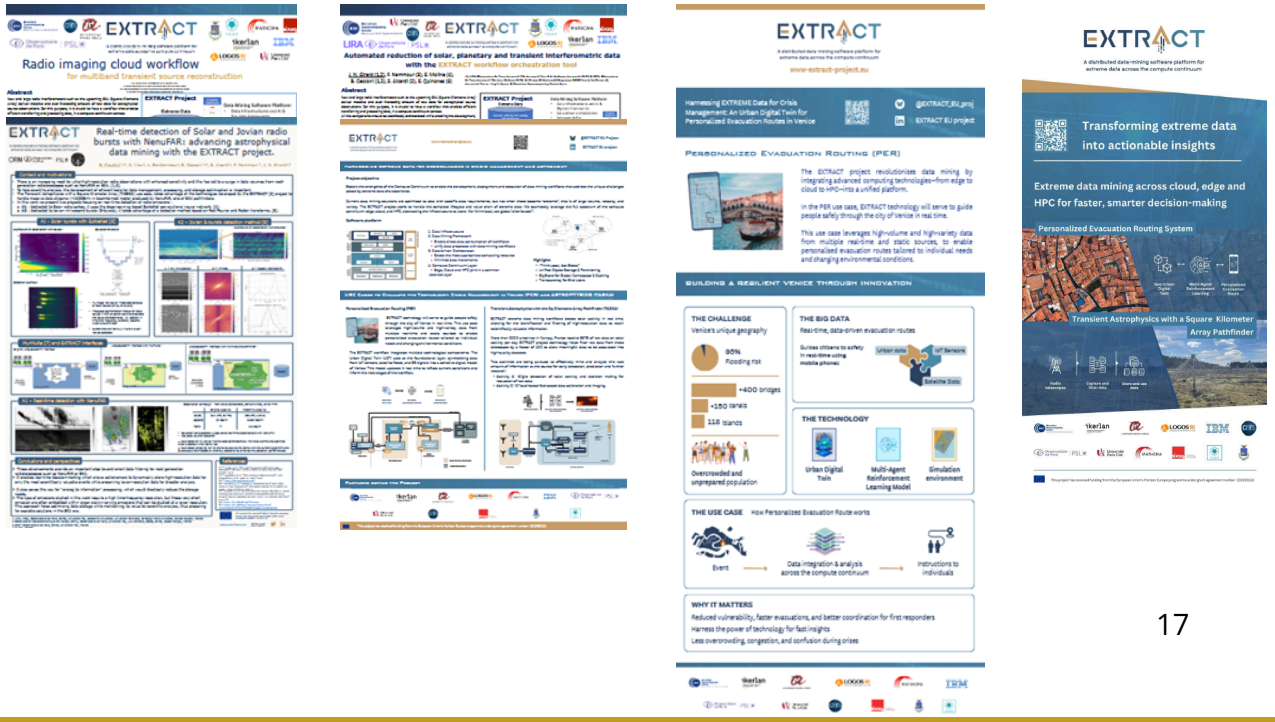
Tailored communication and dissemination material was prepared and adapted for use in the consortium's communication and dissemination activities. In all, four project flyers, nine project posters, ten project videos and nine demos were created to build project recognition, coherence and credibility. All promotional materials uniformly and consistently acknowledged project funding and helped to ensure that project concepts were accessible to a wide range of audiences.

- Overview presentation: Project partners also have access to a general overview PowerPoint presentation of the project, which can be used as an introduction to their presentations that focus on specific aspects of the project. This presentation helped transmit the project's objectives and key messages consistently.
- Flyer: In addition to the project's initial overview, tri-fold flyer, which offered a look into the objectives and the use cases, three more flyers were prepared in digital and physical form. Two were used in the two editions of the Smart City World Expo Congress sessions in 2024 and 2025 to raise awareness prior to the event via social media, email and the project website as well as distributed on-site to attract audiences. A flyer was also created and circulated at the 2025 European Big Data Value Forum and the 2025 EU Science for Preparedness Event showing an overview of the technology and its application in the two use cases. The flyers can be downloaded from the Branding section of the website.

Figure 11: The four EXTRACT flyers produced by the project

- Poster and roll ups: From M19-M39 six increasingly sophisticated and detailed posters were created to highlight project results, for a total of nine posters. Two posters provide information on the project as a whole (Harnessing Extreme Data for Decision-Making in Crisis Management and Astronomy-EGI conference, EBDVF roll up), three posters focused on the TASKA use case (Real-time detection of Solar and Jovian radio bursts with NenuFAR: advancing astrophysical data mining with the EXTRACT project, Radio imaging cloud workflow for multiband transient source reconstruction, Automated reduction of solar, planetary and transient interferometric data with the EXTRACT workflow orchestration tool, and one focused on the PER use case and was presented.

Figure 12: Project posters

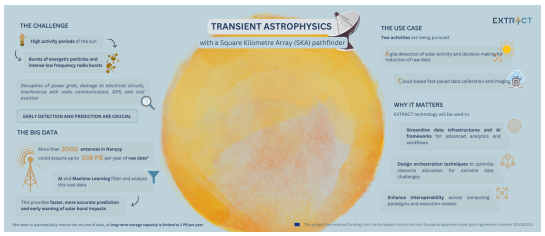


- Videos: Promotional videos offered another means of reaching and engaging with project target audiences. Ten videos were created to contextualise use cases, highlight collaborations and explain project technology. Videos were updated on the project YouTube and shared via the project's social media accounts. Videos ranged from overview videos about the project, to those document user engagement and value of the PER use case in Venice, reports from events and collaboration activities with the DataNexus cluster of sibling EU projects (See Section 5.1 on Community building to learn about this collaboration). Due to the increase in project videos from M21, a new project webpage was created to facilitate access to the project videos. At M39, the project videos had more than 1,000 visualisations. Two videos were produced with the help of the Horizon Results Booster (HRB) service (learn more about this relationship in the section on Section 4.3 Dissemination support from HRB). Moreover, nine demos were updated to help share information on project technologies and their use.

Figure 13: Videos and demos on the EXTRACT project website

- Factsheet/infographic: Several infographics were produced throughout the project to convey its goals, use cases and impact in a visually compelling and accessible format. Each infographic was created to guide non-technical audiences through core elements of the project, including the needs of the use cases, the data available and why results matter, to create engaging narratives. A dedicated factsheet was also developed to showcase the business sustainability plans produced during the EXTRACT Hackathon hosted by BINARE, capturing one of the project's key community-building outputs in an easy to share way.

Figure 14: Infographics about different EXTRACT project areas



Success criteria for promotional materials included two posters, three videos and one factsheet. As the project progressed, it became clear that the volume and variety of dissemination opportunities the consortium participated in warranted a broader range of materials and in the end four flyers, ten videos, one factsheet and two infographics were prepared. These materials are available on the about section of the project. This KPI was exceeded.

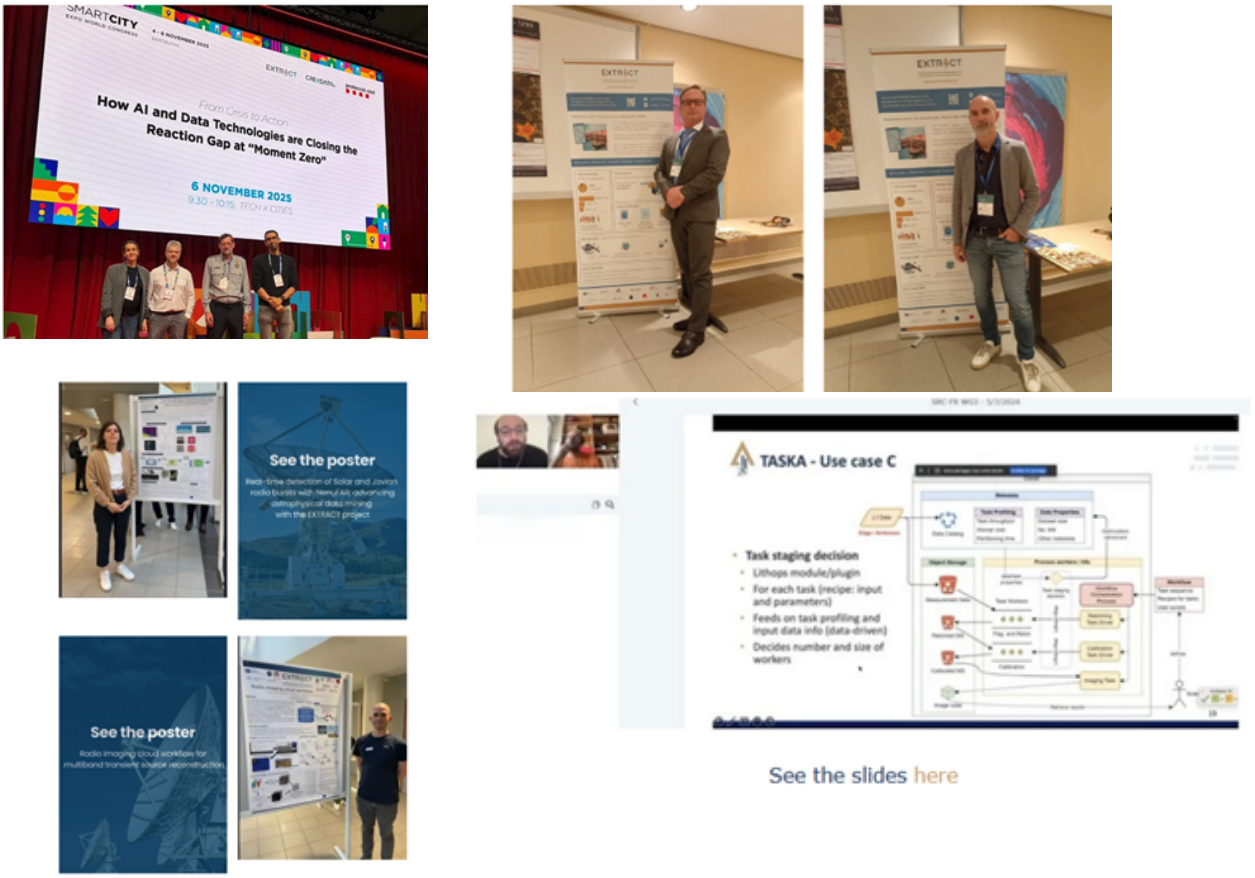
4. Final Report on Dissemination Activities

Building on the awareness and visibility established in the project's first phase, activities from M19 to M39 shifted toward active dissemination and the uptake of results. Beyond communication objectives and impact to target audiences, WP engaged stakeholders through participation in relevant events and publications in peer-reviewed journals and conference proceedings. As project results matured, there were greater exchanges with the audience who could most benefit from them. By M39, the consortium participated in 41 events (26 of them during M19-M39), and published 14 peer-reviewed publications, driving the uptake and sustainability of project technologies.

4.1 Event Participation

From 19-M39, the consortium participated in 26 events: 11 focused on research communities, eight for industry, including presence in industrial booths, and seven events and a booth geared toward policymakers and the general public. For research communities, presentations focused on the details of the data mining workflows, the application of these workflows in research scenarios, like astronomy or urban management, and specific project technologies like the DataPlug, Lithops, the Urban Digital Twin, the MaRL and simulator, Kubernets, SkyStore, etc. For industry, the focus was placed on the ability of technology to exploit the compute continuum and open access to EU-funded technologies that could be tried and further developed in different scenarios. For policymakers and the general public, a great focus was placed on showing the advancements made in technology that could be placed to the service of societal challenges through partnerships with cities and universities who were helping to validate these technologies.

Figure 15: Photos from event participations, left to right, side session at the 2025 Smart City Expo World Congress, poster presentation and networking at the 2025 EU Science for Preparedness JRC event, poster presentations in MaDiCs 2025 conference, presentation to a working group of SKA researchers.



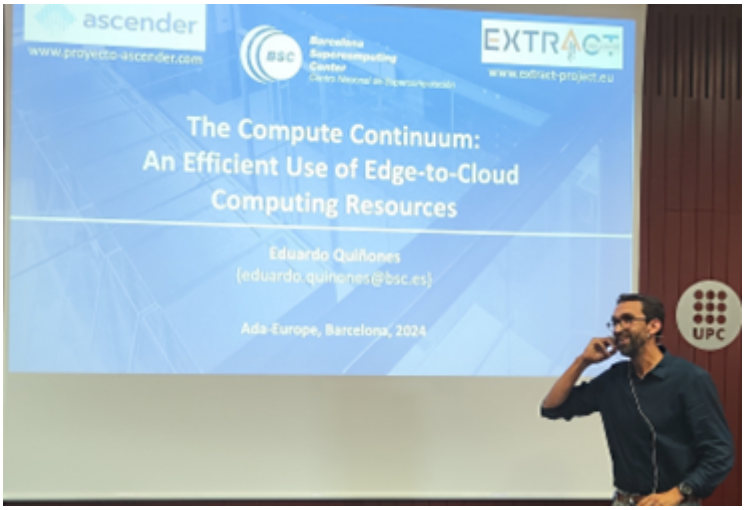
The diversity of these events helped show the application of EXTRACT project technology in different fields and at different levels of understanding. To help ensure the broad reach of the results presented at these events, partners were asked to

inform the WP5 dissemination team of upcoming events so that mini campaigns could be launched to highlight the consortium's role in these events. News pieces summarizing the event are published afterward with a link to the presentations. The events and their details can be found publicly on the project website and internally in the dissemination register (Annex 1).

4.1.1 Keynotes

Additionally, two keynote talks and three final events were carried out to target specific audiences who would be most interested in project results and their use. In the first period of the project, two keynote talks set the stage for establishing credibility and interest in the EXTRACT project among specialized audiences. EXTRACT project coordinator, Eduardo Quiñones delivered two key notes that provided an in-depth and inspiring look at parallel programming, modeling real-time systems, and runtime optimization. His first talk, "Task-based Parallel Programming Models: The Convergence of High-Performance and Edge Computing Domains", took place at the Adaptive Machine Learning at the Network Edge Summer School 2023 directed to PhDs and early-career researchers. The second talk "The Compute Continuum: An Efficient Use of Edge-to-Cloud Computing Resources" was given at the 28th Ada-Europe International Conference on Reliable Software Technologies, an industrial focused conference on software engineering for safety-critical domains. These talks served to generate interest and visibility among high-profile technical audiences and to invite engagement from academic and industrial players in related fields.

Figure 16: Eduard Quiñones delivering the keynote speech at the 28th Ada-Europe International Conference on Reliable Software Technologies



4.1.2 Booths

EXTRACT organized booths in four high-visibility events: the Hardware.io hardware and security conference in 2024, organized by BINARE in 2024, a shared booth at the Smart City Expo World Congress 2024, where the Barcelona Supercomputing Center had a large presence and the project held a side event, and two editions of the European Big Data Value Forum (20244/2025). These booths allowed for a cross-sector presence that reflected the project's goal to reach out beyond the data and AI community but also to industrial

stakeholders and urban innovation actors who could benefit from the project results. Each booth provided an opportunity to position EXTRACT´s results within the specific concerns and priorities of the audiences. As a result, these participations broadened the project's visibility, facilitated networking and contributed to community building.

4.2.3 Final events

In the initial communication and dissemination plan, a final event with more than 50 attendees was envisioned to help share final project results. However, following the interest and opportunities created by the use case leaders, it was decided that instead, two final events would be held to share project results directly with final end users of the project in the crisis management and astronomy communities. Moreover, it was decided that the 2025 European Big Data Value Forum would be leveraged to reach the larger European data community.

These events took place in the last months of the project with the PER validation final event co-located in the policy and general public focused event of the "First Three Days" in Venice and the TASKA event co-located within the 11th Annual Science at Low Frequencies Conference (SALF XI).

In the first PER validation event on 03 October 2025, co-located in a policy-oriented event about crisis management (The First Three Days), the project's end-to-end workflow and mobile application were tested with a group of volunteer retired security officers. These volunteers participated in a live evacuation drill, guided by EXTRACT technology. Key feedback was gathered and proved instrumental in refining the application and paving the way for a second validation, that could be carried out thanks to the project amendment. The second validation was held on 03 February 2026, and represented a refinement of the closing of the EXTRACT technology loop and an updated application interface. An internal training on the integration of the technologies was run on 02 February and the testing of the technology with a group of twenty retired police officers, was held on 03 February. The technology worked on the ground and volunteers were pleased with the app refinement. The test did not go completely as planned, but many important results were gathered and incorporated into refinements in the last month of the project. These final events gathered approximately 80 participants each.

Figure 17: MATHEMA instructing Venetian stakeholders how to use the app and getting feedback once they arrived at the safe spot

In the TASKA final event, a live demonstration of TASKA, delivered by partners from OBSPARIS showcased the project's innovative extreme-data mining workflow for automated data analysis, orchestration and detection. Around 40 participants attended the session with several follow-up discussions occurring after the event, including an invitation to participate in an upcoming addition of the SKA regional network meeting.

Figure 18 Members of OBSPARIS at the SALF XI Conference

Success criteria for this activity had been set at 2 keynote talks, 1 industrial booth and 1 final event with more than 50 attendees. At M39, two keynote talks had been delivered, the project had participated in three industrial booths and four events sharing final project results were held with a total of 200 participants. Moreover, the consortium participated in 67 events throughout the lifetime of the project. Our initial KPIs were exceeded.

4.2 Publications

The project's 14 peer-reviewed publications allowed the consortium to disseminate its findings and ensure that EXTRACT technical results were openly and widely accessible to relevant research communities. The publications focused on key project technologies and grew significantly since M19, reflecting the increasing maturity of project results over time. In additionally to the 14 open publications, four are in the pipeline. The publications from M1-M39 are listed in Table 1, with additional details available in Annex 2.

Several of the publications appear in the EXTRACT Zenodo community, however, some are openly available in partner's institutional repositories or other accepted

repositories. The EXTRACT Zenodo community was established at M6 to facilitate access to project publications and presentations and ensure they would be available beyond the end of the project.

Table 1: EXTRACT publications in M1-M39

Num	Publication type	Title	Author	Venue
1	Conference proceedings	Multi-party Computation for Privacy and Security in Machine Learning: a Practical Review	Bellini, Alessandro; Bellini,Emanuele; Bertini, Massimo; Almhaithawi, Doaa; Cuomo, Stefano	2023 IEEE International Conference on Cyber Security and Resilience
2	Conference proceedings	Is Performance of Object Storage Predictable for Serverless I/O Workloads? A Comparative Study	G. Eizaguirre and M. Sanchez-Artigas	Cloud-Edge Continuum Workshop 2023 (IEEE 31st International Conference on Network Protocols
3	Conference proceedings	Scaling a Variant Calling Genomics Pipeline with FaaS	Aitor Arjona, Arnau Gabriel-Atienza, Sara Lanuza-Orna, Xavier Roca-Canals, Ayman Bourramouss, Tyler K. Chafin, Lucio Marcello, Paolo Ribeca, and Pedro García-López	9th International Workshop on Serverless Computing
4	Conference proceedings	Glider: Serverless Ephemeral Stateful Near-Data Computation	Daniel Barcelona-Pons, Pedro García-López, and Bernard Metzler.	Proceedings of the 24th International Middleware Conference (Middleware '23)
5	Journal	MLLess: Achieving cost efficiency in serverless machine learning training	Pablo Gimeno Sarroca, Marc Sánchez-Artigas	Journal of Parallel and Distributed Computing
6	Journal	Semantic Segmentation of Solar Radio Spikes at Low Frequencies	Pearse C Murphy, Stéphane Aicardi, Baptiste Cecconi, Carine Briand, Thibault Peccoux	The Open Journal of Astrophysics
7	Conference proceedings	Dataplug: Unlocking extreme data analytics with on-the-fly dynamic partitioning of unstructured data	Aitor Arjona, Pedro García-López, Daniel Barcelona-Pons	2024 IEEE 24th International Symposium on Cluster, Cloud and Internet Computing (CCGrid)

8	Conference proceedings	Serverful Functions: Leveraging Servers in Complex Serverless Workflows (industry track)	Germán T. Eizaguirre, Daniel Barcelona-Pons, Aitor Arjona, Gil Vernik, Pedro García-López, Theodore Alexandrov	Proceedings of the 25th International Middleware Conference Industrial Track
9	Article in journal	A Seer knows best: Auto-tuned object storage shuffling for serverless analytics	Germán T. Eizaguirre; Marc Sánchez-Artigas	Journal of Parallel and Distributed Computing
10	Conference proceedings/workshop	Exploring Secure and Efficient Temporary Data Sharing between co-located Kubernetes Containers	Aitor Arjona, Bernard Metzler, Pascal Spörri	2024 IEEE 32nd International Conference on Network Protocols (ICNP)
11	Conference proceedings/workshop	Manifesting the Elasticity of Serverless Data Pipelines: a Metabolomics Use Case	Díez-Cuadrado, E., Eizaguirre, G. T., & Molina-Giménez, E.	Proceedings of the 11th International Workshop on Serverless Computing
12	Conference proceedings/workshop	Flexecutor: Out-of-the-Box Smart Provisioning for Serverless Workflows	Molina-Giménez, E., Barcelona-Pons, D., Iacoponelli, O. H., & García-López, P.	_Proceedings of the 11th International Workshop on Serverless Computing
13	Conference proceedings/workshop	Burst Computing: Quick, Sudden, Massively Parallel Processing on Serverless Resources	Daniel Barcelona-Pons, Aitor Arjona, Pedro García-López, Enrique Molina-Giménez, Stepan Klymonchuk	2025 USENIX Annual Technical Conference
14	Conference proceedings/workshop	Serverless Data Analytics (Finally) Bridging the Gap: Introducing the ORTZI DATAFRAME	Germán T. Eizaguirre, Marc Hostau, Marc Sánchez-Artigas	The IEEE International Conference on Cloud Computing CLOUD 2025

Success criteria for this activity was twelve publications. At M39, the project has 14 publications and five accepted ones. The KPI for this activity has been met.

4.3 Dissemination support from HRB

The Horizon Results Booster (HRB) service, an expert service offered by the European Commission to maximise the impact of Horizon Europe projects, was consulted to help refine the project's dissemination strategy and to help identify

points of improvement. The HRB conducted an analysis of the project's communication and dissemination strategy and gave feedback and recommendations to move forward. Overall, the experts confirmed that the project's strategy was well-structured and effectively executed. Specifically, target audiences were clearly identified, key messages and channels were tailored to each stakeholder group and the link between results, stakeholders and communication outputs was clear. The project scored a 3.8 readiness score, which is above the 2.9 benchmark for comparable research and innovation action projects at a similar stage.

The HRB gave positive feedback on the range and consistency of the project's communication and dissemination activities, highlighting the diversity of audio-visual content, regular organisation of events, the mix and maintenance of communication channels and active engagement with cluster sibling projects. These findings validated the dissemination team's approach and confirmed that the dissemination activities were on track.

The HRB experts also provided recommendations for optimising social media engagement, strengthening the narrative around the TASKA use case to make it more tangible and visible, focusing on impact for non-technical audiences and leveraging platforms such as the Horizon Results Platform for additional visibility. The dissemination team were grateful for these recommendations and incorporated many of them into the final stretch of the project, specifically, tagging partners in social media posts (social media engagement ended with a high average of around 20%), creating more materials focusing on TASKA and its benefit (including a HRB-created video on this case and its importance for scientific advancement), and adding the video to the Horizon Results Platform. The project benefitted from this external insight and recommendations, using them as a valuable reflection on how to fine-tune the dissemination strategy and ensure that results reach the widest and most relevant audiences.

5. Final Report on Community Building and Training Activities

Community building and training have been central to driving uptake and ensuring EXTRACT technologies live on beyond the project. Community building included a hands-on approach, sharing the technology across a wide range of platforms, networks and communities to test its relevance, demonstrate its value and build lasting impact.

Community engagement was ramped up in the second phase of the project, building on the analysis of our target audiences and consortium participation to key events that helped attract interested parties and potential users. Collaborative platforms such as the ADR Awareness Centre, the AIonDemand Platform and networks, including the Big Data Value Association, the European Open Science Cloud ecosystem, HiPEAC and the HORIZON-CL4-2002 DATA-01-05 (DATANEXUS) cluster served to share project technology and refine the dialogue with the wider community. Training opportunities, delivered internally and externally, ensured the transfer of knowledge and tools for using EXTRACT technology.

5.1 Community building

The communication and dissemination channels described in Chapters 3 and 4 were effectively used to build synergies and collaboration with the greater European AI, Data, industrial, academic and Horizon Europe communities, in addition to specific field-specific communities (astronomy, smart cities, security, crisis management). In all, twelve participations in EU AI and Data community events took place, eight of them in M19-M39.

One of the main communities created and developed during the project was the DataNexus cluster of projects, based on the similarities of the project objectives shared by the projects funded after the same HORIZON-CL4-2002 DATA-01-05 call. In the first project period, these projects were beginning to establish this community and by M39 several activities had been carried together, including a press release announcing the community, a video explaining the cluster and its objectives (created with the help of HRB), various social media campaigns announcing these materials, joint sessions at the 2024 and 2025 EBDVF sessions, including one joint booth. Joint participation at the ADRA/AI Future Ready event and DataWeek 2025. Four workshops were organized by the cluster at key community events.

Figure 6: Highlights from EBDVF and Future Ready event participations

Other key community initiatives included the contributions to the AIonDemand Platform with project profiles, cases and publications. This platform and participation in activities and events where it participates have helped ensure the results of EU projects live on beyond projects and facilitate reuse. Presentations at other relevant technological communities include three collaborations with HiPEAC, a hub uniting researchers and institutions to advance computing and foster collaboration, including two workshops, including participations with other EU-funded projects and contributions to its sectorial magazine to share novel approaches and solutions targeting key aspects of the compute continuum design. The project also joined the annual EGI conference discussions on advancing research and innovation in data-intensive processing by sharing how EXTRACT tackles the growing demands of data processing by unifying edge, cloud, and HPC in a seamless *compute continuum*.

TASKA use case owners have also increasingly shared their results with the EOSC community. In November 2025, members of the TASKA use case shared a key innovation in radio astronomy data processing during a presentation to the EOSC

virtual research environments expert group. Two demos were presented demonstrating advanced capabilities for processing radio astronomy data across multiple cloud infrastructures. Following the TASKA final event, members of the TASKA use case were also invited to present the project's interactive workflows at the EOSC Winter School, which provided the opportunity to show the platform's readiness for EOSC Node deployments.

PER use case owners also integrated themselves into policy making and crisis management communities with participation in GovTech and the EU Science for Preparedness Event. Additionally, emphasis was placed on engaging potential end users through three citizen engagement events, including the two final validation. These efforts worked to facilitate endorsement of project technologies and highlight the value of EU funding for societal challenges.

These community activities have helped to establish lasting relationships that the technology developed in the project will continue and will influence future work in these communities.

5.2 Training

Training that demonstrates how project tools work and their potential applications is an important way to go beyond community building to creating the next generation of EXTRACT technology users. For this activity, two training types have been identified: internal and external training.

5.2.1 Internal Training

Internal training took place periodically throughout the project to ensure that the partners, who span the full compute continuum, not only understood each other's technologies, but actively integrated them and user requirements. Regular weekly technical meetings helped keep the project on track while six dedicated trainings held within the project face to face meetings helped build shared understanding.

The first three trainings were held during the first project period and included two trainings held at the site of the project use cases (Venice, Italy and Nançay, France) and a two-day Hackathon, hosted in Barcelona.

The second three trainings were held in Florence in 2025, BSC in 2025 and as part of the final PER validation in 2026. The Florence training focused on providing a hands-on session focused on integration challenges and technical synergies. The BSC training allowed use case owners to delve deeper into the technologies being developed by the project to apply it to different bottlenecks. The training in Venice was focused on ensuring understanding and documenting the steps needed to use the different technological components in the PER extreme mining workflow.

Figure 17: Internal workshop held between IKERLAN and BSC in Florence 2025

5.2.2 External training

External training sought to better prepare end users and members of the aforementioned communities with the knowledge and tools needed to use project technology for their uses and on other platforms. By project end, two tutorials, three workshops and one external Hackathon were held in the framework of the project.

Tutorials

Tutorials were given to mobility experts (M6) and members of the wider EOSC community (M37) interested in deepening collaboration across the wider EOSC community. These training opportunities held potential users with specific needs to better understand the results offered by the project and how to access and implement them.

Workshops

Three workshops were organized within the framework of the project's training activities. EXTRACT participants participated in the Compute Continuum workshop in the 2024 HiPEAC conference, in addition to organizing a workshop in the 2025 HiPEAC workshop with members of related EU projects, namely Verge, NEARDATA, CLOUDSKIN. Additionally, partner BSC included hands-on EXTRACT training on the compute continuum within the 2025 multi-disciplinary Elixirion training focused on emerging Healthcare 4.0 paradigm by leveraging 6G technologies.

Figure 18: Advancing the compute continuum workshop at HiPEAC 2025 and workshop within ELIXIRION MSCA training event

Hackathon

In November 2024, an innovative Hackathon was hosted by BINARE to help drive innovation and business thinking as part of the Nordic Economic Students Union (NESU) Autumn Conference. This Hackathon brought business students together to help them approach real-world opportunities for European-developed technologies.

Success criteria for the training activities were envisioned as three internal trainings, including one Hackathon, one workshop, one tutorial and one Hackathon. The project exceeded this goal with six internal trainings, including a Hackathon and three workshops, two tutorials and one external Hackathon.

6. Final Exploitation Plan and Report

This section presents the final status of the EXTRACT project's exploitation activities over the full project duration (M1–M39). It builds upon the initial exploitation plan developed in D5.1 *Dissemination, exploitation and community building report* (M6) and its subsequent updates, reflecting the evolution from planning to implementation and validation of exploitation strategies.

The EXTRACT project, funded under the Horizon Europe programme, has progressively developed a structured and results-driven approach to exploitation. Originally planned for 36 months and extended to 39 months, the project leveraged the additional period to further mature its Key Exploitable Results (KERs), validate their market relevance, and strengthen pathways towards post-project uptake.

Within the Horizon framework, exploitation refers to the utilisation of project results beyond the project lifetime. This includes their application in further research activities, the development and commercialisation of products and processes, the provision of services, and contributions to standardisation activities.

Throughout the project lifecycle, exploitation activities have encompassed:

- the integration of results into ongoing and future research initiatives,
- the development of innovative products and services,
- engagement with stakeholders and potential adopters,
- and preparation for commercialisation and standardisation.

The primary objective of the exploitation plan has been to maximise the impact of EXTRACT results across scientific, societal, and economic domains. This objective has been operationalised through coordinated partner efforts in innovation management, intellectual property protection, data management, and stakeholder engagement.

Project partners have played a central role in driving exploitation, both individually and collectively. Their activities have included the use of results in follow-up projects, academic outputs (e.g., PhD and post-doctoral research), and the preparation of market-oriented solutions. In addition, mechanisms such as licensing, open access strategies, and collaboration with external stakeholders have been actively considered and, where relevant, implemented.

6.1 Key Exploitable Results (KERs) Identification

A central component of the EXTRACT exploitation strategy has been the systematic identification, monitoring, and refinement of the project's Key Exploitable Results (KERs). KERs represent the most valuable outputs of the project with the highest potential for further development, uptake, and exploitation beyond the project duration.

Throughout the project lifecycle, the consortium has continuously assessed project outcomes in order to identify results with strong scientific, technological, and market

potential. This process has involved evaluating technological innovations, methodological advances, and developed frameworks against criteria such as innovation level, maturity, potential user base, and alignment with emerging market needs.

Within EXTRACT, the identified KERs include advancements in areas such as:

- enhanced data infrastructures,
- novel Machine Reinforcement/Learning AI (MRLAI) and big-data frameworks,
- data-driven orchestration mechanisms,
- and integrated computing technologies designed to address extreme data challenges.

The identification and refinement of these KERs have been carried out collaboratively across the consortium, ensuring that each result is associated with responsible partners and clear exploitation pathways.

An initial list of KERs was defined during the early stages of the project and subsequently refined as the project progressed and technologies matured. The final set of KERs reflects both the technological achievements of the project and their potential for further research, industrial application, and commercial development.

It should be noted that exploitation strategies remain inherently dynamic. Market conditions, technological evolution, and stakeholder needs may influence the long-term development and uptake of specific results even after the completion of the project. Nevertheless, the KERs presented in this report represent the consortium's consolidated view of the most promising exploitable outcomes generated within EXTRACT.

Additional information regarding datasets and data governance associated with several KERs can be found in the latest version of D6.2 Data Management Plan.

The final list of EXTRACT Key Exploitable Results and their corresponding value propositions is presented in Table 1.

Table 1: EXTRACT KERs and their corresponding value.

Partner	**KER name**	**KER type (data, software, hardware, service documents, other IPR)**	**TRL (M18)**	**TRL at the end of the project**	**MRL end of the project**	**KER Description and added value for the project**
BSC	MARL	software	TRL3	TRL6	MRL4	Multi-Agent Reinforcement Learning is a system where multiple agents learn and make decisions together in a shared environment. By interacting with each other and the environment, they solve complex problems, making it great for optimizing distributed tasks. In EXTRACT, the MARL is used to train the agents to

						find the best escape route in the PER use case.
BSC	COMPSs	software	TRL5	TRL6	MRL4	COMPSs is a programming model that boosts the performance of large-scale applications by automatically running tasks in parallel. It makes it easy to scale applications across different environments like cloud, edge, and HPC without needing manual intervention.
IKERLAN	K8s Monitoring tool	software	TRL2	TRL4	N/A	Kubernetes is an open-source tool that automates how we deploy, scale, and manage containerized apps. It handles tasks like scaling, load balancing, and ensuring apps run consistently across environments.
URV	Lithops	software	TRL6	TRL7	N/A	Lithops is a serverless framework that splits large datasets and runs tasks in parallel across different cloud providers. It makes handling data-heavy workloads easier by automatically scaling resources based on need.
URV	Dataplug	software	TRL4	TRL6	N/A	DataPlug facilitates seamless data integration between different systems and platforms. It acts as a bridge that connects disparate data sources, ensuring smooth data flow and compatibility. This tool is particularly useful in environments where multiple data formats, sources, or services need to interact, helping to automate the process of data extraction, transformation, and loading (ETL) while maintaining data integrity across the compute continuum.
OBSPARIS	TASKA-A	other	TRL5	TRL7	N/A	TASKA-A allows us to deploy real time advanced processing pipelines (including ML) directly in the instrument (edge) and provides observers with various types of physical events detected in the real time data flow.
OBSPARIS	TASKA-C	other	TRL1	TRL6	N/A	TASKA-C is a workflow framework for radio astronomy interferometric imaging. It uses EXTRACT libraries and software to optimize, catalog and orchestrate the workflow, on cloud and HPC computing frameworks.
OBSPARIS	TASKA-D	software	TRL1	TRL4	N/A	The dynamic imager software will implement interferometric imaging capabilities (currently only static imaging is available), enabling video reconstruction techniques. TASKA-D is integrated as a step in TASKA-C.
LOGOS	UDT	software	TRL3	TRL4	MRL2	An Urban Digital Twin is a virtual model of a city that mirrors its real-world counterpart. It collects data from various sources, like sensors, traffic systems, and satellites, to simulate and analyse city operations in real-time. This allows for better urban planning, emergency response, and infrastructure management by providing insights and predictions based on both current and historical data.
Venezia	USE case PER	dataset	N/A	N/A	N/A	The PER use case is focused on providing real-time, personalized evacuation routes for citizens in Venice during emergencies. By utilizing advanced data collection and processing technologies, such as a UDT, the system can simulate evacuation scenarios

						and optimize safe routes. This ensures more efficient and coordinated evacuations. A MARL dynamically calculates and updates routes based on real-time conditions, while a Simulator helps test strategies under different emergency situations. Together, these technologies improve crisis management and citizen safety in complex urban environments.
SIXSQ	Nuvla data management features	software	TRL2	TRL6	MRL4	Nuvla.io is a cloud-native platform which allows remote management of edge devices and applications. It simplifies how we manage and deploy containerized applications across edge locations, making remote management and scaling more efficient. Nuvla.io and NuvlaEdge are used in the EXTRACT project for data management features, namely the data catalogue. In addition to this existing functionality, SIX is developing from scratch new features on top of the data catalogue, to enhance edge-to-cloud data management.
BINARE	Compute Continuum Security Analysis	service	TRL1	TRL3	MRL2	This planned component is a consulting + training service type of KER planned by BINARE. It leverages the BINARE Cybersecurity Toolkit for Compute Continuum Software (see other BINARE KERs) and couples it with deep-tech know-how within BINARE in both areas (cybersecurity x compute continuum such as EXTRACT) in order to help both research and industrial complex compute continuum projects and deployments receive end-to-end guidance, training and expert assessment to ensure those projects are following "secure by default" paradigm and methodologies.
BINARE	Compute Continuum Security Toolset BISETO	software	TRL2	TRL3	MRL3	This planned KER envisions to include an essential set of tools, safeguards and checks/procedures needed to secure compute continuum setups and ensure basic and beyond security for a Compute Continuum Site (CCS) deployment. It covers major cybersecurity areas such as from detecting/preventing data tampering and securing machine learning models integrity to secure versioning, and from discovering vulnerabilities from code and containerized deployments to attempting defence against potential/known attacks (HTTPS, man-in-the-middle, credentials stuffing/spraying, etc.), thus ensuring the overall system remains safe and trustworthy.
MATH	MPC	software	TRL2	TRL4	MRL3	Multi-Party Computation is a cryptographic method that allows multiple parties to compute a result from their combined data without revealing their individual inputs to one another. It ensures privacy while allowing collaboration, and it is often used in scenarios where sensitive data from different entities needs to be processed securely.
MATH	Simulator	software	TRL1	TRL5	MRL3	The simulator is the engine that drives the movement and actions of virtual people within the Urban Digital Twin. It replicates

						real-world behaviours to help us study and optimize urban scenarios. It is also helpful to simulate a large amount of scenarios in order to train our MARL.
MATH	DataFusion	software	TRL1	TRL4	MRL3	DataFusion merges data from multiple sources into one cohesive view. It combines, cleans, and integrates datasets, allowing us to make sense of diverse data streams for analysis and decision-making.
IBM	Sky Store	software	TRL5	TRL6	MRL2	Sky Store is a distributed object storage system used for saving large datasets. It provides scalable, secure storage that's ideal for cloud or hybrid environments where we need to store and retrieve large amounts of data.

In addition to the core set of Key Exploitable Results identified during the early stages of the project, the consortium has continued to monitor emerging technological outcomes throughout the project lifecycle. As the research and development activities progressed, several additional results were identified as having potential for further exploitation and development.

These additional KERs reflect technological components, datasets, and software solutions developed within the project that may support future research, innovation activities, or industrial uptake. While some of these results remain at relatively early maturity levels, they represent promising building blocks that can be further developed beyond the project duration.

The Table 2 summarises additional EXTRACT KERs identified during the project, together with their estimated Technology Readiness Levels (TRL) and, where relevant, Market Readiness Levels (MRL) achieved or targeted by the end of the project. These estimations reflect the consortium's final assessment of the maturity and potential exploitation pathways of the respective results.

Table 2: Additional EXTRACT Key Exploitable Results identified during the project and their estimated maturity levels

Partner	**KER name**	**KER type (data, software, hardware, service documents, other IPR)**	**TRL end of the project**	**MRL end of the project**
BSC	MARL	software	TRL6	-
BSC	COMPSs-k8	software	TRL6	-
BSC	Scheduling Solution	software	TRL6	-
BSC	HPC-PER Infrastructure	hardware	TRL6	-
IKERLAN	K8s Monitoring tool	Software	TRL4	
URV				
Venezia	USE case PER	dataset	N/A	N/A
IBM	Advanced Model Serving on K8s	Software	TRL3	MRL1
MATH	Data Security and Privacy Solutions	Software	TRL6	TBD

6.2 Partner Exploitation Strategies

BSC

Barcelona Supercomputing Center has focused on advancing research aimed at integrating High-Performance Computing (HPC) technologies within the compute continuum, with particular emphasis on supporting extreme data processing. Within EXTRACT, BSC has enhanced its HPC frameworks by introducing connectors that facilitate the development of advanced data-mining workflows. These developments significantly expand the applicability of BSC technologies across diverse data-intensive use cases.

An important component of BSC's exploitation strategy involves integrating the outcomes of EXTRACT with results from other European research initiatives in which BSC participates, including CLASS Project, ELASTIC Project, and eFlows4HPC. Through this cross-project knowledge integration, BSC aims to contribute to the development of comprehensive continuum-computing solutions capable of addressing the challenges of extreme-scale data environments.

IKERLAN

IKERLAN has developed the Extract Monitoring Module, an Edge-to-Cloud Continuum Monitoring Platform. The platform builds upon IKERLAN's prior expertise and extends it with new capabilities for monitoring computing infrastructure parameters such as CPU utilisation, power consumption, and memory usage across distributed environments.

The developed platform contributes to the creation of a distributed data infrastructure capable of securely collecting operational information across the computing continuum. From an exploitation perspective, the solution is intended to support industrial customers in automating system monitoring and optimisation processes while improving orchestration and deployment efficiency.

Potential industrial users identified during the project include organisations such as LKS Next, SixSq, Lantek, Mondragon Assembly, CAF, and other companies within the Mondragon industrial ecosystem. Protection of the developed software is foreseen through software registration and copyright mechanisms. The solution is expected to continue evolving beyond the project duration as part of IKERLAN's industrial analytics platform portfolio.

LOGOS Research Institute (LOGOS, LRI)

Link Campus University Research Institute contributes to EXTRACT with research on urban complexity as a Dynamic System of Socio-Technical Systems. The organisation aims to leverage project results to strengthen its position in the fields of urban security, safety, and resilience at both national and European levels.

Within the project, LRI has advanced the development of an Urban Digital Twin prototype, which can be integrated into Smart City control room infrastructures used

by municipalities. The organisation intends to further develop this capability through collaboration with local authorities, urban planners, and research institutions.

LRI follows an innovation approach inspired by the Open Innovation 2.0 and the Quintuple Helix Model, promoting collaboration between academia, industry, government, civil society, and environmental stakeholders. Exploitation activities focus on establishing partnerships with municipalities, integrating open data platforms, and leveraging regional innovation programmes such as European structural funds to support further development of solutions related to the PER use case.

URV

Researchers from Universitat Rovira i Virgili have contributed to EXTRACT through the Cloud and Distributed Systems Research Lab, which maintains strong collaborations with international research institutions and industrial partners.

The group coordinates the open-source mobility initiative CloudStars and collaborates with organisations such as the UK Health Security Agency and the European Molecular Biology Laboratory in the field of bioinformatics.

URV's exploitation strategy focuses primarily on open-source dissemination of the developed software components under widely adopted licenses such as Apache-2.0, enabling both academic and commercial use. In addition, the research group plans to leverage the project outcomes in follow-up research activities and considers the potential creation of a start-up based on cloud analytics technologies developed within the project.

IBM

IBM has focused its exploitation strategy on two technological assets developed or enhanced within the EXTRACT project.

The first asset concerns an advanced model serving framework, which may influence future versions of IBM's AI model serving offerings deployed in both public cloud and private data-centre environments.

The second asset relates to Sky Store, a research prototype originally developed in collaboration with University of California Berkeley. Sky Store functions as a cross-cloud global object storage backbone that enables efficient data transfer and policy-based caching across distributed infrastructures. The insights gained within EXTRACT are expected to inform future IBM cross-cloud solutions.

BINARE

Binare Oy, a deep-tech cybersecurity spin-off from the University of Jyväskylä, has developed an exploitation strategy focused on security aspects of Compute Continuum infrastructures.

The first exploitation direction involves strengthening specialised expertise in security architectures, risk assessment, and best practices for compute continuum environments. This knowledge is expected to be delivered through expert services

and specialised consultancy offerings for organisations deploying distributed computing infrastructures.

The second direction focuses on the potential development of software tools supporting security assessment, configuration verification, and compliance validation for compute continuum deployments. These tools may assist organisations in evaluating and improving the security posture of their infrastructures.

MATH

Mathema SRL is an ICT company specialising in the development of complex digital systems supporting large organisations in their digital transformation initiatives.

Within EXTRACT, Mathema has focused on solutions related to data security and privacy, including the development of a Python-based framework integrating advanced techniques such as:

- Differential Privacy
- Fully Homomorphic Encryption
- Secure Multi-Party Computation

These technologies are particularly relevant for organisations exchanging large volumes of sensitive data and requiring strong privacy and integrity guarantees.

SIXSQ

SixSq develops edge-to-cloud infrastructure management solutions based on its Nuvla technology platform.

The platform consists of two main components:

- Nuvla.io, a Software-as-a-Service platform that manages and orchestrates distributed edge infrastructures
- NuvlaEdge, a containerised software component deployed on edge devices.

Within EXTRACT, these technologies have been utilised for data management functionalities, including the implementation of the project's data catalogue. Although the core platform is already mature (TRL9), additional features developed within EXTRACT aim to enhance data management capabilities across the edge-to-cloud continuum. These capabilities may be particularly valuable for organisations managing large datasets, including institutions such as the European Space Agency.

Venezia

The Personalised Evacuation Routing (PER) System, developed by the Metropolitan City of Venice, represents one of the key application-level outcomes of the EXTRACT project.

The system supports improved public safety during emergency scenarios by providing dynamic evacuation routing based on real-time data and advanced analytics. Beyond its application within the project use case, the PER system has

potential for broader adoption in urban mobility and emergency management contexts.

Future exploitation opportunities include collaboration with regional innovation actors such as Veneto Innovazione and partnerships with mobility analytics companies such as MyTraffic. These collaborations could support the integration of the PER system into broader traffic management solutions used by transportation authorities and urban planners.

6.3 Exploration of Funding and Exploitation Opportunities for KERs

As the EXTRACT project approached its final phase, particular attention was devoted to identifying mechanisms that could support the continued development and exploitation of the project's Key Exploitable Results (KERs) beyond the project lifetime.

A key component of this effort involved mapping relevant funding instruments and innovation support mechanisms at both the European and national levels. These opportunities may support further research, technology maturation, industrial uptake, and potential commercialization of the developed solutions.

The consortium analysed programmes and initiatives provided by European funding frameworks as well as national and regional innovation agencies. This mapping activity aimed to identify suitable instruments that could facilitate follow-up activities such as:

- continuation of research and innovation efforts,
- scaling and validation of developed technologies,
- creation of new collaborations with industrial stakeholders,
- and potential commercialization pathways for selected KERs.

By aligning the project's technological outcomes with existing funding and innovation programmes, the consortium seeks to ensure the sustainability of EXTRACT results and support their transition toward real-world deployment. This approach enables partners to explore complementary funding sources and strategic partnerships capable of extending the impact of the project outcomes.

The following tables present relevant funding opportunities identified across the countries represented in the consortium, highlighting programmes that may support further development, deployment, or commercialization of EXTRACT technologies.

The information on relevant funding opportunities across the countries represented in the consortium is presented in the previous deliverable D5.2.

Throughout the project lifecycle (M1–M39), consortium partners explored a range of potential instruments supporting the exploitation and commercialization of the project's Key Exploitable Results (KERs). This process involved analysing relevant funding programmes, innovation support mechanisms, and commercialization pathways that could facilitate the further development and deployment of project outcomes.

Each partner assessed the relevance of the identified instruments with respect to its specific technological assets, organisational strategy, and target markets. This analysis supported the identification of potential pathways for both joint and partner-specific exploitation activities, including follow-up research initiatives, technology transfer, and industrial collaborations.

The exploration of these instruments provided a structured foundation for the development of post-project exploitation strategies and helped partners prepare for potential future initiatives aimed at scaling and commercialising EXTRACT technologies.

6.4 Market Analysis and Innovation Gap

6.4.1 Market Context

The rapid growth of connected systems and data-driven applications is creating increasing demand for scalable computing architectures capable of processing data across distributed environments. The expansion of cyber-physical systems, artificial intelligence applications, and large-scale Internet of Things (IoT) infrastructures is generating unprecedented volumes of data that require advanced processing, orchestration, and security mechanisms.

According to recent industry analyses from IoT Analytics, the number of connected IoT devices worldwide surpassed 18 billion in 2024 and continues to grow steadily, with projections estimating more than 30 billion connected devices by the end of the decade. This rapid expansion significantly increases the demand for scalable data infrastructures capable of processing large volumes of distributed data generated across cloud, edge, and IoT environments.

At the same time, the increasing distribution of computing resources across cloud, edge, and device environments has led to the emergence of the computing continuum paradigm. This paradigm enables coordinated computing across multiple infrastructure layers, allowing data processing to occur closer to where data is generated while still leveraging large-scale cloud resources.

The development of this ecosystem is strongly supported by initiatives promoted by the European Commission focusing on the Cloud-Edge-IoT Continuum, which aims to strengthen Europe's technological sovereignty and enable secure and efficient data infrastructures.

Innovation Gap

Despite significant technological progress, existing computing infrastructures often struggle to efficiently manage extreme-scale data environments characterized by:

- extremely large data volumes
- high data generation velocity requiring near real-time processing
- heterogeneous and distributed data sources
- intermittent connectivity between edge and cloud environments
- incomplete or sparse datasets

Current solutions typically address these challenges independently rather than providing an integrated framework capable of managing them simultaneously.

The technologies developed within the EXTRACT project aim to address this gap by combining:

- distributed computing orchestration
- scalable big-data infrastructures
- machine learning capabilities
- flexible security mechanisms across multiple infrastructure layers.

This integrated approach enables efficient data processing across the full computing continuum.

6.4.2 Competitive Landscape

The market for distributed computing and edge data processing is evolving rapidly, with several major technology providers offering platforms that support parts of the computing continuum.

Key technology providers include:

- NVIDIA — providing AI-accelerated edge and data-center computing platforms.
- Amazon Web Services — offering edge and hybrid computing solutions such as IoT and edge processing services.
- Microsoft — developing distributed cloud and IoT platforms through Azure-based edge computing solutions.

- Arm Holdings — enabling cloud-to-edge architectures through its Neoverse computing platforms.

- Intel — providing virtualization, security, and distributed infrastructure technologies supporting continuum computing.

However, many existing solutions focus on individual components of the computing continuum, such as cloud infrastructure, edge computing hardware, or AI acceleration. Fully integrated frameworks combining extreme-scale data processing, distributed orchestration, machine learning capabilities, and multi-layer security remain relatively limited.

The technologies developed within EXTRACT therefore aim to complement the existing ecosystem by addressing the integration of these capabilities within a unified framework.

6.4.3 Competitive Advantages

The EXTRACT technological approach offers several potential advantages:

Holistic extreme-data processing
The platform addresses multiple extreme-data characteristics simultaneously, including volume, velocity, variety, and distributed data generation.

Computing continuum integration
The solution supports coordinated computing across cloud, edge, and distributed infrastructures.

Built-in security across layers
Security mechanisms can be integrated across cloud, edge, and operational environments.

Advanced data orchestration and machine learning integration
The framework supports automated data processing and AI-driven analytics across large distributed infrastructures.

Target Markets

Technologies developed within EXTRACT can potentially support multiple data-intensive sectors, including:

- smart cities and urban infrastructure

- industrial automation and manufacturing

- logistics and transportation systems

- energy systems and smart grids

- healthcare and medical imaging

- smart agriculture and environmental monitoring
- large-scale scientific computing.

These sectors are characterised by high volumes of distributed data generation and increasing demand for secure and scalable data processing infrastructures.

6.3.4 Market Segmentation

The IoT ecosystem is expanding rapidly across multiple industry sectors. Key segments include:

- manufacturing and Industry 4.0
- smart cities and infrastructure
- healthcare and medical analytics
- energy and smart grids
- logistics and transportation
- agriculture and environmental monitoring
- telecommunications and critical infrastructure.

These sectors represent major potential users of distributed computing platforms capable of processing large-scale data streams across cloud-edge-IoT environments.

6.4.5 Stakeholders (for Exploitation and Sustainability)

The potential users of EXTRACT technologies include both technical experts and organisations operating complex data infrastructures.

Target users

The primary target users of EXTRACT technologies include:

- data scientists and AI researchers
- system architects and infrastructure engineers
- digital twin developers
- industrial operators managing IoT infrastructures
- research organisations and scientific computing centres.

Key stakeholders

Internal stakeholders include research and development teams, software engineers, IT infrastructure specialists, and project management teams involved in the development and integration of EXTRACT technologies.

External stakeholders include industry partners, policymakers, regulatory bodies, research institutions, and the open-source development community that may benefit from or contribute to the further development and adoption of the project results.

To better understand the broader ecosystem and identify relevant organisations for collaboration and potential uptake of EXTRACT results, a structured stakeholder mapping exercise was conducted with the support of experts from the Horizon Results Booster (HRB) which is presented in Table 3.

The identified stakeholders will support future collaboration, knowledge exchange, and potential uptake of EXTRACT Key Exploitable Results in both research and industrial ecosystems.

Table 3:Stakeholder database.

Nr	Stakeholder group	Stakeholder name	Country	Website link	Email address (publicly available)
1	Cloud-edge infrastructure providers	Tietoevry Tech Services	Finland	https://www.tietoevrytechservices.com/en/	Michael Lenahan, Head of Data, Cloud and Security michael.lenahan@tietoevry.com
2	Cloud-edge infrastructure providers	OpenNebula	Spain	https://opennebula.io/	Alberto P. Martí, VP of Open Source Innovation https://www.linkedin.com/in/albertomarti/
3	Cloud-edge infrastructure providers	European Data Centre Association - EUDCA	EU	https://www.eudca.org/	Lex Coors, President of EUDCA & Policy Committee Chair, Chief Data Center Technology and Engineering Officer, Digital Realty https://www.linkedin.com/in/lex-coors-6356831/?originalSubdomain=nl
4	Cloud-edge infrastructure providers	Gcore	Luxemburg	https://gcore.com/	Pavel Solovev, Vice President of IT, Data Management, and Security https://lu.linkedin.com/in/pavel-solovev-238427a7
5	Cloud-edge infrastructure providers	SUSE	Luxembourg	https://www.suse.com/	Thomas Di Giacomo Chief Technology & Product Officer Tdigiacomo@suse.com
6	IoT/ML/AI research communities	AIOTI – Alliance for AI, IoT and Edge Continuum Innovation	Belgium	https://aioti.eu/	Damir Filipovic, Secretary General sg@aioti.eu
7	IoT/ML/AI research communities	CAIRNE – Confederation of Laboratories for Artificial Intelligence Research in Europe	Netherlands	https://cairne.eu/	Emanuela Girardi, AI, Data and Robotics Association (Adra);Pop AI emanuela.girardi@popai.me
8	IoT/ML/AI research communities	EIT AI Community	EU	https://ai.eitcommunity.eu/	Parul Chaudhary, Strategic Synergies Lead https://www.linkedin.com/in/parul-chaudhary8/
9	IoT/ML/AI research communities	AI-on-Demand Community (AI4Europe)	EU	https://www.ai4europe.eu/	Barry O'Sullivan, Project Coordinator, University College Cork (lead of AI4Europe) b.osullivan@cs.ucc.ie
10	Research Infrastructures	ELIXIR Europe	EU	https://elixir-europe.org/	Jonathan Tedds, Programme Manager – Scientific Computing, Compute Platform / Tools Platform

					jonathan.tedds@elixir-europe.org
11	Research Infrastructures	OpenIoT Research Unit (FBK)	Italy	https://openiot.fbk.eu/	Fabio Antonelli, Head of Unit fantonelli@fbk.eu
12	Research Infrastructures	D4Science	Italy	https://www.d4science.org/	Leonardo Candela, Scientific Director leonardo.candela@isti.cnr.it
13	Research Infrastructures	SovereignEdge.EU	EU	https://sovereignedge.eu/	innovation@sovereignedge.eu
14	Research Infrastructures	dAIEDGE (Network of Excellence for Edge AI)	EU	https://daiedge.eu/	Alain Pagani, Project Coordinator Alain.Pagani@dfki.de
15	Software vendors	Eurotech	Italy	https://www.eurotech.com/it/	Marco Carrer, Chief Technology Officer marco.carrer@eurotech.com
16	Software vendors	Kontron AG	Austria	https://www.kontron.com/en/	Michael Riegert, COO IoT Europe https://www.linkedin.com/in/riegertmichael/
17	Software vendors	Telit Cinterion	UK	https://www.telit.com/	Antonino Sgroi, Head of Global R&D antonino.sgroi@telit.com
18	Industrial R&D	Grafana Labs	Multinational	https://grafana.com	Torkel Ödegaard, Chief Growth Officer torkel@grafana.org
19	Industrial R&D	Honeycomb	UK	https://honeycomb.io	Charity Majors, CTO charity@honeycomb.io
20	Critical infrastructure operators	European Network of Transmission System Operators for Electricity (ENTSO-E)	Belgium	https://www.entsoe.eu/	Edwin Haesen, head of system development edwin.haesen@entsoe.eu
21	Critical infrastructure operators	European Distribution System Operators (E.DSO)	Belgium	https://www.edsoforsmartgrids.eu/	Charles Esser, Secretary General charles.esser@edsoforsmartgrids.eu
22	Critical infrastructure operators	European Network of Transmission System Operators for Gas (ENTSO-G)	Belgium	https://www.entsog.eu/	Panagiotis Panousos Director, System Operation Panagiotis.Panousos@entsog.eu

23	Critical infrastructure operators	RailNetEurope (RNE)	Austria	https://rne.eu/	Harald Reisinger, Chief Information Officer & Chief Financial Officer harald.reisinger@rne.eu
24	Critical infrastructure operators/Research organization	CRESYM	Belgium	https://cresym.eu/	Sébastien Lepy, General Manager sebastien.lepy@cresym.eu

6.5 Characterisation of Key Exploitable Results

To further analyse the exploitation potential of the identified Key Exploitable Results (KERs), the consortium developed a structured characterization of selected results, including their target markets, potential users, value propositions, competitive positioning, and possible exploitation pathways.

This characterization table serves as a practical and operational tool to systematically collect, structure, and consolidate key business, technical, and strategic information for each KER. In particular, it captures aspects such as end-user needs and challenges, market segments, early adopters, competing solutions, unique value propositions, expected use models, and intellectual property considerations. By bringing together these elements in a unified format, the table allows partners to clearly articulate how their results address real-world problems and how they can be positioned within relevant market and technology ecosystems.

The use of this structured approach supports a more market-oriented perspective across the consortium, enabling partners to move beyond purely technical descriptions of results towards concrete exploitation planning. It facilitates internal alignment among partners, improves consistency across KER descriptions, and supports informed decision-making regarding commercialization strategies, potential partnerships, and investment priorities.

In addition, the characterization provides a basis for engaging with external stakeholders, including potential customers, investors, and innovation ecosystems, by translating technical outcomes into clearly understandable value propositions. It also supports the identification of synergies between KERs and helps explore opportunities for joint or complementary exploitation strategies across the consortium.

The structure of the characterization table was developed with the support of experts from the Horizon Results Booster (HRB), whose guidance significantly contributed to refining the exploitation approach of the project. Their expertise helped ensure alignment with best practices in innovation management, market analysis, and result positioning, thereby strengthening the overall quality, consistency, and impact potential of the EXTRACT exploitation strategy.

The characterization table for the main KER EXTRACT platform is presented in Table 4. Other KERs characterisation tables are presented in Appendix 3 and they serve as a living instrument supporting the continuous refinement of exploitation strategies during and beyond the project duration.

Table 4:Characterisation of EXTRACT platform as Key Exploitable Result.

KER name: EXTRACT **Platform**	
	Input from the Beneficiary
Description	The **EXTRACT Platform** (main and foremost result of the EXTRACT project) is a modular orchestration and integration platform designed for the compute continuum (Cloud–Edge–HPC). It provides a unified environment for deploying, managing, and orchestrating distributed workloads across heterogeneous infrastructures. The platform consists of several integrated components. The **Orchestration Engine** manages workload deployment, scheduling, and resource allocation across different infrastructures. The **Monitoring and Telemetry Layer** collects real-time operational metrics from edge, cloud, and HPC nodes. The **Security and Trust Module** provides security orchestration, policy enforcement, and trust scoring mechanisms. The **Data Management Layer** supports federated data access, metadata handling, and streaming data flows. Finally, **Integration Interfaces and APIs** enable interoperability with external tools, industrial protocols, and heterogeneous devices. Together, these components form a full-stack platform that enables scalable, explainable, and sustainable distributed computing across the compute continuum.
Target market/end users	The primary target market includes research infrastructures and labs, public-sector digital initiatives, and particularly companies and organisations adopting compute continuum strategies and product roadmaps. Additional target segments include industrial operators (with main focus on Smart Cities, mobility, energy, and manufacturing sectors), as well as early industrial adopters experimenting with edge–cloud orchestration in a compute-continuum paradigm. Typical end users include cloud–edge infrastructure providers, municipalities and regional authorities, research laboratories and innovation testbeds, as well as developers building distributed systems, automation pipelines, and large-scale data-processing AI-based workflows.
End-users needs / problems	Organisations operating heterogeneous Cloud–Edge–HPC infrastructures often face fragmented orchestration environments where tools operate only within a single domain, such as cloud-only, edge-only, or HPC-only environments. Deploying distributed workflows across these infrastructures requires complex configuration and integration work. At the same time, interoperability between platforms, industrial protocols, and legacy systems remains limited, or is often limited by both commercial and technological lock-in of established players (e.g., Amazon AWS, Microsoft Azure, Google GCP). Many organisations also lack unified monitoring and telemetry across the compute continuum, while security and trust mechanisms are often implemented in separate

	systems, increasing operational complexity and compliance risks. These challenges make it difficult to deploy scalable and policy-compliant/policy-flexible distributed infrastructures aligned with European Digitalisation Strategies. As a result, deployment processes remain with large manual interventions, sub optimal approaches, and are difficult to scale. The EXTRACT Platform addresses these challenges by providing a unified orchestration and integration framework that supports the entire compute continuum in an end-to-end fashion.
Competitive advantages	The EXTRACT Platform provides full-stack orchestration capabilities across heterogeneous environments spanning edge, cloud, and HPC infrastructures. By integrating orchestration, monitoring, security/trust management, and data handling into a single modular platform, it significantly reduces operational complexity compared to fragmented point solutions. The core of the platform is released as open-source software, making it particularly suitable for immediate uptake and exploitation by research infrastructures, public-sector deployments, and experimental pilot environments. Expected performance indicators to be validated during project pilots include orchestration overhead below 20% compared to bare-metal execution, and improvements of more than 40% in deployment automation efficiency for complex distributed workflows.
Use model	The EXTRACT Platform will be released as an open-source software platform under a permissive license. It is intended to be integrated into Horizon Europe research pilots, public-sector testbeds, and academic research infrastructures. Initial adoption is expected among research institutions, municipalities running Smart City pilots, and Horizon Europe projects that require orchestration across the compute continuum.
Early Adopters	Early adopters are expected to include research infrastructures and national research computing centres, Horizon Europe and Digital Europe projects requiring distributed orchestration capabilities, municipalities deploying Smart City pilots, public-sector digital innovation hubs, and industrial R&D laboratories experimenting with edge–cloud–HPC integration. Adopters will primarily be reached through Horizon Europe clustering events, European Digital Innovation Hubs, scientific conferences on edge computing and distributed systems, open-source developer communities, and industrial networks of the project partners.
Adopters' problems/needs	Early adopters typically need to experiment with compute continuum orchestration without developing custom integration layers from scratch. They require interoperable and standards-compliant solutions aligned with EU digital sovereignty initiatives, while also seeking to reduce integration costs and deployment time in pilot environments. In addition, they require transparent orchestration mechanisms suitable for public-sector governance requirements, along with integrated cybersecurity, trust evaluation, and policy enforcement mechanisms. Flexible and modular architectures are also needed to support diverse use cases such as Smart Cities, research infrastructures, and distributed industrial systems.
Alternative solution	Currently, organisations address these challenges by combining multiple independent technologies such as Kubernetes for container orchestration, OpenStack for private cloud management, and Slurm for HPC workload scheduling. These tools are often connected through custom middleware layers designed to bridge edge, cloud, and HPC environments. In many cases, organisations rely on proprietary vendor solutions or manual configuration processes when deploying experimental infrastructures. Such approaches tend to be

	fragmented, complex to maintain, and poorly suited for unified orchestration across the compute continuum.
Unique Value Proposition	The EXTRACT Platform provides a unified orchestration framework designed specifically for the compute continuum. It integrates monitoring, telemetry, security checks, trust management, and data handling within a single modular architecture. Compared to solutions that rely on loosely integrated tools, the platform reduces integration complexity while improving transparency and automation in distributed environments. The architecture supports orchestration across edge, cloud, and HPC resources while maintaining low overhead and improved deployment automation efficiency. Its open-source nature further reduces vendor lock-in and makes it particularly attractive for public-sector digital infrastructures and research-driven innovation ecosystems.
Competitors	Potential competitors include large hyperscale cloud platforms such as Amazon Web Services (AWS), Microsoft Azure (Azure), and Google Cloud (GCP). These platforms provide mature ecosystems and highly scalable cloud infrastructures but remain largely cloud-centric and may introduce vendor lock-in, with limited native support for integrated HPC–edge orchestration. Other competing technologies include edge orchestration frameworks such as KubeEdge and Open Horizon. While these solutions provide strong edge-native orchestration capabilities, they generally lack integrated support for HPC infrastructures and do not offer unified trust and security frameworks. n the HPC domain, workload managers such as Slurm are widely used for high-performance scheduling but are not designed for integrated orchestration across edge–cloud–HPC environments.
Partners	The end-to-end development and validation of the EXTRACT Platform involve ALL project partners. These partners contribute expertise in orchestration technologies, monitoring systems, cybersecurity, data management, and industrial integration.
Timing	Adoption of the platform is expected in the medium term, with an estimated timeline of two to three years for uptake in research infrastructures, public-sector digitalisation initiatives, and Horizon Europe pilot deployments, as well as by EXTRACT partners in the upcoming European and National projects.
IP Strategy	The EXTRACT Platform will be released as open-source software under a permissive license. Intellectual property management will focus primarily on platform branding, governance models, licensing strategy, and long-term sustainability of the open-source ecosystem.

6.6 Intellectual Property Rights (IPR) Management

6.6.1 Consortium IPR Framework

The protection and management of Intellectual Property Rights (IPR) represent an essential component of the exploitation strategy of the EXTRACT project. Effective IPR management ensures that the knowledge and technological results generated during the project can be protected, valorised, and transferred to industry, research communities, and society.

In accordance with the Horizon Europe Model Grant Agreement, ownership of project results belongs to the partner that generates them, unless otherwise agreed in cases of joint ownership. The management of background and foreground intellectual property within the EXTRACT consortium is governed by the project Consortium Agreement, which defines access rights, ownership rules, and procedures for the protection, dissemination, and exploitation of project results.

The consortium continuously monitors potential intellectual property arising from project activities. Depending on the nature of the Key Exploitable Results (KERs), different protection mechanisms may be considered, including patents, copyrights, trade secrets, software licensing, or open-source strategies.

6.6.2 IPR and Exploitation Contact Point Identification

As part of its exploitation strategy, EXTRACT has established a structured approach for the identification, management, and protection of Intellectual Property Rights (IPR), alongside the designation of dedicated Exploitation Contact Points within each partner organisation.

Throughout the project, the identification of potential intellectual property has been carried out in a systematic manner, covering patentable technologies, advanced algorithms, software components, and proprietary methodologies developed within the project. This process supports informed decision-making regarding protection mechanisms, ownership arrangements, and potential exploitation pathways for each identified Key Exploitable Result (KER).

In parallel, designated Exploitation Contact Points from each partner organisation have been actively involved in coordinating exploitation-related activities both within the consortium and in interactions with external stakeholders. These contact points facilitate:

- alignment of partner-specific exploitation strategies
- communication with potential users, investors, and collaborators
- exploration of licensing, technology transfer, and commercialisation opportunities

This structured approach strengthens the consortium’s capacity to protect its intellectual assets while enabling efficient and coordinated exploitation activities across the project.

The list of exploitation and IPR contact points for all consortium partners is provided in Table 5.

Table 5: EXTRACT exploitation contacts for consortium partners

Partners	Contact
BSC	Alba Abia Hernandez <alba.abiahernandez@bsc.es>
IKERLAN	Jesus Ruano <jmruano@ikerlan.es>
URV	Àngels Luque <angels.luque@fundacio.urv.cat> Daniel <daniel.barcelona@urv.cat> Pedro <pedro.garcia@urv.cat>
OBSPARIS	Thomas Rochette-Castel <thomas.rochette-castel@cnrs.fr> Nikolina Domacinovic <nikolina.domacinovic@univ-paris-diderot.fr> Paraskevi Triantafyllidou <Paraskevi.Triantafyllidou@obspm.fr>
LOGOS	Rafaella Santucci <raffaella.santucci@logos-ri.eu>
Venezia	Enrico Gavagnin <enrico.gavagnin@comune.venezia.it>
SIXSQ	Agathe Veillon <agathe.veillon@sixsq.com>
BINARE	Andrei Costin <andrei.costin@binare.io>
IBM	Erez Hadad <erezh@il.ibm.com>
MATH	Stefano Cuomo <stefano.cuomo@mathema.com>

6.6.3 Partner IPR Policies

Several consortium partners have established internal institutional policies governing the protection, management, and commercialisation of intellectual property generated through their research and innovation activities. A summary of relevant policies is provided below. Some industrial partners may manage intellectual property through internal corporate policies that cannot be publicly disclosed. In such cases, the management of intellectual property within the project follows the rules defined in the Grant Agreement and the Consortium Agreement.

LOGOS Research and Innovation (LRI)

LOGOS Research and Innovation has established internal structures to support the protection and valorisation of intellectual property generated through research activities. The organisation distinguishes between independent research, collaborative research, and commissioned research, with ownership arrangements defined according to the contribution of each participating party.

LOGOS operates a dedicated Research Valorization Unit responsible for supporting the evaluation, protection, and commercialisation of intellectual property. The unit assists researchers with patentability assessments, patent filing procedures, confidentiality agreements, and the identification of potential commercialisation pathways such as licensing, technology transfer, or joint ventures.

The organisation applies a revenue-sharing model that distributes benefits derived from intellectual property between inventors, research units, and institutional innovation funds, ensuring incentives for researchers while supporting future innovation activities.

Universitat Rovira i Virgili (URV)

The intellectual and industrial property generated by researchers of Universitat Rovira i Virgili (URV) is governed by the university's internal regulations on research, development, innovation, and knowledge transfer activities.

In general, inventions created by researchers within the scope of their professional duties are owned by the university, while authors retain recognition of authorship. Software and databases developed as part of research activities are typically owned by URV, with exploitation rights managed through the university's technology transfer structures.

Revenue generated through the commercial exploitation of intellectual property is distributed between inventors and the institution according to predefined institutional rules. In collaborative research projects, ownership and exploitation conditions are defined in the relevant collaboration agreements.

Barcelona Supercomputing Center (BSC)

As a public research institution, the Barcelona Supercomputing Center manages intellectual property in accordance with Spanish national legislation governing research and innovation activities. These include laws regulating authorship, intellectual property ownership, and the protection of trade secrets.

Internal institutional procedures are aligned with the national legal framework and govern the management of research results, including their protection, dissemination, and potential commercialisation.

SIXSQ

SixSq manages intellectual property related to its software platforms and services, including the Nuvla platform, according to its corporate legal framework and publicly available Terms and Conditions. Intellectual property related to platform software and associated applications is owned by SixSq or its licensors, while customers retain ownership of their data.

The company ensures that intellectual property rights are protected through contractual agreements and licensing mechanisms governing the use of its technologies.

BINARE

BINARE owns any background and foreground IPR developed individually before and/or during the EXTRACT project. Intellectual property related to Binare Platform, Binare software and associated applications is owned by BINARE, while the customers or users retain ownership of their data processed through BINARE's software and applications.

BINARE ensures that its intellectual property rights (IPR) are protected through contractual agreements and licensing mechanisms governing the use of its technologies. When using third-party IPR (including IPR resulting from EXTRACT project), BINARE follows the standard contractual agreements mechanisms to secure clear usage, licensing, ownership, benefits, ownership, and other rights relevant to the use of the specific IPR covered by the agreement.

Overall, BINARE manages intellectual property internally according to its corporate policies and in compliance with applicable national and international intellectual property regulations. Nevertheless, BINARE follows and observes any IPR clauses specified in the Grant Agreement (GA) and Consortium Agreement (CA) of the EXTRACT project, especially to the results that qualify as "Joint IPR" between BINARE and at least another organization/partner of the EXTRACT project.

As a Finnish company, BINARE's employee invention rights are governed by the Act on the Right in Employee Inventions, which stipulates that original rights belong to the inventor (employee), but employers can claim rights to inventions made during employment duties. Employees are entitled to "reasonable remuneration" if the employer acquires the rights, a right that cannot be waived by contract. Key Aspects of Finnish Inventor Law:

- Definition of Employee Invention: The invention must be patentable and result from the employee's work duties or significantly utilize the employer's experience/resources.
- Notification Obligation: Employees must notify the employer in writing without delay about any potentially patentable invention.
- Employer Acquisition: Employers have a right to acquire the invention, in whole or in part, by notifying the employee.
- Reasonable Remuneration: When an employer acquires rights, the inventor is entitled to reasonable compensation, considering the invention's value, the scope of rights transferred, and the employment terms.
- Disputes: Inventor rights to claim remuneration generally expire 10 years after the employer notifies the employee they are acquiring the rights.
- Limitations: The Act does not apply to employees at universities or higher education institutions, which have their own specific rules.

- International Transfer: If an inventor changes employers, any agreement regarding a previous invention is invalid if made more than one year after employment termination.

Last but not least, BINARE is bound to and follows the Finnish IPR Laws and guidelines as follows:

- Patents: Governed by the Patents Act, protecting technical solutions, new inventions, and industrial applications for up to 20 years.
- Trademarks: Protected under the Trademarks Act, safeguarding brand names and logos against unauthorized use.
- Copyrights: Protected under the Copyright Act (1961), applying to literary and artistic works, including performers' rights (commonly 50 years).
- Trade Secrets: Protected under special legislation covering non-public information of business value.
- Utility Models: Often called "short-term patents," these protect smaller technical inventions, notes the PRH.

IBM and MATH

IBM and Mathema did not provide detailed institutional IPR policy descriptions for this deliverable. Both organisations manage intellectual property internally according to their corporate policies and in compliance with applicable national and international intellectual property regulations.

Detailed descriptions of partner-specific IPR policies and institutional frameworks are provided in Appendix 4 of this deliverable.

7. Final Sustainability Plan

EXTRACT project sustainability work started M16, in April 2024, after the release of an initial Sustainability Plan of Results, available in *D5.*2. All the sustainability work was carried as part of Task 5.4 Sustainability.

The definition of the word “sustainability” within the framework of European Research projects is the following: "A project is sustainable when it continues to deliver benefits to the project beneficiaries and/or other constituencies for an extended period after the Commission's financial assistance has been terminated."

The sustainability of the results of a project can be assessed thanks to:

- The diversity and intensity of activities/outputs maintained or developed after the end of the project;
- The intensity and potential enlargement of the cooperation between partners, and the consolidation of the project's ecosystem.

In the context of EXTRACT, two pillars have been identified to ensure the project's sustainability:

- EXTRACT results, which can be leveraged and promoted as part of the EXTRACT platform framework, as well as independently.
- EXTRACT community, built on solid foundations thanks to task 5.2 Community building. The community now relies on strong bonds and knowledgeable actors who can share EXTRACT outcomes in relevant situations.

The first iteration of the Sustainability Plan of Results described preliminary ideas and actions to be investigated, and eventually put in place to ensure the sustainability of the project.

This second iteration provides an overview of initiatives taken during the course of the project to raise awareness on the project and its results, then describes concretely the ongoing and upcoming actions to ensure the use of project results after the end of the project.

7.1 Raising awareness on EXTRACT Sustainability

The sustainability activity started at Month 16 of the project, with the clear objective to ensure the use of EXTRACT results after the end of the project.

To complete this objective, the team started raising awareness about sustainability even before the official start of the activity, initiating Consortium discussions during face-to-face meetings, and identifying potential synergies with activities carried out within Dissemination, Community Building and Exploitation tasks.

Thanks to this preliminary work, several actions could be conducted by Month 36, giving more substance to what sustainability means in the project, and alowing a smooth transition with planned actions defined in the plan, which will happen after the end of EXTRACT.

7.1.1 Internal and external awareness around project results

First, results of the project have been shared broadly during events, conferences, pitches, ensuring clear communication of the intermediary milestones reached by the project and proper messaging around EXTRACT components' interoperability and easy integration within European systems.

Section 5 of D5.3 shows that community building played a central role in this activity, actively embedding EXTRACT technologies within Europe's wider AI, data, research, and industrial ecosystems.

Throughout the second phase of the project, partners strengthened engagement with key communities by participating in major European networks such as the ADRA ecosystem, AIonDemand, the Big Data Value Association, EOSC, HiPEAC, and the DataNexus cluster, thereby increasing the project's visibility and fostering long-term collaboration opportunities.

These activities helped identify potential adopters, refine dialogues around extreme-data technologies, and create bridges with scientific communities such as astronomy and crisis-management stakeholders. The project also invested heavily in internal and external training with hands-on sessions, workshops, tutorials and hackathons that equipped community members with practical skills to apply EXTRACT tools in their own environments.

A highlight was the EXTRACT hackathon hosted by the NESU university, where business students had the opportunity to pitch the EXTRACT platform through new use cases they had defined. All students were also asked to provide a business sustainability plan with business models, financial projection and investors' mapping. All the groups provided valuable insights, investigating Software-as-a-Service (SaaS) pricing models as well as public and private Finnish and European funding sources.

Together, these sustained interactions, joint events, and capacity-building efforts form a robust ecosystem around EXTRACT, positioning its results to continue generating impact well beyond the end of the project.

7.1.2 EXTRACT Results individual exploitation

In addition to raising awareness on the importance of educating the community to the EXTRACT project results and resulting possibilities, another key aspect to tackle during the project's lifetime was the components' exploitation. The topic is key, both for project exploitation and sustainability perspectives as it acts as a foundation for discussions on how to exploit and keep using the results, at the platform level.

EXTRACT got selected to benefit from the Horizon Booster Results consulting services. As part of this process, consortium partners reflected on their Key Exploitable Results and future exploitation, after the end of the project.

All partners provided information according to the template available in Table 6, focusing on the next 6 months and describing main actions and achievements to be expected for their components, as well as coverage.

Table 6: Exploitation roadmap template for Key Exploitable Results.

Exploitation roadmap for KER [title of the KER]	
Actions	*Briefly describe actions planned to be executed 3-6 months after the end of the project. Make sure you do not just focus on technical activities (realisation of a prototype, software interface, etc) but also consider the finalisation of a business plan, the protection of intellectual property, the collection of authorisations, all it will be needed to start implement what is in your exploitation plan*
Roles	*Roles of partners involved in the actions defined above. Start with the identification of the organisation which will be the main responsible for the exploitation roadmap (exploiting the KER)*
Milestones	*List the milestones and KPIs to be used for monitoring the implementation of the actions listed above. Add timeline.*
Costs	*Cost estimation to implement planned activities (1 year, 3 years). Provide information on the costs/investments needed to bridge the end of the project to the next steps planned and increase TRL or go to market (you may invest in a patent, in the realisation of a prototype, etc.).*
Revenues	*Projected revenues and eventual profits once the KER will be used (1 and 3 years after use). Consider revenues you will expect to collect by licensing, or thanks to service provision or sale of devices. They generate the cash flow that will make the use of the result sustainable over time (provide an estimation concerning the first year and what is expected after 3 years, if possible). It is recommended that you estimate the revenues according to your early adopters and potential customers and include the information in the draft exploitation plan.*
Other sources of coverage	*Resources needed to bridge the investment needed to increase TRL and ensure the result is used. Financial resources to cover costs incurred before collecting the first revenues (during the "time to market" – see costs) and their sources. Sources can be partners` own budget, other project grants, national/regional incentives, risk capital, loans, etc. Make sure to obtain them at the right timing.*

KERs' exploitation roadmaps are available in Annex 4, below the characterization tables.

Across the consortium, partners adopt a clear and consistent exploitation strategy centred on maintaining, maturing, and operationalising their Key Exploitable Results (KERs) in the months following the project's end.

Most open-source software KERs, such as the EXTRACT Platform, Kubecompss, Lithops, DataPlug, and the various TASKA components, prioritise continued codebase maintenance, consolidation of documentation, and integration into research infrastructures or scientific observatories, with scientific partners planning journal publications, pilot deployments, and extensions of functionality as early milestones.

Commercial-oriented KERs, such as Nuvla and BISETO-CC, focus on strengthening existing market offerings, expanding features (e.g., data-warehouse additions for Nuvla), and leveraging established commercial channels to reach early adopters.

For use-case specific KERs like the Urban Digital Twin, PER use case, Device Simulator, MPC module, and radio-astronomy TASKA pipelines, exploitation actions emphasise deployment in operational environments, alignment with public-sector or scientific use cases, and progressive TRL increase through further R&D or pilot integrations.

Several partners also plan business-plan refinement, IP strategy definition, and acquisition of required authorisations (notably in the PER use case), while others rely on future Horizon Europe funding, institutional budgets, and community contributions to sustain development.

Overall, the consortium's approach combines open-source ecosystem building, scientific dissemination, commercial exploitation where applicable, and targeted deployment within real-world testbeds to ensure maintenance, improvement, long-term adoption and impact of the EXTRACT results.

In addition to individual roadmaps, exploitation of the EXTRACT platform as a whole is also considered, to continue, and potentially expand, the use of the project results.

7.2 EXTRACT sustainability plan: availability of results for future use

When EXTRACT ends, in March 2026, results of the project will continue to be used, either together or individually, for research and commercial purposes. This section describes the ongoing and performed actions to maximise the use of results by making sure of their availability and replicability.

7.2.1 EXTRACT platform availability

After the end of the EXTRACT project, the platform's components will continue to be maintained and stored within the infrastructures of the partners responsible for their development, ensuring their long-term availability and reuse.

Each Key Exploitable Result (KER), including software modules, datasets, workflow engines, and supporting tools, will remain hosted by the partner owning or maintaining that technology, typically within their institutional repositories, cloud or edge platforms, or existing open-source channels.

Components already integrated into mature platforms, such as Nuvla or HPC and continuum-computing frameworks, will remain accessible and managed through these environments.

Research-oriented components will be preserved within partners' research infrastructures, Git repositories, or academic platforms like Zenodo.

This decentralised but structured storage model ensures that EXTRACT technologies remain available, reproducible, and ready for further exploitation, integration, and evolution well beyond the project's lifetime.

All components repositories have been captured in D4.5 (*Final release of the EXTRACT Software Platform)*, as well as their installation processes for anyone to be able to reproduce the EXTRACT platform, or to adapt it to other use cases.

Furthermore, the EXTRACT platform will continue to run in the OBSPARIS lab, to support further astrophysical research within the framework of the TASKA use cases. OBSPARIS is also in contact with other teams interested in deploying the EXTRACT TASKA framework for their own use cases.

7.2.2 EXTRACT platform replicability

The EXTRACT platform is designed to be fully replicable thanks to its modular, open-source architecture and the detailed installation procedures provided in D4.5. The platform is composed of well-defined layers described in D4.5: Programming Models, Data Infrastructure, Orchestration, and Compute Continuum Abstraction. Each of them is implemented through standalone components that can be deployed independently or combined into a full continuum-spanning environment.

Every component is distributed through accessible GitLab or GitHub repositories, and D4.5 documents step-by-step installation workflows, including container images, configuration scripts, Ansible playbooks, and Kubernetes deployment instructions. This structure allows any organisation with access to standard cloud, edge, or on-premise Kubernetes infrastructure to reproduce the platform's functionality.

Replicability is further reinforced by the use of widely adopted technologies (e.g., Kubernetes, Docker Buildx, Prometheus, Ceph/S3-compatible storage), avoiding vendor-specific dependencies. As a result, the EXTRACT platform can be redeployed in diverse environments—from HPC clusters to edge devices—provided that the corresponding hardware and service prerequisites are met. While some components require advanced setup (e.g., multi-cluster networking with Linkerd or Skupper, monitoring via Ansible roles, specialised SkyStore deployments), D4.5 provides sufficient guidance to ensure reproducibility by technical teams familiar with distributed systems.

Overall, the platform's open repositories, clear installation paths, and reliance on standard container-native tools make EXTRACT a replicable, portable, and reusable software stack for extreme-data workflows.

7.2.3 Additional storage possibilities for EXTRACT components

Edge commercial marketplaces like Nuvla

Edge marketplaces such as Nuvla provide a catalogue of ready-to-use, containerised applications that can be deployed automatically on distributed edge devices. These marketplaces offer end-users a simple, uniform way to install, update, and manage applications without needing to understand the underlying infrastructure. In practice, an edge marketplace acts as an app store for edge computing: users select an application, Nuvla deploys it to the target devices, monitors it, and handles lifecycle operations such as upgrades and configuration. This model works well for self-contained services that can run autonomously on an edge node, require minimal manual configuration, and expose a clear user-facing function (e.g., AI inference, data acquisition, security monitoring, industrial gateways, etc.).

In the context of the EXTRACT project, however, it was not appropriate to host the components listed in D4.5 on the Nuvla marketplace. Most of the EXTRACT software modules are not standalone applications, but instead form parts of a distributed software stack that must be jointly deployed across HPC clusters, cloud infrastructures, and edge devices. They require complex orchestration, multi-cluster Kubernetes deployments, and bespoke networking configurations that cannot be encapsulated in a single deployable unit. Many components, such as COMPSs, Kubernetes multi-cluster configurations, SkyStore deployments, the monitoring infrastructure, or the bespoke TASKA-C pipelines, need system-level configuration, privileged cluster access, or custom runtime environments that fall outside the standard capabilities of an edge marketplace. Additionally, these components were designed as research artefacts rather than packaged products, meaning they are not prepared with the metadata, APIs, lifecycle scripts, or user-facing interfaces normally expected for marketplace publication.

The main obstacles are therefore structural. First, EXTRACT components span multiple architectural layers and cannot be meaningfully deployed as isolated edge applications. Second, their installation requires manual steps, cluster-specific adaptations, and integration with other parts of the EXTRACT platform, none of which can be automated through a marketplace's standard deployment model. Third, several EXTRACT modules rely on HPC or cloud-side capabilities that a marketplace cannot provision or manage. Finally, the EXTRACT platform itself is not targeted at end-users in the sense assumed by edge marketplaces: its components are intended to be operated by engineers and researchers within controlled compute environments, and not consumed as turnkey edge applications. For these reasons, packaging EXTRACT technologies into Nuvla marketplace apps would not have been technically feasible or aligned with the project's objectives.

Open Source storage initiatives

In addition to the Nuvla marketplace, several Open Source options have been considered to host the EXTRACT results after the end of the project.

Several open-science and open-source initiatives offer suitable options for hosting the EXTRACT platform components beyond the project's lifetime.

The European Open Science Cloud (EOSC) is a natural candidate, as it provides a federated environment where research data, tools and services can be published, discovered and reused across Europe, and has been explicitly considered within the project as a potential home for EXTRACT outputs.

For long-term preservation of software source code, Software Heritage is also relevant and already listed among the storage options explored within EXTRACT's sustainability planning, as it guarantees durable archiving, version tracking and persistent identifiers for research software.

In addition, OpenAIRE and Zenodo (operated by CERN and indexed in the Open Science ecosystem) are solid options for publishing software releases, documentation and datasets, and they are commonly used in EU-funded projects to ensure FAIR access and citability of research outputs. Zenodo is already used by the project to ensure EXTRACT publications and public deliverables are identified and can be accessed easily.

Together, these initiatives provide a range of complementary solutions, from code archiving to service exposure and research-output publication, allowing EXTRACT partners to ensure long-term accessibility, discoverability and reuse of the project's components.

Selection of additional storage and ongoing actions

After careful assessment of the different options, the project has decided to investigate storage and referencement possibilities within the EOSC EGI cluster.

EOSC is the most suitable platform for exposing EXTRACT outputs to the wider European research ecosystem, as it is explicitly designed to publish, catalogue, and make research data, tools, and services discoverable across disciplines and countries, and it was already identified within the project as a potential home for EXTRACT results. Its FAIR-aligned environment ensures that EXTRACT components—particularly those intended for reuse, training, or integration in scientific workflows—can be accessed and used by researchers beyond the project's duration.

Furthermore, the project already has links with the EOSC federation thanks to the TASKA use case, which actively engaged with EOSC communities during the project. Members of the TASKA team contributed technical demonstrations and presentations to EOSC-related expert groups, notably sharing their innovations in radio astronomy data processing with the EOSC Virtual Research Environments expert group, where they showcased advanced capabilities for processing radio astronomy data across distributed cloud infrastructures.

They were also invited to present the TASKA interactive workflows during the EOSC Winter School, an opportunity that highlighted the operational maturity of the TASKA pipeline and its readiness for deployment within EOSC nodes. These interactions strengthened the visibility of EXTRACT technologies within the EOSC ecosystem and demonstrated how TASKA's extreme-data workflows can support broader scientific communities seeking scalable, cloud-enabled processing solutions.

OBSPARIS shared contact at EGI, and introduced the EXTRACT project team. A meeting will be organised in the next weeks to discuss a potential collaboration between the two entities and a deployment of the EXTRACT platform on EGI to foresee deployment within EOSC.

8. Evaluation of Performance and KPIs (until M39)

Activity	KPI	Measure (by end of Project)	Status
Corporate image	Logo and poster creation	1 PPT, Word, and poster template	Met
Promotional material	Flyers, posters, Project videos	At least 2 posters/infographic At least 3 Project videos	Surpassed: 7/2 posters Surpassed: 10/3 videos Surpassed: 3/1 infographics, 1 factsheet
Website	Visitor statistics	Average 1000 sessions per year	Surpassed: 13k/3k sessions
Social media	Number social media followers	300 total across social media accounts	504/400 social media followers 494 on LinkedIn
Press strategy	Number of press releases/ press clippings	One press release /year At least 25 press clippings	On track: 2/3 PR On track: 18/25 clippings
Event participation	Key notes, organized events with significant attendance (more than 30 ppl)	At least 2 keynotes, at least one organized event, at least one booth in an industrial event	Met: 2/2 key notes 3/1 organized events 3/1 booths in industrial event 47 event participations
Publications	Papers in conferences and journals	At least 4 publications per year	Met: 14/12 publications 4 in progress

Training courses	Organization of internal training courses for Project partners and participation in external trainings such as summer schools, tutorials, workshops	1 internal training course per year At least 1 workshop, 1 tutorial and 1 summer school	Met: 2/1 tutorial 6/3 internal training course 3/1 workshop 1/1 external Hackathon
Final event	Organization of a final Project event	Significant attendance 50 or more	Met: 4/1 final events Total of 200 attendees

9. Conclusion

The EXTRACT project, as detailed in this deliverable D5.3, has successfully achieved its goals in terms of dissemination, exploitation, community building and sustainability. Since its initiation on January 1, 2023, the dedicated efforts of the Work Package 5 team, led by partners Binaré, Barcelona Supercomputing Center, and SixSq, have been instrumental in achieving multifaceted objectives.

The EXTRACT project concludes its dissemination, exploitation, community building and sustainability activities with a strong, coherent set of achievements that demonstrate the project's maturity and long-term relevance.

Over the full project duration, the consortium successfully established a clear project identity, deployed effective communication channels, and significantly exceeded initial KPIs in website engagement, social media performance, promotional material production, and participation in European research and industry events.

Dissemination activities intensified as the technology matured, resulting in extensive event participation, impactful final validation activities for both use cases, and a broad portfolio of peer-reviewed publications.

Community-building efforts strengthened EXTRACT's presence across key European AI, data, HPC and crisis-management ecosystems, while internal and external training events ensured that knowledge and tools developed in the project can continue to be used and transferred beyond its lifetime.

The exploitation work identified and refined a comprehensive set of Key Exploitable Results and mapped the pathways, stakeholders and funding mechanisms needed to support their future adoption.

Finally, the sustainability strategy outlines a clear vision based on leveraging EXTRACT's advanced algorithms and maintaining a strong stakeholder ecosystem, ensuring that the project's technological, scientific and societal benefits continue to grow after the end of the funded period.

In summary, the EXTRACT project stands as a testament to innovation, collaboration, and impact, navigating a well-informed trajectory toward success in the dynamic domain of extreme data management and analytics. The comprehensive evaluation of performance and KPIs reinforces the project's commitment to excellence and signals its continued advancement in the pursuit of its goals.

10. Acronyms and Abbreviations

- ADRA – AI, Data and Robotics Association
- AI – Artificial Intelligence
- BDVA – Big Data Value Association
- BE – British English
- BIN – BINARE
- BSC – Barcelona Supercomputing Center
- D – deliverable
- DoA – Description of Action (Annex 1 of the Grant Agreement)
- EXTRACT – A distributed data-mining software platform for extreme data across the compute continuum
- EB – Executive Board
- EC – European Commission
- EOSC – European Open Science Cloud
- HE – Horizon Europe
- HPC – High-Performance Computing
- IPR – Intellectual Property Right
- KER – Key Exploitable Result
- KPI – Key Performance Indicator
- M – Month
- ML– Machine Learning
- SIXSQ – Six Square
- SME – Small and Medium-sized Enterprise
- STEM – Science, Technology, Engineering and Math
- WP – Work Package
- 8M – Eight of March, International Women's Day
- HRB - Horizon Results Booster

Appendices

Annex 1 - EXTRACT Dissemination Register

Partner	Activity	Date
ALL	Press release: Introducing DATANEXUS: the EU Cluster Harnessing Extreme Data using Advanced European Technologies	01-Jul-24
OBSPARIS	Presentation to the French SKA-SRC WP3 of EXTRACT- TASKA and performed a live demo of the current version of the TASKA Use Case MVP	10-Jul-24
URV	News piece: Pursuing the ultimate serverless experience	19-Jul-24
ALL	Factsheet: Innovative Approaches to Extreme Data Challnges-DataNexus Cluster	27-Aug-24
ALL	Newspiece on Factsheet release: NEW FACTSHEET: Learn ABOUT the DATANEXUS Cluster	27-Aug-24
ALL	Nit de la Recerca flyer	09-Sept-24
VEN	Video: Transforming Crisis Management in Venice Through Innovative EXTREME DATA Solutions	17-Sept-24
All	Website news: Transforming Crisis Management in Venice Through Innovative EXTREME DATA Solutions	17-Sept-24
MATH	EBDVF 2024 Session	02-Oct-24
ALL	Horizon Results Booster Pitch Video	02-Oct-24
ALL	100 flyers distributed at EBDVF 2024	02-Oct-24
BSC/BINARE/ MATH	EBDVF 2024 Booth	02-Oct-24
IKERLAN	Newspiece: EXTRACT´s monitoring platform: The infrastructure behind the use cases	07-Oct-24
ALL	Newspiece: Highlighting DataNexus cluster at EBDVF 2024	09-Oct-24
BINARE	EXTERNAL Hackathon at 2024 Nordic Economics Students Union	16-Oct-24
BINARE	Exploitation/Sustainability - Nordic Innovation House REACH Investors Event Stockholm	16-Oct-24
BSC	News on website: Optimizing Compute Continuum Workflows with COMPSs Orchestration	21-Oct-24
BINARE	Hardwear.io NL 2024 Booth/Banner	23-Oct-24
BINARE	Attendance and EXTRACT Flyer distribution at ACM IMC (Internet Measurement Conference) 2024 Madrid	03-Nov-24

ALL	News on website: BINARÉ-hosted EXTRACT hackathon drives innovation at the NESU Autumn Conference: Students deliver impressive long-term sustainability plans	04-Nov-24
BSC	News piece on BSC website :BSC showcases the potential of supercomputing and AI to create the city of the future at the Smart City Expo World Congress	04-Nov-24
ALL	Flyers at Smart City World Expo Congress	05-Nov-24
BSC/Venezia	Breakout session Smart City World Expo Congress	05-Nov-24
BSC	Booth representation at Smart City World Expo Congress	05-Nov-24
ALL	News on website: Successful presentation of the PER case at the Smart City Expo World Congress	08-Nov-24
ALL	News on website: "TASKA "A" Use Case: "A"gile detection and analysis of solar activity using NenuFAR radiotelescopes"	11-Nov-24
ALL	News on website: TASKA "C" Use Case: Fast-paced data calibration and imaging on the "C"loud	12-Nov-24
OBSPARIS	French Astrophysics at radio frequencies, toward the SKA: "EXTRACT: preparing large-scale radio interferometric data processing in the SKA era"	14-Nov-24
ALL	News on website: EXTRACT deliverables now available	28-Nov-24
ALL	News on website: Year 2 Wraps Up with Geneva F2F	29-Nov-24
BINARE	New on website: Toward "Cybersecurity blueprints" that address modern security challenges	11-Dec-24
EKINOX	News on website:Strengthening EXTRACT project success through external expertise and collaboration	16-Dec-24
OBSPARIS	News on website: TASKA use case presented at key astronomy events	19-Dec-24
BSC	Interview with Catalan Mobilicat magazine La intel·ligència artificial ha de ser oberta o no serà	07-Jan-25
BSC/URV	Organized Workshop-2024 HiPEAC Compute Continuum	20-Jan-25

BSC	News on website: Advancing the compute continuum at EXTRACT-organized WS at HiPEAC	05-Feb-25
OBSPARIS	Presentation: Workshop participation-TASKA-C at the– SKA-France Workshop	11-Feb-25
ALL	WS at Future-Ready On-Demand Solutions with AI, Data, & Robotics Event	18-Feb-25
SIXSQ	News on website: Sustainability in the EXTRACT project	20-Feb-25
ALL	News on website: DataNexus Cluster Shares Common Challenges & Shared Successes @ Future Ready BOF	28-Feb-25
VEN	Presentation at GovTech Forum 202, "Technology to manage disasters"	13-March-25
ALL	News on website: What's Next for EXTRACT as part of the DATANEXUS Cluster?	21-March-25
VEN	News on website: Harnessing technology for disaster management: The EXTRACT Project at GovTech Forum 2025	03-Apr-25
LOGOS/MATH	News on website: Venice's Personalized Evacuation Route Use Case Advances with Mobile Interface and 3D Urban Digital Twin Development	07-Apr-25
BSC	Panel participation: European Convergence Summit (ECS) 2025	09-Apr-25
ALL	News on website: EXTRACT consortium gathers in Florence to strengthen collaboration and align on progress	10-Apr-25
BSC	News on website: Building European Resilience: AI, Robotics, and Big Data at the Forefront of Crisis Management	15-Apr-25
URV	News piece: A Crucial Step Towards Optimization: On-the-Fly Partitioning	21-May-25
BSC	Presentation at CIDAI Conference: Data and AI Proof of Concept	21-May-25
BSC	Workshop at DATAWEEK 2025- "Extreme data driven solutions and services – Updated perspectives and Challenges for Europe"	27-May-25
OBSPARIS	2 poster presentations at 7th MADICS Symposium	27-May-25
OBSPARIS	AI Seminar at HPA pole "Working with AI"	05-Jun-25
BSC	Presentation and poster at EGI2025	05-Jun-25
OBSPARIS	International Workshop on Planetary, Solar, and Heliospheric Radio Emissions	11-Jun-25
OBSPARIS	Presentation: "Les données de NenuFAR, éclaireur de SKA: un défi dans un paysage contraint" at Semain de l'Astrophysique Française	02-July-25
BSC	News on website: Deploying and Executing Workflows Across the Compute Continuum	15-July-25

IKERLAN	News on website: Swim-nsm monitoring module for network state monitoring	21-July-25
IBM	Presentation in 51st International Conference on Very Large Data Bases	1-Sept-25
MATH	Presentation at IEEE International Conference on Cyber Humanities (IEEE CH): "Latent Space Security – Implications and Challenges"	08-Sep-25
BSC	Presentation at ELIXIRION Schools 3/4 Training	15-Sep-25
BINARE	Stand at TechEx Europe at RAI Amsterdam	24-Sep-25
Venezia/BSC/ LOGOS	PER Final Event: The First Three Days- technical validation and 3 presentations	03-Oct-25
VEN/BSC/LOGOS	News on website: EXTRACT shares innovative crisis-tech at F3D Event in Venice	15-Oct-25
BSC	Participation in the BDVA Workshop on Evolving Data Architectures, "From Workflow Description To Compute Continuum Deployment"	22-Oct-25
ALL	News on website: EXTRACT SHARES dATA MINING WFs at BDVA WS on new data architectures	22-Oct-25
BINARE	Participation/attendance OpenInfra Europe: Informal dissemination of EXTRACT and technology uptake/scouting	17-Oct-25
BINARE	EXTRACT Stand/Booth at Valencia Digital Summit DVS2025	20-Oct-25
Venezia/LOGOS	Poster Presentation at Copernicus Emergency Management Service and the Disaster Risk Management Knowledge Centre EU Preparedness event	04-Nov-25
BSC	Smart City World Expo session on "How AI and Data Technologies are Closing the Reaction Gap at 'Moment Zero')	06-Nov-25
Venezia/LOGOS	News on website: EU Science for Preparedness: A Strategic Convergence of Knowledge and Action	10-Nov-25
BINARE	Participation/attendance SLUSH 2026 Helsinki: Dissemination, Technology Demo Scouting and Stakeholder Gathering + Stakeholder Awareness, Flyers Distribution	09-Nov-25
BSC/BINARE	EBDVF 2025 Booth	12-Nov-25

BSC	News on website: EXTRACT PROJECT COORDINATOR PARTICIPATES IN INSIGHTFUL PANEL AT THE 2025 SMART CITY EXPO WORLD CONGRESS	12-Nov-25
BSC	Presentation in cluster session on "Extreme Data Architectures Empowering AI Innovation Across the Compute Continuum" and booth in European Big Data Value Forum	12-Nov-25
ALL	News on website "EXTRACT results presented at EBDVF'25"	20-Nov-25
OBSPARIS	Presentation and Demo for OAEG4 EOSC working group meeting	25-Nov-25
ALL	News on website "EXTRACT Showcases Distributed WFL Orchestrator for Radio Astronomy at OAEG4 Meeting"	01-Dec-25
OBSPARIS	Presentation and booth in 11th Annual Science At Low Frequencies (SALF) Conference	07-Dec-25
ALL	News piece: TASKA final event at SALFXI	08-Dec-25
ALL	News piece: ASNUM 2025- Journées de l'Action Numérique Spécifique	15-Dec-25
OBSPARIS	Presentation at ANSUM 2025	15-Dec-25
SIXSQ	News on website: Organizing extreme data at the Edge	16-Dec-25
ALL	News on website: TASKA Celebrates its Final Event at SALF XI: Advancing Radio Astronomy with EXTRACT	23-Dec-25
OBSPARIS	EOSC Winter School	28-Jan-26
BSC, MATH, Venezia, LOGOS	Final Tech Validation in Venice	04-Feb-26
ALL	News on website:EXTRACT Showcases Distributed Workflow Orchestrator at EOSC Winter School 2026	17-Feb-26
ALL	News on website: EXTRACT Project Completes Second PER Validation in Venice	19-Feb-26
ALL	News on website:A View from the Field: Partner Venezia Weighs in on the PER use case	12-Mar-26
ALL	News on website: Engineering Students Extend EXTRACT TASKA C with a Visual Interface for Astronomy Workflows	20-Mar-26

Annex 2 - EXTRACT Detailed list of publications

Publication type	Title	Authors	DOI	Repository	Journal title or equivalent	Publisher	Year
Conference proceedings	Multi-party Computation for Privacy and Security in Machine Learning: a Practical Review	Bellini, Alessandro; Bellini,Emanuele; Bertini, Massimo; Almhaithawi, Doaa; Cuomo, Stefano	10.1109/CSR57506.2023.10224826	https://iris.polito.it/retrieve/handle/11583/2982820/682640	2023 IEEE International Conference on Cyber Security and Resilience	IEEE	2023
Conference proceedings	Is Performance of Object Storage Predictable for Serverless I/O Workloads? A Comparative Study	G. Eizaguirre and M. Sanchez-Artigas	10.1109/ICNP59255.2023.10355617	https://zenodo.org/records/11382387	Cloud-Edge Continuum Workshop 2023 (IEEE 31st International Conference on Network Protocols	IEEE Computer Society	2023
Conference proceedings	Scaling a Variant Calling Genomics Pipeline with FaaS	Aitor Arjona, Arnau Gabriel-Atienza, Sara Lanuza-Orna, Xavier Roca-Canals, Ayman Bourramouss, Tyler K. Chafin, Lucio Marcello, Paolo Ribeca, and Pedro García-López	https://doi.org/10.1145/3631295.3631393	https://arxiv.org/html/2312.07090v1	9th International Workshop on Serverless Computing	Association for Computing Machinery	2023

Conference proceedings	Glider: Serverless Ephemeral Stateful Near-Data Computation	Daniel Barcelona-Pons, Pedro García-López, and Bernard Metzler.	https://doi.org/10.1145/3590140.3629119	https://zenodo.org/records/10026859	Proceedings of the 24th International Middleware Conference (Middleware '23)	Association for Computing Machinery	2024
Journal	MLLess: Achieving cost efficiency in serverless machine learning training	Pablo Gimeno Sarroca, Marc Sánchez-Artigas	https://doi.org/10.1016/j.jpdc.2024.104871	https://arxiv.org/abs/2206.05786	Journal of Parallel and Distributed Computing	Science Direct	2024
Journal	Semantic Segmentation of Solar Radio Spikes at Low Frequencies	Pearse C Murphy, Stéphane Aicardi, Baptiste Cecconi, Carine Briand, Thibault Peccoux	https://doi.org/10.33232/001c.120317	https://arxiv.org/abs/2403.08546v2	The Open Journal of Astrophysics	Maynooth Academic Publishing	2024
Conference proceedings	Dataplug: Unlocking extreme data analytics with on-the-fly dynamic partitioning of unstructured data	Aitor Arjona, Pedro García-López, Daniel Barcelona-Pons	https://doi.org/10.5281/zenodo.14161769	https://zenodo.org/records/14161823	2024 IEEE 24th International Symposium on Cluster, Cloud and Internet Computing (CCGrid)	IEEE	2024
Conference proceedings	Serverful Functions: Leveraging Servers in Complex Serverless Workflows (industry track)	Germán T. Eizaguirre, Daniel Barcelona-Pons, Aitor Arjona, Gil Vernik, Pedro García-López, Theodore	https://doi.org/10.1145/3700824.3701095	https://zenodo.org/records/15019182	Proceedings of the 25th International Middleware Conference Industrial Track	Association for Computing Machinery	2024

		Alexandrov					
Article in journal	A Seer knows best: Auto-tuned object storage shuffling for serverless analytics	Germán T. Eizaguirre; Marc Sánchez-Artigas	https://doi.org/10.1016/j.jpdc.2023.104763	https://zenodo.org/records/7311277	Journal of Parallel and Distributed Computing	Elsevier	2024
Conference proceedings/ workshop	Exploring Secure and Efficient Temporary Data Sharing between co-located Kubernetes Containers	Aitor Arjona, Bernard Metzler, Pascal Spörri	10.1109/ICNP61940.2024.10858528	https://zenodo.org/records/15097005	2024 IEEE 32nd International Conference on Network Protocols (ICNP)	IEEE	2025
Conference proceedings/ workshop	Manifesting the Elasticity of Serverless Data Pipelines: a Metabolomics Use Case	Díez-Cuadrado, E., Eizaguirre, G. T., & Molina-Giménez, E.	https://doi.org/10.1145/3774899.3775018	https://zenodo.org/records/18661222	WoSC11 '25: Proceedings of the 11th International Workshop on Serverless Computing	ACM	2025
Conference proceedings/ workshop	Flexecutor: Out-of-the-Box Smart Provisioning for Serverless Workflows	Molina-Giménez, E., Barcelona-Pons, D., Iacoponelli, O. H., & García-López, P.	https://doi.org/10.1145/3774899.3775013	https://zenodo.org/records/18661418	WoSC11 '25: Proceedings of the 11th International Workshop on Serverless Computing	ACM	2025

Conference proceedings/ workshop	Burst Computing: Quick, Sudden, Massively Parallel Processing on Serverless Resources	Daniel Barcelona-Pons, Aitor Arjona, Pedro García-López, Enrique Molina-Giménez, Stepan Klymonchuk	https://doi.org/10.5281/zenodo.17662627	https://zenodo.org/records/17662627	2025 USENIX Annual Technical Conference	USENIX	2025
Conference proceedings/ workshop	Serverless Data Analytics (Finally) Bridging the Gap: Introducing the ORTZI DATAFRAME	Germán T. Eizaguirre, Marc Hostau, Marc Sánchez-Artigas	https://doi.org/10.5281/zenodo.17701352	https://zenodo.org/records/17701352	The IEEE International Conference on Cloud Computing CLOUD 2025	IEEE	2025

Annex 3 - EXTRACT KER Characterisation Tables and Exploitation roadmaps

KER name: **Nuvla + NuvlaEdge**	
	Input from the Beneficiary
Description	Nuvla.io is a cloud-native platform designed for the **remote management of edge devices and applications** using its agent software, NuvlaEdge. Beyond simplifying the deployment of containerized applications across edge locations, it features a **metadata catalogue** for categorising, mapping, and easily retrieving data collected at the edge.
Target market/end users	**Target market:** Industrial adopters of edge computing, telecoms, IoT platform providers, public agencies. **End users:** Enterprises managing distributed infrastructure, research infrastructures, Smart City platforms.
End-users needs / problems	Enterprises and public organisations operating distributed edge infrastructures face the following challenges: Remotely managing large fleets of geographically distributed edge devices, Deploying and updating containerised applications consistently across heterogeneous environments. Nuvla supports both Docker and Kubernetes deployments, Handling high volumes of edge-generated data without structured metadata management, Lack of unified visibility over infrastructure, applications, and datasets, Difficulty ensuring secure, policy-compliant data processing at the edge, Limited automation between data generation events and downstream workflows (e.g., AI pipelines, storage ingestion), Fragmented toolchains combining separate orchestration, monitoring, and data cataloguing solutions. Research infrastructures and Smart City platforms particularly struggle with dataset discoverability, provenance tracking, and cross-team data reuse. The Nuvla.io platform, together with NuvlaEdge, addresses these needs by integrating edge orchestration and metadata-driven data catalogue capabilities in one solution.
Competitive advantages	**Market-ready technology** (TRL 5+) with commercialization pathway (SaaS and on-premise). Combines remote orchestration of edge devices and applications, and metadata-driven data management in one solution. Nuvla is already exploited in remote orchestration business projects, such as: • Traffic light smart management. • Hospital network for health data. Advantages of Nuvla: • With Nuvla, companies can reduce their infrastructure costs by up to 75% compared to cloud-only solutions. • Nuvla has been validated at scale with 10,000 edge devices and 10,000 deployed applications.
Use model	Commercial service: Software-as-a-Service (SaaS). Early adopters: industrial pilots in IoT, smart cities, and public agencies seeking secure data processing.
Early Adopters	**Primary early adopters:** Industrial IoT pilots generating large volumes of distributed data, Smart City platforms deploying edge-based services, Public agencies requiring secure and sovereign edge data processing, Telecom operators deploying distributed edge infrastructure, Research infrastructures operating distributed testbeds

	Channels to reach them: Direct commercial channels of SixSq, Commercial channels of mother company Ekinops (network provider), IoT and edge computing industry events, Smart City and GovTech conferences,Horizon Europe and Digital Europe project networks, Partnerships with telecom and infrastructure providers, Open-source and developer community engagement Early adopters are typically organisations already experiencing operational complexity due to distributed edge environments.
Adopters' problems/needs	Early adopters need to:Reduce operational overhead in managing distributed edge fleets, Ensure reliable remote deployment and lifecycle management of applications, Automatically register and categorise datasets generated at the edge, Improve data discoverability, traceability, and reuse, Integrate edge-generated data with storage systems such as MinIO, Enable event-driven automation (e.g., via MQTT notifications) when new datasets are created, Maintain security, compliance, and visibility across distributed infrastructure. They require a market-ready, commercially supported solution rather than building custom integrations.
Alternative solution	Currently, adopters solve these problems by: Using container orchestration tools such as Kubernetes for application deployment, Deploying edge extensions such as KubeEdge., Relying on cloud-native platforms from providers like Amazon Web Services or Microsoft Azure, Implementing standalone storage solutions (e.g., MinIO) without integrated metadata governance, Developing custom middleware and internal metadata databases. These approaches are often fragmented, require significant integration effort, and lack unified orchestration + metadata lifecycle management.
Unique Value Proposition	Nuvla + NuvlaEdge uniquely combine: Edge device lifecycle management, Remote deployment of containerised applications, Integrated metadata-driven data catalogue, Continuous dataset mapping and indexing, Event-driven automation flows, SaaS and on-premise deployment options Compared to alternative solutions, Nuvla: Reduces integration complexity by unifying orchestration and data management, Shortens time-to-deployment for edge pilots, Improves dataset discoverability and reuse across teams, Supports sovereign and hybrid deployments (important for public agencies), Is already commercially exploited and validated beyond research pilots. It is not only a technical framework but a commercially supported platform with an established ecosystem.
Competitors	1. Hyperscaler edge platforms • Amazon Web Services • Microsoft Azure • Google Cloud **Strengths:** scalability, mature ecosystems, integrated cloud services. **Weaknesses vs Nuvla:** vendor lock-in, cloud-centric architecture, limited flexibility for sovereign/on-prem deployments. 2. Edge orchestration frameworks

	• KubeEdge • Rancher **Strengths:** strong edge orchestration capabilities, open-source community. **Weaknesses vs Nuvla:** lacks integrated metadata catalogue and commercial full-stack solution. 3. Standalone data governance tools • Apache Atlas **Strengths:** strong metadata management. **Weaknesses vs Nuvla:** not integrated with edge device orchestration; requires additional integration layers.
Partners	Owner: SixSq Technical integration partners in the EXTRACT project: - OBSParis for data collection. - BSC CNS for data layer integration. - IBM Israel for cloud patterns and storage.
Timing	**Short-term:** Available on the market today (TRL 5+; commercially exploited). **Mid-term (1–2 years):** Enhanced data catalogue features and automation capabilities. **Long-term:** Expansion into large-scale industrial edge and telecom deployments.
IP Strategy	Commercial IP managed by SixSq. Most of the Nuvla and NuvlaEdge software are Open Source, only the billing component is proprietary.

Exploitation roadmap for Nuvla (Nuvla.io platform + NuvlaEdge)	
Actions	The Nuvla.io platform and NuvlaEdge software are already exploited commercially and available as Open Source software on GitHub. In addition to the data catalogue, SixSq foresees the implementation of an Open Source data warehouse, as part of another European project (with the European Space Agency). This edge data warehouse will allow users to monitor and to access historical data on their edge devices health, availability, as well as the containerised applications running on it.
Roles	SixSq is already solely exploiting the Nuvla platform as a PaaS (Platform-as-a-Service).
Milestones	Number of users of the Nuvla solution.
Costs	The main costs for Nuvla today are the hosting of the platform in the cloud, and the development of the data warehouse feature.
Revenues	Nuvla operates with a Pay-as-you-go model, in which users are billed on a monthly or yearly basis depending on the subscription they chose. The different pricing models are defined by the size of the customer's edge devices fleet.
Other sources of coverage	Funding from ongoing and future European research projects.

KER name: SkyStore Distributed Object Storage	
	Input from the Beneficiary

Description	**SkyStore** is an open-source S3 service that is jointly developed by IBM (Israel Research Lab) and University of California, Berkeley. It aggregates multiple concrete S3 services from different locations across the compute continuum and makes all data accessible under a single unified namespace. Access to SkyStore itself is S3-compatible (s3-proxy) and can be distributed across the continuum as well. Last, SkyStore enables policies for replicating and caching S3 data so it can accelerate access to remote S3 data subject to various usage patterns such as repeated local access to remote data.
Target market/end users	Target market: Medium and large-scale enterprises with globally-distributed data. Industrial foundations based on edge-to-cloud continuum such as Telecom services, Weather services, etc. Given the open-source and evolving nature, also R&D and productization. Users: Admin/DevOps/CDO on enterprise use (transparent to end-users).
End-users needs / problems	Access your S3 data efficiently, continually and consistently, regardless of the data location (different S3 or cloud regions, private/public premises)
Competitive advantages	SkyStore enables users from either industry or academia to efficiently use S3 data at global scale, federating and making use of both large S3 deployments in datacenters as well as limited edge deployments. It can be deployed cloud-native, stand-alone, or dockerized. Global, unified and managed access to entire S3 data is the bottom line. Most competitive "data federation" offerings general but proprietary API to access data (data lake - all types of data) instead of simplifying federated S3 access to the most common cloud-native type - S3 data.
Use model	Use model: provision of a federated data service that unifies access to multiple globally-distributed S3 regions at different compute continuum scale (e.g., cloud/edge/private premises) The KER is available as an open-source repository with a permissive license.
Early Adopters	Early adopters will likely include further EU projects at different calls, researchers and pilot deployments from interested enterprises dealing with S3-distributed data.
Adopters' problems/needs	Adopters of SkyStore are looking to use their data consistently across different sites. Their data is likely in S3 - private deployments (e.g., Minio etc), cloud regions (AWS S3, OVH S3, etc), on-prem storage (Ceph S3). They would like to access all S3 data through the same S3 interface using consistent global location names (buckets) and efficiently - such as ocal reads and writes.
Alternative solution	All Data Lake solutions are alternatives, but specialized - as they require integration of non-S3 APIs such as data catalog lookup, notifications etc
Unique Value Proposition	This is the "not too big, not too small" advantage. On one hand, we get data federation and consistent access, similar to a data lake (but restricted to S3). On the other hand, We provide standard S3-compliant service, that works with all S3 clients (API/CLI/UI) allowing seamless integration into applications. On top of that, there is the policy asepct - ability to mobilize data transparently to improve access where needed.
Competitors	All major hyperscalers as well as several open-source projects offer generic data federation solutions in the "Data Lake" style.
Partners	Candidates - OBSParis, LOGOS.
Timing	Short term: Code and documentation available today in the repository Any further plans may depend on how adoption unfolds
IP Strategy	Much relevant IP is public - papers, code, documentation etc. Of course, IBM IPR strategy, as defined in the EXTRACT GA, overrules in case of any conflict.

Exploitation roadmap for SkyStore	
Actions	Ongoing community maintenance of the SkyStore codebase, continued use in scientific publications.

Roles	No responsible partner. IBM, UC Berkeley and SkyStore community may appoint leads for development, research, and commercialization effort if so happens.
Milestones	Depend on community interest in SkyStore
Costs	No proclaimed plans yet, so costs are TBD.
Revenues	No revenue model yet. Depends on future commercialization intents.
Other sources of coverage	Time and effort of open-source contributors and/or future European research projects.

KER name: IKERLAN Monitoring Platform (for Kubernetes)	
	Input from the Beneficiary
Description	The IKERLAN Monitoring Platform is a lightweight and extensible software component designed to provide intelligent and energy-aware observability across heterogeneous compute continuum infrastructures. The platform collects performance and energy metrics, including CPU usage, power consumption, and memory utilisation, using Prometheus and stores them in time-series databases for further analysis. In addition to infrastructure metrics, the platform can gather traces and logs from instrumented applications, enabling fine-grained performance analysis of distributed workflows. A built-in API allows seamless integration with orchestration frameworks such as COMPSs, allowing monitoring data to be directly used by scheduling and deployment logic. Through its modular architecture, the platform enables performance- and energy-aware monitoring across edge, cloud, and HPC environments, supporting improved infrastructure management and orchestration decisions.
Target market/end users	The primary target market includes research projects developing orchestration solutions for distributed infrastructures, industrial orchestration frameworks, and operators of critical infrastructures requiring reliable monitoring across heterogeneous environments. Typical end users include researchers working with distributed computing systems, industrial R&D teams developing orchestration technologies, infrastructure administrators managing Kubernetes environments, CIOs and CDOs responsible for infrastructure operations, and operators in sectors such as energy, transport, and manufacturing.
End-users needs / problems	Organisations operating Kubernetes-based and heterogeneous compute continuum infrastructures often face limited visibility across distributed environments spanning edge, cloud, and HPC resources. Monitoring systems typically provide performance metrics but lack integrated energy consumption monitoring, making it difficult to optimise workloads for both performance and sustainability. In many infrastructures, monitoring stacks are fragmented and require manual integration of multiple tools. This makes it difficult to correlate infrastructure metrics with orchestration decisions or to implement energy-aware workload placement strategies. Existing enterprise monitoring platforms can also be complex to configure and costly to operate.

	For operators of critical infrastructures such as energy grids, transport systems, and manufacturing environments, these challenges are compounded by the need for reliable anomaly detection, rapid incident response, and high service continuity with minimal operational overhead. The IKERLAN Monitoring Platform addresses these challenges by providing lightweight, real-time, and energy-aware observability integrated with orchestration frameworks.
Competitive advantages	The platform provides real-time monitoring of both performance and energy consumption across distributed infrastructures, enabling energy-aware observability that is rarely available in conventional monitoring solutions. Its modular and lightweight design allows deployment across heterogeneous environments, including resource-constrained edge devices as well as cloud and HPC infrastructures. In addition, the platform is released as open-source software, which facilitates integration with existing infrastructures and encourages adoption by research projects and industrial pilots.
Use model	The IKERLAN Monitoring Platform will be released as an open-source software component and made available through a software registry and project repositories. It is intended to be integrated into orchestration frameworks and infrastructure management solutions used in research projects and industrial pilot environments. Initial adoption is expected among research infrastructures and pilot industrial deployments operating critical systems that require advanced monitoring capabilities.
Early Adopters	Early adopters are expected to include research infrastructures operating Kubernetes-based orchestration systems, Horizon Europe and Digital Europe pilot deployments, industrial R&D teams exploring energy-aware orchestration strategies, operators of critical infrastructures such as energy networks and smart transport systems, and research initiatives working on HPC–edge orchestration. These adopters will primarily be reached through European research project networks, Kubernetes and cloud-native community events, industrial technology clusters, open-source dissemination channels, and direct collaboration through the networks of IKERLAN.
Adopters' problems/needs	Early adopters typically require detailed monitoring capabilities that allow them to observe CPU utilisation, memory usage, and power consumption across distributed nodes in real time. They also need monitoring outputs that can be directly integrated into orchestration engines to support scheduling decisions and energy-aware workload placement. In addition, adopters seek tools that provide trace-level observability for distributed applications while avoiding the need to deploy heavy external monitoring frameworks. The ability to detect anomalies quickly through dashboards and alerts, while reducing integration complexity between monitoring and orchestration systems, is also a key requirement. These organisations therefore need monitoring solutions that support both performance optimisation and sustainability objectives in distributed infrastructures.
Alternative solution	Many organisations currently rely on combinations of open-source monitoring tools such as Prometheus for metrics collection and Grafana for dashboards and visualisation. Others use commercial observability platforms such as Datadog or deploy native Kubernetes monitoring stacks.

	In some environments, organisations also rely on custom scripts for collecting energy consumption data or application performance monitoring tools that are not designed for energy-aware orchestration. These approaches are often cloud-centric, expensive to operate in enterprise environments, and poorly integrated with orchestration frameworks such as COMPSs.
Unique Value Proposition	The IKERLAN Monitoring Platform provides energy-aware observability specifically designed for compute continuum environments. Unlike conventional monitoring stacks that focus primarily on performance metrics, the platform integrates energy consumption monitoring with real-time infrastructure observability and exposes this information directly to orchestration frameworks. Its lightweight and modular architecture allows deployment across heterogeneous infrastructures, including edge devices, cloud nodes, and HPC resources. The built-in API enables direct integration with orchestration tools, supporting real-time feedback loops between monitoring systems and scheduling decisions. By combining performance monitoring, energy awareness, and orchestration integration in a single open-source solution, the platform reduces integration complexity while enabling more efficient and sustainable resource management.
Competitors	Potential competitors include open-source monitoring ecosystems such as Prometheus and Grafana, which provide mature monitoring capabilities and large community adoption but do not inherently support energy-aware observability or orchestration feedback loops. Commercial observability platforms such as Datadog and New Relic offer enterprise-grade monitoring and analytics but are typically cloud-centric, costly to deploy, and difficult to adapt to research-driven orchestration workflows. Kubernetes-native monitoring stacks also provide integrated monitoring within container environments but generally lack support for heterogeneous edge–cloud–HPC infrastructures and do not prioritise energy-aware infrastructure optimisation.
Partners	IKERLAN lead is coherent; call out BSC CNS (COMPSs integration), SIXSQ (packaging for marketplace), IBM Israel (hybrid cloud patterns).
Timing	Medium-term (technology validated at TRL 5; adoption expected in 2–3 years).
IP Strategy	The platform follows an open-source software strategy aimed at encouraging community adoption and collaborative development. No strict intellectual property barriers are imposed, and the exploitation approach focuses on ecosystem growth, research collaboration, and potential industrial integration.

Exploitation roadmap for IKERLAN Monitoring Platform for Kubernetes	
Actions	Open-source release through project repositories and software registries; integration with orchestration frameworks; pilot validation with research and industrial users; dissemination via EU project networks, cloud-native events, and IKERLAN collaboration channels.
Roles	IKERLAN leads technical development and exploitation; BSC-CNS supports orchestration/COMPSs integration; SIXSQ supports packaging and marketplace-oriented uptake; pilot users validate functionality and adoption needs.
Milestones	Release of the open-source component; integration with orchestration APIs; validation in pilot environments; engagement of early adopters; medium-term market uptake within 2–3 years from TRL5.
Costs	Main costs relate to software maintenance, integration, packaging, pilot deployment, documentation, community support, and dissemination to research and industrial adopters.

Revenues	No direct software-license revenues are foreseen under the open-source model; expected value comes from service-based activities such as integration, customization, technical support, collaborative R&D, and industrial deployment projects.
Other sources of coverage	European research and innovation projects, industrial pilots, collaborative agreements with infrastructure operators, and follow-up exploitation through IKERLAN's service and partnership network.

KER name: Kubecompss (formerly COMPSs-k8s)	
	Input from the Beneficiary
Description	Kubecompss is a software solution that helps users run complex workflows on distributed computing systems without having to manage the infrastructure themselves. In Kubecompss, COMPSs is integrated with Kubernetes running across multiple clusters. Kubernetes provides the computing resources and manages containers, scaling, and deployment, while COMPSs decides when and where each task is executed. This allows workflows to run across different clusters as if they were a single system. COMPSs is a task-based programming model aimed at simplifying the development of applications for distributed environments, such as clusters and clouds. It enables automatic parallelization and efficient execution of complex workflows that may consist of multiple interdependent tasks. In EXTRACT, we have adapted COMPSs to work within Kubernetes multiclusters. The solution reduces the technical effort required to run distributed applications. Users focus on writing their workflows, while Kubecompss handles task scheduling, resource usage, and execution across clusters. This helps researchers run scalable workflows more easily and efficiently in the compute continuum.
Target market/end users	Research and scientific computing sector, specifically organizations and groups that run task-based workflows on distributed infrastructures using Kubernetes. This includes research institutes, universities, and HPC centers that need to execute complex workflows across clusters or geographically distributed resources without investing heavily in infrastructure management.
End-users needs / problems	Researchers and scientists need to run complex workflows on distributed computing resources, but doing so usually requires advanced skills in parallel programming, cluster management, and Kubernetes configuration. Setting up and operating multi-cluster environments, managing containers, and manually scheduling tasks is time-consuming and error-prone. The KER addresses the need for higher automation and simpler execution of distributed workflows, abstracting the infrastructure management from end users.
Competitive advantages	The KER provides automatic task parallelization and scheduling across Kubernetes multiclusters, removing the need for users to manually manage resources, deployments, and execution logic. Unlike alternative solutions that require users to adapt their applications to specific execution models or to handle infrastructure configuration directly, the KER lets users describe workflows sequentially and handles execution transparently. It supports general workflows with complex data dependencies, not limited to a specific application domain, while Kubernetes abstracts the underlying compute continuum. This results in lower technical complexity, better portability across infrastructures, and more efficient use of distributed resources compared to current solutions.
Use model	Researchers currently using COMPSs will be informed when Kubecompss becomes operational, so that they can use this tool in their research activities.
Early Adopters	Research groups related with EXTRACT
Adopters' problems/needs	- Create and configure Kubernetes clusters. - Create a COMPSs Docker image. - Install the CLI. Everything you need is open source.
Alternative solution	Currently, the only alternative if you want to use COMPSs is to execute in Docker in a single machine. There are alternatives to COMPSs that can be deployed in Kubernetes, but don't have explicit support for multicluster configurations (e.g. Ray, Apache Spark).
Unique Value Proposition	The main competitors are Ray, Apache Dask, Apache Spark, and Apache Flink, which all provide distributed execution frameworks but with different design goals.

	Ray is strong in AI and machine learning workloads and offers simple APIs for parallel execution. Apache Dask and Apache Spark are widely used for data analytics and data-parallel workloads. They provide mature ecosystems and good performance for large-scale data processing, but they require applications to fit their data-centric models and are less suitable for generic task-based workflows. Apache Flink excels at stream processing and low-latency data pipelines. Its strengths lie in continuous data streams rather than batch or irregular task-based workflows, making it less flexible for general scientific workflows. Multi-cluster deployments typically require additional manual configuration and operational effort when deploying these frameworks.
Competitors	The main competitors are Ray, Apache Dask, Apache Spark, and Apache Flink, which all provide distributed execution frameworks but with different design goals. Ray is strong in AI and machine learning workloads and offers simple APIs for parallel execution. Apache Dask and Apache Spark are widely used for data analytics and data-parallel workloads. They provide mature ecosystems and good performance for large-scale data processing, but they require applications to fit their data-centric models and are less suitable for generic task-based workflows. Apache Flink excels at stream processing and low-latency data pipelines. Its strengths lie in continuous data streams rather than batch or irregular task-based workflows, making it less flexible for general scientific workflows. Multi-cluster deployments typically require additional manual configuration and operational effort when deploying these frameworks.
Partners	BSC
Timing	COMPSs users are expected to adopt it during 2026.
IP Strategy	Since the result is written code, it will be protected by copyright and distributed under a permissive open source license.

Exploitation roadmap for Kubecompss	
Actions	The result requires further development before its full potential can be realized and commercial exploitation considered. Currently, it has a permissive open-source license that will allow for faster technological development. Contact is already underway with various groups interested in using the result. Work will continue on identifying these groups with the aim of developing a product that is useful to as many people as possible.
Roles	BSC is the leading developer of COMPS and Kubernetes and will be responsible for their protection, operation and sustainability
Milestones	Number of research groups using the results.
Costs	The results have already been transferred to the FORGE project to facilitate further development. It is expected that the project will advance by one or two TRL levels and that the number of research groups testing it will increase.
Revenues	The outcome is not expected to generate financial revenue. However, it is expected to bring about a significant improvement in efficiency in line with the latest developments, which in turn will result in lower costs for the centres that adopt it, thereby reducing the cost of future developments.
Other sources of coverage	BSC has an interest in utilising the results in its own operations. This translates into a guarantee that the results will be utilised, which, together with any improvements that may be developed, will ensure their long-term use

KER name: Lithops	
	Input from the Beneficiary
Description	Lithops is an open-source serverless computing framework designed to simplify the execution of large-scale, data-intensive workloads across multiple cloud providers. It

	executes tasks in parallel using serverless and managed compute services (e.g., AWS Lambda, Google Cloud Functions, Azure Functions, Kubernetes, virtual machines). Lithops automatically handles scaling and backend provisioning with almost zero configuration, allowing users to focus on application logic rather than infrastructure management.
Target market/end users	**Target market** • Cloud computing • Serverless data analytics • Scientific computing and data science • Large-scale batch and embarrassingly parallel workloads **Customers / End-users** • Researchers and scientists running data-intensive experiments • Data scientists and ML engineers executing large batch jobs • Cloud engineers and developers building scalable analytics pipelines • SMEs and startups requiring elastic compute without managing clusters • Academic institutions and research infrastructures
End-users needs / problems	End users face several recurring problems: • Difficulty scaling compute resources dynamically for bursty or short-lived workloads • High operational overhead when managing clusters (Spark, Dask, Ray, Kubernetes) • Over-provisioning and underutilization of cloud resources • Complexity in deploying parallel workloads across heterogeneous cloud backends • Lack of portability across cloud providers Lithops addresses the need for **simple, scalable, and cost-efficient parallel computing** without requiring deep expertise in cloud infrastructure.
Competitive advantages	Lithops provides a true serverless abstraction with no cluster management, automatic massive parallelism, and fine-grained pay-per-execution billing. It is cloud-agnostic, supports multiple backends, offers a Python-native API, and is highly efficient for embarrassingly parallel workloads, significantly reducing setup time, operational complexity, and idle costs compared to traditional big data frameworks.
Use model	Lithops is made available as open-source software under a permissive license, enabling direct adoption by industry and research organizations. It is used through public cloud infrastructures following a pay-per-use model and can be transferred to industrial and academic environments. Additional value is created through integration services, consulting, training, and scientific dissemination via publications and benchmarks. There is also one production platform called PyRun (https://pyrun.cloud) which makes (even more) easier to use Lithops.
Early Adopters	Early adopters include research groups with bursty computational needs, data science teams running large batch analytics, cloud-native startups optimizing costs, and HPC users transitioning to cloud-based execution models. These adopters can be reached through open-source communities, scientific and cloud-native conferences, technical blogs, tutorials, and academic–industry collaborations.
Adopters' problems/needs	Adopters need faster experimentation cycles, elastic compute without infrastructure management, predictable and reduced cloud costs, and seamless integration with existing Python-based workflows. They often seek alternatives to complex cluster-based solutions that slow down development and increase operational risk.
Alternative solution	Current alternatives include managed cluster services such as Spark on EMR, distributed frameworks like Dask or Ray deployed on Kubernetes, traditional HPC batch schedulers, and custom orchestration built directly on cloud APIs. While functional, these solutions typically require significant setup, tuning, and ongoing maintenance.
Unique Value Proposition	Lithops offers clusterless parallel computing that combines the simplicity of local Python programming with the elasticity and cost efficiency of serverless cloud execution. It enables

	users to run large-scale parallel workloads with minimal configuration, no infrastructure management, and strong portability across cloud providers.
Competitors	Competing solutions include Apache Spark, Dask, Ray, managed cloud analytics services, and raw serverless platforms. While these alternatives offer scalability and mature ecosystems, they often suffer from high operational overhead, vendor lock-in, or development complexity. Lithops differentiates itself by unifying serverless execution and distributed computing in a lightweight and flexible framework.
Partners	URV
Timing	TRL6
IP Strategy	Lithops follows an open-source IP strategy under the Apache 2.0 license, promoting wide adoption and reuse. Value is generated through community, research, and services rather than patents.

Exploitation roadmap for Lithops	
Actions	Ongoing maintenance of the Lithops codebase, continued use in scientific publications, and initial go-to-market efforts for PyRun.cloud as a commercial entry point.
Roles	URV is the sole responsible partner, leading development, research, and commercialization efforts.
Milestones	Consolidate users in pyrun.cloud platform
Costs	Lithops/pyrun.cloud project will require a full time engineer 50,000€ developing/supporting the platform per year.
Revenues	pyrun.cloud has no pricing defined yet, but the key idea is to earn money via registrations/usage/consultoring the platform.
Other sources of coverage	URV research budget and future European research projects.

KER name: DataPlug	
	Input from the Beneficiary
Description	DataPlug facilitates seamless data integration between different systems and platforms. It acts as a bridge that connects disparate data sources, ensuring smooth data flow and compatibility. This tool is particularly useful in environments where multiple data formats, sources, or services need to interact, helping to automate the process of data extraction, transformation, and loading (ETL) while maintaining data integrity across the compute continuum.
Target market/end users	DataPlug targets data-intensive computing environments where large volumes of heterogeneous and unstructured data must be processed efficiently. Its primary market includes scientific computing infrastructures, cloud and hybrid-cloud platforms, and organizations operating large-scale analytics and AI workloads. The main end users are data engineers, research teams, and platform operators who need to integrate diverse data sources with distributed computing frameworks while maintaining scalability and performance.
End-users needs / problems	End users increasingly face challenges related to the ingestion, preprocessing, and efficient access of unstructured data. Traditional ETL pipelines introduce significant latency, high preprocessing costs, and strong coupling between data layout and analytics frameworks. These limitations hinder elasticity, slow down experimentation, and prevent efficient exploitation of distributed and serverless compute resources. DataPlug addresses these needs by enabling flexible, low-overhead data access and integration across heterogeneous environments.
Competitive advantages	DataPlug offers a significant advantage over existing solutions by enabling dynamic, on-the-fly data partitioning without requiring data transformation or rewriting. Its ability to operate directly on unstructured data and adapt partitioning strategies at runtime allows workloads to scale efficiently while minimizing preprocessing costs. The framework is designed for extreme-scale analytics and is natively compatible with modern distributed computing paradigms, which distinguishes it from static and platform-specific data integration solutions.

Use model	DataPlug is made available as an open-source software framework that can be adopted by research institutions, infrastructure providers, and industrial users. It can be deployed as a client-side library integrated into existing analytics pipelines or offered as part of data processing services in cloud and HPC environments. The framework also serves as a vehicle for technology transfer, research dissemination, and future standardization activities in data-intensive computing.
Early Adopters	Early adopters of DataPlug include research groups, HPC centers, and cloud-native organizations that work with extreme-scale or unstructured data and experience strong limitations with existing data integration approaches. These adopters can be reached through scientific publications, open-source communities, technical demonstrations, and collaborations within European research and innovation projects.
Adopters' problems/needs	Adopters struggle with inefficient data pipelines and high preprocessing overhead. They require a solution that allows data to be reused across multiple workloads and infrastructures without repeated preprocessing or data reorganization. DataPlug directly addresses these challenges by providing a flexible, reusable, and infrastructure-agnostic data integration layer.
Alternative solution	Current alternatives include static ETL pipelines, data lakes, warehouses, and cloud-specific integration services. These often require upfront transformations, are hard to scale, and can be costly to maintain. Many organizations end up building ad-hoc solutions that create further complexity.
Unique Value Proposition	DataPlug enables analytics on unstructured data without repeated preprocessing. Its dynamic partitioning adapts to workloads, lowers costs, boosts performance, and works across diverse infrastructures.
Competitors	DataPlug's competitors are general-purpose frameworks and cloud-native ETL or ingestion services. These are widely adopted but rely on static data layouts and struggle with unstructured or extreme-scale datasets. DataPlug differentiates itself by supporting unstructured data, offering runtime adaptability, and operating smoothly across heterogeneous compute environments.
Partners	URV
Timing	TRL5
IP Strategy	DataPlug's core components are released under the Apache 2.0 license, ensuring open-source availability while allowing contributions and commercial use.

Exploitation roadmap for DataPlug	
Actions	Ongoing maintenance of the DataPlug codebase, submission of a journal publication showcasing its capabilities and results obtained during the EXTRACT project. DataPlug is also part of the PyRun.cloud stack, contributing to its commercial roadmap alongside Lithops.
Roles	URV is the sole responsible partner, leading development and research dissemination.
Milestones	Journal paper submitted within 6 months after project end. New implementations for new formats.
Costs	~€25,000 (part-time development and dissemination) per year.
Revenues	Economical exploitation via pyrun.cloud. pyrun.cloud has no pricing defined yet, but the key idea is to earn money via registrations/usage/consultoring the platform.
Other sources of coverage	URV research budget and future European research projects.

KER name: TASKA-A	
	Input from the Beneficiary

Description	TASKA-A allows us to deploy real time advanced processing pipelines (including ML) directly in the instrument (edge) and provides observers with various types of physical events detected in the real time data flow.
Target market/end users	The main end users are research engineers, research teams, telescope operators, and students, mainly for reducing data volume in jovian and solar high resolution beamformed observations.
End-users needs / problems	Modern radiointerferometers produce giant datasets that are complicated to store, to move and to reduce. The final scientific products are an observation file with a lower resolution and a complementary file with portions of interest in the observation at high resolution. Currently the processing of these observations is done afterwards and the data storage space limits the observation possible if not post-processed fast enough. End-users need a tool that processes data on the go, managing storage issues in real-time.
Competitive advantages	TASKA-A allows complex signal processing on a real-time data stream. The MurMuRe interface that was originally developed for NenuFAR and this use case is now working on several other instruments. This means that future detection algorithms could be adapted to work in real-time easily, or to adapt existing algorithms to other instruments.
Use model	TASKA-A packages are made available as open source codes that can be adopted by research institutions for both on-disk data or real-time data streams coming from the telescope.
Early Adopters	Same as end-users
Adopters' problems/needs	Same as end-users
Alternative solution	There is no current alternative solution as the main objective of this use case was to create the "real-time processing" feature which is new. Alternative would be to use a pre-existing algorithm to post process on-disk data.
Unique Value Proposition	TASKA-A is unique in the sense that there is currently no other real-time processing tool of this sort for the NenuFAR radio telescope.
Competitors	Other scientific teams in different research facilities are working on other real-time detection algorithms.
Partners	OBS
Timing	TRL5 ?
IP Strategy	TASKA-A are published on an open source licence (MIT)

Exploitation roadmap for TASKA use case A	
Actions	Maintenance of the TASKA-A codebase, deployment of the framework in production at Nançay Radioastronomy Observatory (NenuFAR), Observatoire de Paris (OBS).
Roles	OBS is the sole responsible scientific partner, leading scientific exploitation and research dissemination.
Milestones	First Journal paper within 6 months after the end of the project to demonstrate performances. Full deployment for production of outputs with NenuFAR radiotelescope (2026). Adaptation to other radiotelescope.
Costs	~5k€ / per year / per paper (Engineering student internship and non waived publication costs)
Revenues	None
Other sources of coverage	OBS research budgets and future european/national research projects and funding.

KER name: TASKA-C

	Input from the Beneficiary
Description	TASKA-C is a workflow framework for radio astronomy interferometric imaging. It uses EXTRACT libraries and software to optimize, catalog and orchestrate the workflow, on cloud and HPC computing frameworks.
Target market/end users	The main end users are research engineers, research teams, telescope operators, and students, mainly for imaging astrophysical sources from any interferometric dataset.
End-users needs / problems	Modern radiointerferometers produce giant dataset that are complicated to store, to move and to reduce. The final scientific products are small images cubes requiring the data cleaning, reduction and calibration of these massive datasets. Current struggle to perform all these tasks (often done manually) slow down the process the scientific production and waste useful scientific time. End-users needs a seamless reduction framework to reduce the clutter and computing time.
Competitive advantages	TASKA-C enable the automatic and flexible run of data processing workflows of heavy data set that are distributed in multiple sites. The unique features that TASKA-C provide are: - **Workflow composability**: made easier on the client side to compose a large variety of various workflow for different tasks - **Ressource optimization**: thanks to dataplug and Lithops: the ressource provisionning and data partitionning helps to dynamic adjust to the requirements of the input data set in order to reduce computing waste as well as processing time. - **Workflow dynamics**: capability to insert quality gates that will suspend the workflow for human verification, based on quality criteria or decision gates. Capability to cancel and redo portion of the workflow (as a "goto" function), even in the situation of a parallel workflow graph. - **Seamless user interface:** the user run workflows and use data "as if" computing ressources and data were local whereas both can be spread in various CLOUD/HPC/Storage centres. - **Adaptation to other science cases**: the framework prepared in the context of EXTRACT can also serve to other domain of science involving heavy data reduction.
Use model	TASKA-C software is made available as an open-source software framework that can be adopted by research institutions, infrastructure providers, and industrial users. It can be deployed as a client-side library integrated into existing analytics pipelines or offered as part of data processing services in cloud and HPC environments. The software is also organized to allow further development for new type of source or other telescopes.
Early Adopters	Same a End-users
Adopters' problems/needs	Same a End-users
Alternative solution	Current alternatives solutions involve workflow on multiple and heterogeneous systems, each attached to a given particular radio telescopes. Contrary to the EXTRACT platform, there is no coordinated workflow that could connect to multiple computing centers and S3 storage nodes.
Unique Value Proposition	TASKA-C is unique in the sense of it is the only framework that could adapt to virtually -any- science case. Each individual task can be any scientific task (using community tools) of any scientific domain involving big data. The source code also provide tools for developing new flavor of workflow in a different scientific environment.
Competitors	Other scientific and data processing teams associated with other radio telescopes (ASTRON, SKA office in South Africa, SKA-LOW in Australia) who are also developing similar tools.
Partners	OBS, URV, BSC
Timing	TRL1
IP Strategy	The source code is published under a open source license (Apache 2.0)

Exploitation roadmap for TASKA use case C	
Actions	Maintenance of the TASKA-C codebase, deployment of the framework in production at Nançay Radioastronomy Observatory (NenuFAR), Observatoire de Paris (OBS), BRGM (Orléans) and possibly on EOSC infrastructure. EXTRACT TASKA-C is also a very good candidate for being a solution for the upcoming French SKA Regional Centre deployed on multiple heterogenous site in France.

Roles	OBS is the sole responsible scientific partner, leading scientific exploitation and research dissemination. URV and BSC are the main developers to add new feature not already existing in the current version.
Milestones	First Journal paper with 6 months after the end of the project to demonstrate performances. Full deployment for production of images with NenuFAR radiotelescope (2026). Presentation as a solution for the upcoming French SKA Regional Center (>2027).
Costs	~5k€ / per year / per paper (Engineering student internship and non waived publication costs)
Revenues	None
Other sources of coverage	OBS research budgets and future european/national research projects and funding.

KER name: TASKA-D	
	Input from the Beneficiary
Description	TASKA-D is a new imager based on trained deep learning networks to image radio interferometric data. The network is fully compatible with the community standard and in specialized in detecting and imaging variable sources.
Target market/end users	The main end users are research engineers, research teams, telescope operators, and students, mainly for imaging astrophysical sources from any interferometric dataset.
End-users needs / problems	Classic radiointerferometric imaging assumes a "steady sky" composed of astrophysical sources that do not varies in the scope of the human life. However, some strong sources might vary (the Sun, flare stars, planets & exoplanets, microquasars, pulsars, etc.) and current imaging software can only do frame-by-frame imaging with limited signal-to-noise ratio.
Competitive advantages	This new imager is based on the throughout learning of the instrumental response as well as training data containing variable sources.
Use model	TASKA-D software is made available as an open-source software framework that can be adopted by research institutions, infrastructure providers, and industrial users. It can be deployed as a client-side library integrated into existing analytics pipelines or offered as part of data processing services in cloud and HPC environments. The software is also organized to allow further development for new type of source or other telescopes.
Early Adopters	Early adopters of TASKA-D include research groups, HPC centers, radio observatories and research laboratories computing infrastructures that work with massive amount of radio interferometric data that can be processed locally.
Adopters' problems/needs	Early adopters already struggle with a lot of radio interferometric data that contains information about a static sky. In order to search for transient sources in a large data volume, they would need to process again and again the same data at various time resolutions to capture 1s, 1min, 1hour varying sources. With TASKA-D software, the data need to be visited once as it is meant to detect variable sources in the data in one pass.
Alternative solution	Current alternative solutions are performing improved frame-by-frame imaging but do not incorpore time as a dedicated data axis.
Unique Value Proposition	TASKA-D software is designed to learn the specific instrumental properties of radiointerferometers that is used for training. The TASKA-D solution provide: i) sky model simulation modelling a dynamic sky, ii) AI training tools to adapt the imager to each radiotelescope features and data, and iii) the imager to do the inference on real interferometric data.
Competitors	Other scientific and data processing teams associated with other radio telescopes (ASTRON, SKA office in South Africa, SKA-LOW in Australia) who are also developing similar tools.
Partners	OBS
Timing	TRL1
IP Strategy	The source code is published under a open source license (Apache 2.0)

Exploitation roadmap for TASKA use case D	
Actions	Maintenance of the TASKA-D codebase, multiple submission of journal publication: i) proof of concept on detecting fixed transient (EXTRACT results), ii) study of moving transient and iii) extended transients. Possibly a future PhD program based on this imager.
Roles	OBS is the sole responsible partner, leading development and research dissemination
Milestones	First Journal paper with 6 months after the end of the project. Second submitted before the end of 2026. Deployment and application to real data streams (NenuFAR) in the course of 2026.
Costs	~5k€ / per year / per paper (Engineering student internship and non waived publication costs)
Revenues	None
Other sources of coverage	OBS research budgets and future european/national research projects and funding.

KER name: UDT	
	Input from the Beneficiary
Description	An Urban Digital Twin (UDT) is a dynamic digital model that continuously collects data from multiple heterogeneous sources, such as sensors, mobility systems, simulations, and external data services. It processes this data to create a real-time state of the city, representing how the urban environment is behaving at a given moment. This real-time city state can be consumed by other components to perform advanced computations, such as intelligent routing, emergency management, and optimisation tasks.
Target market/end users	**Target market** The Urban Digital Twin targets the public sector and urban innovation market, in particular cities and regional authorities adopting data-driven approaches for urban management, resilience, and planning. It is also relevant for research and innovation projects and for organisations developing advanced smart-city and decision-support solutions. **Customers / End users** Municipalities and regional authorities Civil protection and emergency management agencies Urban planning and mobility departments Smart City platform operators and system integrators Research institutions and innovation labs working on urban analytics and simulation
End-users needs / problems	Urban authorities and organisations face increasing difficulty in understanding and managing complex cities in real time. Data relevant to urban operations is typically fragmented across multiple systems, sources, and departments, making it hard to obtain a coherent and up-to-date view of the city. **End users need:** - A unified and real-time representation of the city state built from heterogeneous data sources - The ability to share a consistent city state across multiple services and applications - Support for advanced analytics, simulation, and optimisation, rather than simple monitoring

	- A flexible foundation that can be extended with new capabilities (e.g. intelligent routing, emergency response, "what-if" analysis) without redesigning the whole system The UDT addresses these needs by transforming raw and heterogeneous data into a usable, consistent, and continuously updated city state that can support decision-making and downstream computations.
Competitive advantages	The Urban Digital Twin offers a set of advantages over traditional urban monitoring and planning solutions: - Real-time city state: instead of static maps or periodic reports, the UDT maintains a continuously updated representation of the city, reflecting current conditions as they evolve. - Integration of heterogeneous data: the UDT natively combines data from multiple sources (sensors, mobility data, simulations, external services) into a single coherent model, reducing fragmentation and manual data handling. - Computation-ready model: the city state produced by the UDT is designed to be directly consumed by other components (e.g. intelligent routing, optimisation, emergency management), not only visualised. - Extensibility and modularity: new capabilities such as "what-if" analysis, prediction models, or learning-based components can be added without redesigning the system. - Domain-agnostic foundation: unlike many existing solutions focused on a single domain (traffic, energy, or GIS), the UDT provides a general-purpose foundation that can support multiple urban use cases. Overall, the UDT moves beyond dashboards and siloed digital twins by acting as a shared, real-time computational backbone for advanced urban decision-support systems.
Use model	The Urban Digital Twin is made available as a software-based platform and reference architecture that can be deployed and customised for different urban contexts. The main use model includes: - Deployment in pilot cities and public-sector environments as a decision-support backbone for urban management and resilience - Technology transfer to municipalities, public authorities, and system integrators for integration with existing city platforms - Use within research and innovation projects as a reusable framework for developing and testing advanced urban services - Provision of services, including system integration, customisation, training, and technical support for end users -
Early Adopters	Early adopters are **municipalities and public authorities** facing complex urban management and resilience challenges, particularly in Smart City and emergency preparedness contexts. They also include **civil protection agencies** and **urban innovation and research centres** experimenting with real-time city models.
Adopters' problems/needs	Adopters such as municipalities and public authorities need tools that help them operate and plan cities using real-time data, rather than fragmented and static information systems. They face difficulties in integrating heterogeneous data sources, sharing a consistent city state across services, and supporting advanced analytics and decision-making. Adopters need a reusable and extensible digital foundation that can be integrated into existing city platforms and can support multiple use cases, such as emergency management, mobility optimisation, and urban planning, without high integration or maintenance costs.
Alternative solution	At present, adopters mainly rely on static GIS systems, dashboards, and isolated domain-specific platforms (e.g. traffic or infrastructure monitoring tools).

	Data from different sources is often integrated manually or through ad-hoc scripts, with limited real-time capabilities. Advanced simulations and analytics are typically performed offline and separately from operational systems, making it difficult to reuse results in real-time decision-making. This fragmented approach increases complexity, limits scalability, and reduces the effectiveness of urban management and emergency response.
Unique Value Proposition	The Urban Digital Twin provides adopters with a single, real-time and computation-ready city state that can be shared across multiple services and applications. Unlike current fragmented solutions, it acts as a common digital backbone rather than a standalone tool or dashboard. By enabling real-time analytics, simulation, and optimisation on top of the same city state, the UDT allows cities to move from static monitoring to dynamic, data-driven decision-making, while remaining flexible and extensible for future use cases.
Competitors	Competitors include commercial Smart City platforms, GIS-based systems, and domain-specific digital twins focused on areas such as traffic, energy, or infrastructure monitoring. These solutions are generally mature and well-integrated within their specific domains. However, most competing solutions are siloed, static, or limited to visualisation, and they often lack a unified real-time city state that can be reused across multiple services. In contrast, the UDT provides a cross-domain, extensible, and computation-ready foundation, enabling advanced analytics, simulation, and optimisation beyond traditional dashboards and single-purpose digital twins.
Partners	LOGOS
Timing	Time to market not defined at this stage
IP Strategy	IP Strategy not defined at this stage

KER name: USE case PER	
	Input from the Beneficiary
Description	The PER use case is focused on providing real-time, personalized evacuation routes for citizens in Venice during emergencies. By utilizing advanced data collection and processing technologies, such as a UDT, the system can simulate evacuation scenarios and optimize safe routes. This ensures more efficient and coordinated evacuations. A MARL dynamically calculates and updates routes based on real-time conditions, while a Simulator helps test strategies under different emergency situations. Together, these technologies improve crisis management and citizen safety in complex urban environments.
Target market/end users	The PER solution addresses the market of **digital emergency-management systems** and **smart city resilience tools**. It is designed for urban contexts with high population density, significant tourist flows, and exposure to environmental risks. The primary market includes public authorities seeking data-driven platforms to support evacuation planning, real-time decision-making, and crisis coordination. **Institutional customers:** Metropolitan City of Venice, Municipality of Venice, Civil Protection agencies, and transport/territorial management bodies. These actors use the system to simulate scenarios, plan evacuation strategies, and coordinate emergency responses. **Final end-users:** Residents, commuters, and tourists. They benefit from real-time, personalized evacuation routes that enhance safety and improve the effectiveness of emergency operations.
End-users needs / problems	Urban areas like Venice face increasing challenges in managing emergency evacuations due to complex mobility patterns, high tourist density, and rapidly evolving risk conditions. Public authorities lack tools capable of providing **real-time, personalized, and dynamically updated evacuation routes** that account for changing scenarios, congestion, and environmental constraints. End-users—residents, commuters, and tourists—need **clear, timely, and context-aware guidance** to move safely during

	emergencies, while institutional actors require **data-driven decision-support systems** to coordinate responses efficiently. The PER KER addresses this need by integrating advanced data processing, simulation, and multi-agent routing to deliver optimized evacuation strategies that enhance safety and operational effectiveness.
Competitive advantages	The PER KER outperforms existing evacuation and routing solutions by providing **real-time, personalized, and dynamically updated evacuation routes**, unlike static plans or generic navigation tools. Its integration of **multi-agent reinforcement learning**, continuous data ingestion, and high-fidelity simulations enables more accurate and adaptive decision-making. The system also offers **scenario-based testing** for authorities and **individualized guidance** for citizens, improving both preparedness and operational efficiency. Its interoperability with urban data ecosystems further enhances coordination across agencies, distinguishing it from current market alternatives.
Use model	The PER KER will be deployed as a **digital service** integrated into existing urban-data platforms and emergency-management systems. Its use model is based on the **provision of a decision-support service** for public authorities, complemented by **training activities** for operational staff. The technology can be made available through **licensing agreements**, **technology transfer**, or **service contracts** with municipalities and civil-protection agencies. Simulation tools and routing components may also be shared through **research collaborations** or **contract research**. This approach ensures scalable adoption and seamless integration into local governance workflows.
Early Adopters	Early adopters include **municipalities and civil-protection agencies** in cities with high tourist pressure, complex mobility patterns, or exposure to climate-related risks. These actors experience the evacuation-management problem more acutely and are therefore more inclined to adopt innovative, data-driven solutions. Additional early adopters may include **transport authorities** and **regional emergency-response centres** seeking integrated decision-support tools. They can be reached through **institutional networks**, targeted **workshops and training sessions**, **pilot deployments**, and dissemination via **urban-resilience platforms** and **EU innovation channels**. This approach accelerates trust building and facilitates early-stage adoption.
Adopters' problems/needs	Potential adopters—such as municipalities, civil-protection agencies, and transport authorities—struggle to manage evacuations in complex urban environments due to limited real-time information, static evacuation plans, and fragmented decision-support tools. They need solutions that can **integrate heterogeneous data**, simulate scenarios, and provide **dynamic, personalized routing** to improve coordination and reduce response times. The PER KER addresses this gap by offering an adaptive, data-driven platform that enhances operational efficiency and citizen safety during emergencies.
Alternative solution	Potential adopters currently rely on **static evacuation plans**, periodic risk-assessment documents, and manual coordination procedures that are often slow to update and not responsive to real-time conditions. Some municipalities use **generic navigation tools** or basic GIS systems, which lack predictive capabilities and cannot provide personalized routing during emergencies. Decision-making is therefore fragmented, with limited integration of heterogeneous data sources and no dynamic simulation of evolving scenarios. This results in reduced situational awareness and sub-optimal coordination during critical events.
Unique Value Proposition	The PER KER provides a major improvement over current evacuation-management practices by offering **real-time, adaptive, and personalized routing**, overcoming the limitations of static plans and non-specialized navigation tools. Its integration of **multi-agent reinforcement learning**, high-fidelity simulations, and continuous data ingestion enables more accurate predictions and faster operational decisions. Unlike existing solutions, it supports **scenario testing**, cross-agency coordination, and seamless interoperability with urban data platforms. These capabilities give adopters a **more reliable, scalable, and context-aware system**, significantly enhancing preparedness and response effectiveness in complex emergencies.
Competitors	Competitors include providers of **static evacuation-planning tools**, **generic navigation apps**, and **standard GIS-based emergency platforms**. These solutions are strong in

	basic mapping and wide user adoption but lack **real-time adaptation**, **predictive simulation**, and **personalized routing**.
Partners	Potential adopters and future partners in the Venetian territory include:**Municipality of Venice** and its operational units involved in mobility, safety, and emergency response. **Protezione Civile regionale e locale**, responsible for coordinating evacuation procedures. **ACTV / AVM Venezia**, managing public transport and mobility flows during emergencies. **Port of Venice Authority** and **airport operators**, due to their need for coordinated evacuation and crisis-management tools. **Tourism and cultural-site management bodies**, which require real-time guidance for large visitor flows in emergency situations.
Timing	The PER KER can reach the market in a **short-to-medium timeframe (1–2 years)**, as its core components are already validated within the project. Deployment mainly requires integration with existing urban-data platforms
IP Strategy	The PER KER relies on partners' **background** in data processing, simulation, and routing, while generating **foreground** software components and integration workflows jointly owned by contributors. An IP strategy is relevant to define ownership, licensing options, and exploitation rights.

Exploitation roadmap for PER Use Case	
Actions	In the 3–6 months following the project, planned actions include the **finalisation of the business and exploitation plan**, refinement of the **IP protection strategy**, and the definition of licensing or service models for adopters. Partners will work on **collecting necessary authorisations** from local authorities, validating compliance with data-governance and safety regulations, and preparing agreements for future deployments. Technical activities will focus on **stabilising the prototype**, integrating feedback from pilot tests, and ensuring readiness for operational use. Dissemination and engagement with potential adopters in the Venetian territory will also continue to support early uptake.
Roles	The **Metropolitan City of Venice** will be the main responsible organisation for the exploitation roadmap, coordinating adoption, authorisations, and integration into local governance workflows. Technical and research partners will support **prototype stabilisation**, data-governance compliance, and refinement of the operational model. Industry partners will contribute to **service packaging, licensing options, and business-plan finalisation**, ensuring readiness for market uptake and scalability across the territory.
Milestones	**Months 1–3:** Business plan, IP mapping, initial stakeholder meetings **Months 3–5:** Licensing/IP agreements, authorisations, service model definition **Months 5–6:** Prototype stabilisation, final validation, adopter engagement **Milestones (within 3–6 months post-project)** • **M1 – Business & Exploitation Plan finalised** (by Month 3) • **M2 – IP and licensing agreements defined** (by Month 4) • **M3 – Local authorisations secured** (by Month 5) • **M4 – Stabilised KER prototype released** (by Month 6) • **M5 – Engagement with at least two potential adopters** (by Month 6)
Costs	**1-Year Horizon:** Estimated costs include **€80k–€120k** for prototype stabilisation, integration with local data platforms, and initial operational testing. Additional investments of **€20k–€40k** will be required for **IP protection**, licensing preparation, and legal support. Around **€30k–€50k** will support business-plan finalisation, stakeholder engagement, and obtaining local authorisations. **3-Year Horizon:** A broader investment of **€250k–€400k** is expected to scale the KER toward full market readiness, including continuous technical upgrades, maintenance, and TRL increase. Further **€60k–€100k** may be needed for extended IP management (e.g., patenting,

	licensing models), regulatory compliance, and long-term service packaging. Engagement with adopters, training, and deployment across the Venetian territory may require an additional **€100k–€150k**.
Revenues	In the first year of adoption, expected revenues range from **€60k–€120k**, mainly from licensing, integration services, and early operational support, with modest profits (10–20%). After three years, with wider uptake across the Venetian territory and potential replication in similar cities, annual revenues could reach **€200k–€350k**, with improved profit margins (25–35%) thanks to recurring service contracts and reduced development costs.
Other sources of coverage	To cover the pre-revenue phase and increase TRL, the KER will require **€120k–€180k in the first year**, covering prototype stabilisation, IP protection, authorisations, and early deployment. These resources may come from **partners' own budgets**, **regional/national innovation incentives**, and **additional project grants**. If scaling accelerates, **risk capital or targeted loans** may complement funding. Securing these resources at the right time ensures continuity toward market readiness.

KER name: Binare's Security Toolkit (BISETO) for Compute Continuum (BISETO-CC)	
	Input from the Beneficiary
Description	The BISETO-CC is a cybersecurity support framework developed by Binare to check, follow-up and enhance security in compute continuum environments. The toolkit combines specialised tools, security procedures, and expert consulting services designed to protect complex distributed infrastructures operating across edge, cloud, and HPC environments. BISETO-CC addresses a broad set of cybersecurity domains relevant to compute continuum deployments. These include mechanisms for detecting cases where data tampering and machine learning model integrity could be possible (source code or running environments), secure versioning of software components used across the complex layers and modules within a deployment, and vulnerability/weakness discovery in code and containerised/kubernetes deployments. The toolkit also supports checks against common attack vectors such as man-in-the-middle attacks, credential stuffing, and insecure communications. Beyond the technical toolkit itself, BISETO-CC includes consulting, training, and expert assessment services that help research and industrial teams implement secure-by-design infrastructures. By combining cybersecurity expertise with compute continuum knowledge developed within projects such as EXTRACT, the toolkit supports secure deployment practices across the entire lifecycle of distributed infrastructures. This sub-component is a consulting + training service type of KER planned by Binare. It leverages the Binare Cybersecurity Toolkit for Compute Continuum Software (see above) and couples it with deep-tech know-how within Binare in both areas (cybersecurity x compute continuum such as EXTRACT) in order to help both research and industrial complex compute continuum projects and deployments receive end-to-end guidance, training and expert assessment to ensure those projects are following "secure by default" paradigm and methodologies.
Target market/end users	The primary target market includes research and industrial initiatives developing or deploying compute continuum infrastructures. This includes Horizon Europe projects, industrial edge–cloud deployments, and organisations operating distributed infrastructures that require advanced cybersecurity practices. Typical end users include research teams deploying distributed applications, industrial R&D departments building edge and cloud platforms, DevOps and DevSecOps teams responsible for infrastructure security, and project managers or technical leaders responsible for secure infrastructure governance. Additional users include organisations operating Smart City infrastructures, critical digital services, and large-scale IoT systems.
End-users needs / problems	Organisations deploying compute continuum infrastructures often struggle to ensure consistent cybersecurity across distributed environments spanning edge devices, cloud

	services, and high-performance computing resources. These infrastructures typically integrate multiple technologies, making it difficult to implement unified security policies and detect vulnerabilities across the full system lifecycle. Research and industrial teams frequently lack specialised security expertise in distributed environments, which can lead to insecure configurations, weak access control mechanisms, or vulnerabilities in containerised deployments. At the same time, new attack vectors targeting machine learning pipelines, container environments, and distributed infrastructures are becoming increasingly common. These challenges create a strong need for security frameworks that can guide teams in implementing secure-by-default infrastructures while providing practical tools and procedures for vulnerability detection, threat mitigation, and secure deployment practices.
Competitive advantages	The BISETO-CC Security Toolkit combines cybersecurity expertise, technical tools, and consulting services tailored specifically for compute continuum environments. Unlike general cybersecurity solutions, BISETO-CC focuses on the security challenges of distributed infrastructures spanning edge, cloud, and HPC systems, especially in highly complex&hybrid environments. The toolkit supports multiple security layers including vulnerability scanning, container security, secure deployment procedures, machine learning integrity checks/detections, and detections and checks against common network, penetration and authentication attacks. In addition, BISETO-CC provides structured training and expert consulting that enable research and industrial teams to adopt secure-by-design practices from the earliest phases of infrastructure development. Through this combination of tools, expertise, and training, BISETO-CC helps organisations reduce cybersecurity risks while improving the resilience and trustworthiness of compute continuum deployments.
Use model	BISETO-CC will be offered as a proprietary service and toolkit distributed through commercial licensing by Binare. The offering combines security tools with consulting, assessment, and training services designed to support organisations deploying distributed infrastructures. The toolkit can be integrated into research projects, industrial pilot deployments, and production infrastructures where advanced cybersecurity support is required. In addition to project-specific consulting, the service may include structured training programmes for development and operations teams working with compute continuum environments.
Early Adopters	Early adopters are expected to include Horizon Europe research projects deploying compute continuum architectures, industrial organisations experimenting with edge–cloud infrastructures, and technology teams responsible for securing distributed applications and containerised platforms. Additional adopters may include Smart City initiatives, research infrastructures, and digital innovation hubs exploring large-scale distributed computing deployments. These adopters will be reached primarily through the industrial network and commercial activities of Binare, as well as through Horizon Europe collaboration networks, cybersecurity conferences, and technology partnerships within the compute continuum ecosystem.
Adopters' problems/needs	Early adopters typically require structured cybersecurity guidance when deploying distributed infrastructures, particularly in environments combining edge devices, cloud services, container orchestration systems, and machine learning components. They often

	need support in identifying vulnerabilities in code and container deployments, protecting data integrity, and implementing secure communication and authentication mechanisms. In addition, organisations seek practical guidance on implementing DevSecOps practices and secure infrastructure governance across complex distributed systems. Training and expert assessment are often necessary to help teams understand emerging cybersecurity threats and implement mitigation strategies effectively. BISETO-CC addresses these needs by providing a combination of security tools, expert consulting, and training that supports secure infrastructure design and operation.
Alternative solution	Many organisations currently rely on general-purpose cybersecurity tools such as vulnerability scanners, container security platforms, or penetration testing services offered by vendors such as CrowdStrike or Palo Alto Networks. Others implement internal security practices using open-source tools integrated into their DevOps pipelines. While these solutions provide valuable security capabilities, they are typically not tailored to the specific requirements of compute continuum environments and often require significant integration effort. In many cases, organisations also rely on ad hoc consulting or internal expertise to design secure infrastructures, which can lead to inconsistent security practices across projects.
Unique Value Proposition	The BISETO-CC Security Toolkit provides a security framework specifically designed for compute continuum infrastructures. It integrates cybersecurity tools, expert consulting, and training services into a unified offering that supports the secure deployment of distributed infrastructures across edge, cloud, and HPC environments. By focusing on the security challenges of compute continuum systems, BISETO-CC enables organisations to implement secure-by-default architectures, improve vulnerability detection, and strengthen protection against emerging cyber threats. The combination of practical tools, expert guidance, and structured training enables teams to adopt security best practices while reducing operational complexity and improving infrastructure resilience.
Competitors	Potential competitors include cybersecurity consulting and platform providers such as CrowdStrike, Palo Alto Networks, and Check Point Software Technologies. These companies offer advanced security products and enterprise consulting services but typically focus on enterprise IT or cloud environments rather than distributed compute continuum infrastructures. Other competing solutions include DevSecOps toolchains and container security platforms integrated into Kubernetes environments. While these solutions provide strong technical capabilities, they often lack specialised expertise in compute continuum architectures and do not combine security tooling with dedicated consulting and training services tailored to distributed infrastructure deployments.
Partners	The BISETO-CC Security Toolkit is developed and maintained by Binare, which contributes expertise in cybersecurity, firmware security, and distributed infrastructure protection.
Timing	Medium-term (expected to reach TRL 5 in 1–2 years; early adoption in 2–3 years at TRL6+; and broader adoption in 3–4 years at TRL7+)
IP Strategy	The BISETO-CC toolkit follows a proprietary intellectual property strategy managed by Binare, and the main IPR protections methods include Copyrights and Trade Secrets type of protection. The solution will be distributed through commercial licensing and proprietary service offerings, including consulting, training, and security assessment services tailored to compute continuum deployments.

KER name: Multi-Party Computation (MPC) for Privacy-Preserving Collaborative Analytics	
	Input from the Beneficiary
Description	Multi-Party Computation is a cryptographic method that allows multiple parties to compute a result from their combined data without revealing their individual inputs to one another. It ensures privacy while allowing collaboration, and it is often used in scenarios where sensitive data from different entities needs to be processed securely. Within the project, the proposed solution consisted of a research prototype integrating MPC capabilities into a distributed analytics workflow. The implementation was based on an existing open-source MPC library, which was integrated and evaluated in the project context rather than developing a new MPC framework from scratch. The work included benchmarking activities aimed at evaluating the performance and feasibility of MPC protocols in representative collaborative data analytics scenarios. The resulting prototype reached Technology Readiness Level (TRL) 4 (technology validated in laboratory environment) and Market Readiness Level (MRL) 3, reflecting an early-stage technology validation and initial exploration of potential application contexts.
Target market/end users	The target market includes organizations that need to collaboratively analyse sensitive data while preserving confidentiality, particularly in environments where data sharing is restricted due to privacy, regulatory, or competitive constraints. Relevant sectors include: • Public authorities and smart city operators, requiring privacy-preserving analytics across multiple departments or municipalities. • Critical infrastructure and emergency management systems, where sensitive operational data must be processed securely. • Healthcare and biomedical research, where patient data cannot be directly shared between institutions. • Financial institutions, requiring secure joint risk analysis or fraud detection. • Industrial data spaces and collaborative AI platforms, where companies need to compute shared insights without revealing proprietary data. End users include: • Data analysts and AI engineers working with sensitive distributed datasets • Public administrations and smart city operators • Research organizations and healthcare institutions • Companies participating in secure data sharing ecosystems or data spaces Although the project work remains at an early validation stage, these sectors represent potential future application domains for MPC-based privacy-preserving analytics solutions.
End-users needs / problems	Organizations increasingly need to extract insights from distributed datasets owned by different parties, but data sharing is often limited by: • Privacy regulations (e.g., GDPR) • Confidentiality constraints • Commercial sensitivity of data • Lack of trust between collaborating entities

	As a result, many potential collaborative analytics initiatives cannot be implemented because participants are not willing or legally allowed to share raw data. Multi-Party Computation addresses this challenge by enabling joint computation on distributed data while ensuring that each party keeps its input data private. The project investigated this capability by experimentally validating MPC-based approaches within a controlled laboratory environment.
Competitive advantages	The MPC-based approach developed in the project offers several advantages compared to conventional collaborative analytics solutions: • Strong privacy guarantees, as raw data never leaves the owner's infrastructure. • Cryptographic protection of intermediate computation results, preventing data leakage during processing. • Compliance with data protection regulations, supporting privacy-by-design approaches required by frameworks such as GDPR. • Applicability to distributed environments, where multiple organizations need to jointly compute results without trusting a central authority. Compared to traditional centralized data aggregation solutions, MPC enables secure collaboration without requiring data pooling or full trust among participants. However, the current results correspond to a prototype validation (TRL4) and further development would be required to reach operational deployment.
Use model	The technology was developed and evaluated as a prototype integration of MPC capabilities within distributed analytics workflows. Possible exploitation models include: • Integration into secure data processing platforms used by public authorities and organizations handling sensitive data. • Licensing of the MPC-based solution as a software module enabling privacy-preserving analytics in distributed environments. • Technology transfer and integration into AI and data space infrastructures. • Provision as a secure analytics service for organizations that need privacy-preserving collaborative computations. At the current stage, the work mainly demonstrates technical feasibility rather than a fully deployable service.
Early Adopters	Early adopters are expected to be organizations that face strong regulatory or confidentiality constraints when sharing data but still require collaborative analytics capabilities. Examples include: • Public administrations and smart city operators • Healthcare institutions and research consortia • Organizations participating in European data spaces • Critical infrastructure operators • Security and emergency management agencies

	These stakeholders represent potential early adopters once the technology progresses beyond the current prototype stage.
Adopters' problems/needs	Potential adopters require solutions that allow them to: • Collaborate on data analysis without sharing sensitive raw data • Ensure compliance with privacy and data protection regulations • Maintain confidentiality of proprietary or regulated information • Enable secure cross-organizational analytics Existing solutions often require centralized data storage or trusted intermediaries, which are not acceptable in many regulated or competitive environments.
Alternative solution	Currently, organizations typically address these challenges through: • Centralized data aggregation platforms • Data anonymization or pseudonymization techniques • Trusted third-party data processors • Federated learning approaches These approaches often involve privacy risks, regulatory limitations, or trust issues between participating entities.
Unique Value Proposition	The MPC-based approach investigated in the project demonstrates the potential to enable secure collaborative analytics without requiring data sharing. Compared to existing solutions, the approach offers: • Cryptographically guaranteed data confidentiality • Privacy-preserving joint computation across organizations • Elimination of the need for trusted central intermediaries • Support for privacy-by-design data processing architectures The project contribution mainly lies in demonstrating and benchmarking the feasibility of MPC integration within distributed analytics scenarios using an existing open-source MPC framework.
Competitors	Potential competitors include providers of privacy-preserving computation frameworks and secure analytics platforms, such as: • MP-SPDZ • Sharemind • Duality Technologies • SecretFlow Strengths of existing solutions include mature cryptographic implementations and active developer communities. However, many require significant expertise to deploy, high computational overhead, or complex infrastructure integration.

	The project work therefore focused on evaluating and benchmarking such MPC technologies in practical integration scenarios rather than developing a completely new framework.
Partners	MATHEMA
Timing	Technology maturity reached in the project: • TRL4 – Technology validated in laboratory environment • MRL3 – Early market validation and exploration
IP Strategy	The exploitation strategy considers both background and foreground intellectual property. Background IP includes existing open-source MPC libraries and cryptographic protocols used for implementation and benchmarking. Foreground IP may include: • Software components implementing the MPC integration within the EXTRACT architecture • Optimizations for distributed analytics workflows • Integration methods with edge and data platform infrastructures • Benchmarking results and evaluation methodologies produced during the project The IP strategy prioritizes software ownership by the contributing partner (MATHEMA), while enabling technology transfer, licensing opportunities, and potential open-source contributions.

KER name: Device Simulator	
	Input from the Beneficiary
Description	The simulator is the engine that drives the movement and actions of virtual people within the Urban Digital Twin. It replicates real-world behaviours to help us study and optimize urban scenarios. It is also helpful to simulate a large amount of scenarios in order to train our MARL.
Target market/end users	The simulator operates in the market of **Urban Digital Twin platforms**, more specifically in the intersection of: 1. **Smart city software and urban intelligence systems** 2. **Mobility and transport simulation platforms** 3. **Decision-support systems for public-sector planning** 4. **Synthetic environment generation for AI/ML training**, especially **multi-agent reinforcement learning (MARL)** 5. **Scenario analysis and resilience planning tools** for complex urban environments In practical terms, this is not a generic simulation engine for games or visualization. It belongs to the category of **high-fidelity urban behavioural simulation infrastructure**,

	where the core value lies in reproducing the dynamics of human movement, agent interactions, and system-level emergent effects inside a digital representation of a city.
End-users needs / problems	Urban planners, mobility engineers, and infrastructure operators often rely on **static models, simplified traffic simulations, or historical datasets** when planning interventions. While these approaches can capture macro-level flows, they typically fail to represent **fine-grained behavioural dynamics**, such as individual movement decisions, crowd interactions, adaptive responses to policy changes, or emergent collective behaviours. This creates a significant gap between **predicted outcomes and real-world behaviour**. When changes are implemented—such as new mobility policies, infrastructure redesigns, public transport modifications, or emergency response strategies—the actual system response may differ substantially from expectations. Testing urban interventions directly in the physical environment is **expensive, time-consuming, and often politically or operationally risky**. Examples include: • redesigning pedestrian or traffic flows in dense urban areas • modifying transport schedules or mobility regulations • implementing crowd-management strategies during large events • evaluating evacuation procedures or emergency scenarios • deploying new adaptive traffic control policies Because these interventions affect real populations and infrastructure, **trial-and-error experimentation in the real world is rarely feasible**. As a result, decision-makers often operate under high uncertainty when introducing changes. Another critical need is the ability to **simulate and compare a large number of potential scenarios**. Urban systems are sensitive to many variables, including population behaviour, time-of-day patterns, infrastructure constraints, and policy incentives. Traditional modelling approaches struggle to efficiently explore this high-dimensional scenario space. End users need tools that allow them to: • simulate multiple policy alternatives • evaluate infrastructure redesigns before implementation • assess system performance under disruptions or extreme events • test operational strategies across varying demand patterns Without such capabilities, strategic planning remains largely reactive rather than proactive.

	With the increasing adoption of **AI-driven control and optimization systems**, another emerging problem concerns the **lack of safe and realistic environments in which to train these algorithms**. Cities and infrastructure operators are increasingly adopting **Urban Digital Twin frameworks** to integrate spatial data, sensor streams, and predictive analytics. However, many digital twin platforms still lack a robust behavioural simulation layer capable of reproducing how populations move and interact within the city. End users therefore need **simulation engines that can populate the digital twin with realistic virtual agents**, enabling them to transform static digital representations into **dynamic, behaviour-driven analytical environments**. The simulator addresses these challenges by providing an engine capable of **replicating the movement and actions of virtual people within an Urban Digital Twin**, enabling the realistic modelling of human behaviour and interactions in complex urban environments. By enabling **large-scale behavioural simulation and scenario exploration**, the system allows end users to: • test urban interventions before real-world deployment • evaluate the impact of policy or infrastructure changes • analyse crowd and mobility dynamics under different conditions • simulate emergency and disruption scenarios • generate synthetic environments for training MARL systems In this way, the simulator reduces uncertainty, lowers experimentation risks, and supports **data-driven decision making for the design, optimization, and management of urban systems**.
Competitive advantages	The KER provides a **high-fidelity behavioural simulation engine** designed to model the movement and interactions of virtual individuals within an Urban Digital Twin. Compared to existing urban modelling and simulation solutions, the KER introduces several technical and operational advantages that significantly improve the ability of end users to analyse, predict, and optimize complex urban systems.
Use model	The KER will be put into use through a **multi-channel exploitation strategy**, combining: • **software licensing and integration** within Urban Digital Twin platforms • **provision of simulation and scenario analysis services** for cities and infrastructure operators • **AI training environment services** for multi-agent reinforcement learning applications • **technology transfer and contract research collaborations**

	• **scientific dissemination and standardization activities** This diversified use model enables the KER to reach different categories of adopters while maximizing its impact in both **urban decision-support applications and AI-enabled urban system development**.
Early Adopters	The early adopters of the KER are organizations that **experience the limitations of current urban modelling and decision-support tools most strongly** and that are already investing in **Urban Digital Twin technologies, advanced mobility management, or AI-driven urban optimization systems**. These actors typically face complex operational environments and require **high-fidelity simulation capabilities to support planning, policy evaluation, and intelligent system development**. The following groups represent the most likely early adopters. 1. Cities and metropolitan authorities implementing Urban Digital Twins Municipal governments and metropolitan authorities that are actively deploying **Urban Digital Twin platforms or smart city infrastructures** represent a primary group of early adopters. These organizations need tools capable of simulating **population behaviour, mobility dynamics, and infrastructure interactions** in order to support data-driven decision-making. These actors are early adopters because they already recognize the limitations of static modelling tools and are actively seeking **predictive simulation capabilities to support planning and policy testing**. 2. Mobility and transport authorities facing congestion and system optimization challenges Public transport operators and mobility authorities are another key group of early adopters. These organizations operate complex systems where **human mobility patterns directly affect system performance**, making behavioural simulation particularly valuable. These organizations need to **optimize mobility flows, test infrastructure or scheduling changes, and evaluate congestion mitigation strategies**, making advanced behavioural simulation tools highly relevant. 3. Engineering and mobility consulting firms Consulting and engineering firms involved in **transport modelling, urban planning, and infrastructure design** are also strong early adopters. These organizations continuously perform **scenario analysis and feasibility studies** for public-sector clients and benefit from tools that improve analytical capabilities. For these organizations, adopting the simulator provides a **competitive advantage in delivering more realistic scenario analyses to their clients**. 4. Research laboratories and AI teams developing MARL-based urban control systems Includes **research institutions and R&D teams working on AI-driven urban management systems**, particularly those using **Multi-Agent Reinforcement Learning (MARL)**.
Adopters' problems/needs	Potential adopters of the KER—such as **cities implementing Urban Digital Twins, mobility authorities, infrastructure operators, consulting firms, and research groups developing AI-based urban systems**—face significant challenges when trying to

	understand, predict, and optimize the behaviour of complex urban environments before implementing real-world interventions. One major problem is that many existing urban analysis tools rely on **simplified or aggregated models** of human mobility. These approaches typically represent population movement as static flows across a network and do not capture **heterogeneous behaviour, adaptive decision-making, or interactions between individuals**. As a result, these tools provide only limited insight into **emergent dynamics such as congestion formation, crowd behaviour, or behavioural responses to policy changes**. Another challenge faced by adopters is the **high cost and risk associated with testing urban interventions directly in the real world**. Infrastructure modifications, mobility regulations, or operational changes can affect thousands or millions of users and often involve significant financial investments. Without reliable predictive tools, decision-makers must implement changes with considerable uncertainty regarding their potential impacts. Adopters also face difficulties in **exploring multiple alternative scenarios efficiently**. Urban systems are influenced by numerous variables, including infrastructure configurations, mobility demand, environmental conditions, and policy incentives. Traditional modelling approaches often limit the number of scenarios that can be tested, making it difficult to systematically evaluate different options and identify optimal solutions.
Alternative solution	At present, potential adopters—such as **municipal authorities, mobility operators, infrastructure managers, consulting firms, and research institutions**—address the challenges of analysing and predicting urban dynamics through a combination of **traditional modelling tools, data-driven analytics, and limited simulation approaches**. While these methods provide valuable insights, they often lack the capability to fully capture the **complex behavioural interactions and emergent dynamics** that characterize modern urban systems.
Unique Value Proposition	The KER provides a **behavioural simulation engine for Urban Digital Twins** capable of modelling the movement and interactions of virtual people in complex urban environments. Compared to the solutions currently used by potential adopters—such as traditional transport models, isolated simulation tools, or static urban analytics platforms—the KER offers several competitive advantages that significantly improve the ability to analyse, predict, and optimize urban systems.
Competitors	Potential adopters currently rely on a range of **urban simulation, transport modelling, and analytics tools** to address the challenges of predicting and optimizing urban dynamics. These solutions constitute the main competitors or alternative approaches to the proposed KER. They can be broadly grouped into **traffic simulation software, general-purpose simulation platforms, urban system modelling tools, and simplified research simulators**.
Partners	MATHEMA
Timing	TRL5
IP Strategy	[Is IP strategy relevant for your KER? What is the Background (type/ partner)? What is the Foreground (type/ partner)? Describe any particular activity you have planned to manage IP].

KER name: DataFusion	
	Input from the Beneficiary

Description	DataFusion merges data from multiple sources into one cohesive view. It combines, cleans, and integrates datasets, allowing us to make sense of diverse data streams for analysis and decision-making.
Target market/end users	The **DataFusion module** operates in the broader market of **data integration and data management solutions**, particularly within domains where organizations need to **combine heterogeneous data streams from mobile devices, sensors, and digital platforms into a unified dataset for analysis and decision-making**. This market is growing rapidly due to the increasing deployment of **IoT systems, smart city infrastructures, mobile sensing technologies, and data-driven analytics platforms**. DataFusion competes within the segment of **data fusion and integration middleware**, providing capabilities to collect, clean, and integrate data coming from different sources into a cohesive and usable dataset. The module's implementation using **Android APIs and web technologies (jQuery/JavaScript)** positions it particularly well for applications involving **mobile platforms, distributed sensing systems, and web-based analytics environments**.
End-users needs / problems	Potential end users—such as **data engineers, software developers, urban data analysts, IoT platform operators, and digital twin managers**—often need to work with data coming from **multiple heterogeneous sources**. In environments such as smart cities, mobile sensing platforms, and IoT systems, data is continuously generated by **different devices, applications, and services**, each producing information in different formats and structures. While this data is valuable for analysis and decision-making, end users face significant challenges in **combining and preparing these diverse datasets for effective use**.
Competitive advantages	The **DataFusion KER** provides a module designed to merge, clean, and integrate heterogeneous datasets originating from multiple sources, including mobile devices, web services, and distributed sensing systems. Compared to current approaches used by end users—such as manual data integration workflows, custom scripts, or generic data processing tools—the DataFusion module offers several competitive advantages that improve the efficiency, reliability, and usability of data integration processes. - Native support for mobile and distributed data sources - Simplified integration of heterogeneous datasets - Lightweight and flexible architecture - Improved data consistency and quality
Use model	The **DataFusion KER** consists of a software module designed to **merge, clean, and integrate heterogeneous datasets originating from multiple sources**, including mobile devices, sensors, and web-based services. The module is implemented using **native Android APIs together with web technologies such as JavaScript and jQuery**, allowing it to operate within mobile and web-based data collection environments. The exploitation and deployment of this KER will follow a **software-oriented use model**, combining technology integration, service provision, and research collaboration.
Early Adopters	The early adopters of the **DataFusion module** are organizations and teams that **frequently deal with heterogeneous, fragmented data streams and experience strong operational friction when integrating data from multiple sources**. These actors are typically already managing complex data environments (e.g., IoT systems, mobile sensing platforms, or urban data infrastructures) and therefore **feel the need for automated data integration and cleaning tools more strongly than other potential users**. - Smart city technology teams and urban data platform operators - IoT solution providers and sensor network operators - Mobile sensing application developers

	- Research institutions working with heterogeneous datasets
Adopters' problems/needs	
Alternative solution	Potential adopters of the **DataFusion KER**—such as smart city platform operators, IoT solution providers, mobile application developers, and research teams—already attempt to manage heterogeneous data sources through a variety of **data integration and preprocessing approaches**. While these methods allow organizations to partially address the challenge of combining multiple datasets, they often require **significant technical effort, manual intervention, or complex infrastructures**.
Unique Value Proposition	The **DataFusion KER** provides a module for merging, cleaning, and integrating heterogeneous datasets originating from multiple sources such as mobile devices, sensors, and web-based services. Compared with the approaches currently used by potential adopters—such as custom data pipelines, enterprise ETL platforms, or domain-specific integration tools—the DataFusion module offers several competitive advantages that address key limitations of existing solutions.
Competitors	
Partners	Mathema
Timing	TRL3
IP Strategy	IP strategy not defined at this stage

Annex 4 - Institutional IPR Frameworks of EXTRACT Partners

This appendix provides additional information on the institutional Intellectual Property Rights (IPR) frameworks and policies of selected EXTRACT consortium partners. These frameworks support the management, protection, and exploitation of research results generated during the project.

The information presented in this appendix complements the summary of partner IPR policies described in Section 6 of this deliverable.

LOGOS Research and Innovation – IPR Policy

Recognizing and protecting the value of research results beyond their scientific and social contributions is an essential component of the research and innovation process. Safeguarding intellectual property represents the first step in establishing a value chain that facilitates technology transfer and the transformation of research outputs into accessible innovation.

LOGOS Research and Innovation acknowledges the importance of intellectual property protection and has established internal structures to support the safeguarding and valorisation of research outcomes. By integrating IPR protection into a broader strategy of knowledge valorisation, technology transfer, and economic return, LOGOS aims to enhance the impact of research throughout its lifecycle.

Types of Research and Ownership

LOGOS distinguishes between three types of research activities:

Independent Research
If a researcher conducts research using resources allocated by LOGOS, the researcher retains ownership of patentable inventions. Researchers may file for a patent independently while informing LOGOS or may transfer economic rights to the institution while retaining moral rights to authorship.

Collaborative Research
When research is conducted in collaboration with third parties, ownership of results is shared among the participating partners according to their respective contributions.

Commissioned Research
If research is conducted under contract for an external entity, LOGOS retains ownership of the intellectual property. The commissioning entity may subsequently acquire the rights or obtain an exclusive license for specific applications.

LOGOS ensures that both permanent and temporary research staff are treated equally regarding rights and obligations related to intellectual property.

Distribution of Economic Benefits

LOGOS applies a revenue-sharing model to distribute economic benefits derived from intellectual property:

- 50% allocated to the inventor(s)
- 20% allocated to the research institute or department
- 10% allocated to the affiliated research unit
- 20% allocated to a central fund supporting future IPR protection

This distribution model incentivizes researchers while ensuring institutional reinvestment in intellectual property protection and innovation activities.

Research Valorization Unit

The Research Valorization Unit at LOGOS plays a central role in coordinating intellectual property protection and commercialization activities. This unit provides multidisciplinary expertise covering scientific, administrative, legal, and economic aspects of intellectual property management.

Key responsibilities include:

- evaluation and processing of invention proposals
- patentability assessment and prior-art analysis
- management of patent filing procedures
- support for licensing, joint ventures, and technology transfer
- advisory support for confidentiality agreements and collaborative arrangements

Patent Protection Process

The internal procedure for protecting intellectual property includes several steps:

1. **Preliminary Notification**
   The inventor submits an initial disclosure form to the IPR Unit.
2. **Internal Assessment**
   Experts assess novelty, inventive step, and industrial applicability.
3. **Patent Filing**
   If the invention is considered viable, the IPR Unit coordinates the filing process with legal experts.
4. **Market and Value Assessment**
   Commercial potential is evaluated and opportunities for commercialization are explored.
5. **Ongoing Portfolio Management**
   The intellectual property portfolio is periodically reviewed to assess market relevance and strategic value.

All individuals involved in the evaluation and management process are bound by confidentiality obligations to ensure protection of intellectual property.

Universitat Rovira i Virgili (URV) – IPR Regulations

The intellectual and industrial property generated by researchers and students at Universitat Rovira i Virgili (URV) is governed by internal institutional regulations defining the legal framework for research, development, innovation, and knowledge transfer activities.

The official regulations are available at:

https://www.urv.cat/media/upload/arxius/normatives/propia/activitat_universitaria/investigacio/norm_prop_intelectual.pdf

Ownership of Industrial Property Rights

URV researchers:
URV holds ownership and exploitation rights for inventions developed by its researchers as part of their professional duties in teaching and research activities, in accordance with applicable industrial property law.

Unless proven otherwise, inventions for which protection is requested within one year after termination of employment are presumed to have been created during the employment period.

URV students:
Students retain ownership and exploitation rights over inventions developed during their academic activities. For jointly developed inventions, ownership is shared proportionally according to each contributor's involvement.

Ownership of Intellectual Property Rights

Researchers retain authorship rights over works created in the course of their research or teaching duties. However, the university generally holds exploitation rights in accordance with intellectual property legislation.

Students retain authorship and exploitation rights over academic works produced during their studies unless specific collaboration agreements define otherwise.

Software, Databases, and Collective Works

Software and applications developed within the scope of research activities are typically owned by URV. For collective works coordinated and published by the university, both authorship and exploitation rights belong to URV unless alternative agreements are established.

The university also retains exploitation rights for databases and structured datasets developed by researchers within their professional duties.

Confidentiality and Protection of Results

Members of the URV community involved in research activities must maintain confidentiality regarding information related to potentially protectable research results. Dissemination of such information must be coordinated through the relevant institutional structures to ensure proper protection of intellectual property.

Revenue Distribution

Income generated through commercial exploitation of intellectual property is distributed between inventors and the university according to predefined institutional rules.

Barcelona Supercomputing Center (BSC)

Internal procedures governing intellectual property management at the Barcelona Supercomputing Center cannot be publicly disclosed.

As a public research institution, BSC manages intellectual property according to Spanish national legislation governing research, innovation, and intellectual property protection.

Relevant legislation includes:

- **Law 14/2011 on Science, Technology and Innovation**
- **Law 21/2014 on Intellectual Property**
- **Law 1/2019 on Trade Secrets**

These legal frameworks regulate the authorship, ownership, and protection of intellectual property generated through research activities.

SixSq

SixSq manages intellectual property related to its software platforms and services according to its corporate legal framework and publicly available Terms and Conditions.

Intellectual property associated with the company's platforms and services is owned by SixSq or its licensors. Customers retain ownership of their data processed through the platform.

Users acknowledge that intellectual property associated with the platform software, applications, and related content is protected under copyright, trademark, and other applicable laws. Unauthorized modification, reproduction, or redistribution of such materials is prohibited unless explicitly permitted.

IBM

IBM did not provide a detailed description of its internal Intellectual Property Rights (IPR) policy for inclusion in this deliverable.

The company confirmed that its intellectual property commitments within the EXTRACT project are governed by the project Grant Agreement (GA) and Consortium Agreement (CA), which define the rules regarding ownership of results, access rights, and exploitation obligations for all consortium partners.

As a large industrial research organisation, IBM manages intellectual property according to its internal corporate policies and applicable international legal frameworks. Any project-related intellectual property generated by IBM within EXTRACT is handled in accordance with the provisions defined in the GA and CA.

MATH

Mathema did not provide specific institutional IPR policy descriptions for inclusion in this deliverable. Intellectual property generated by the company is managed internally according to corporate policies and applicable legal frameworks.

BINARE

BINARE owns any background and foreground IPR developed individually before and/or during the EXTRACT project. Intellectual property related to Binare Platform, BINARE's software and associated applications is owned by Binare, while the customers or users retain ownership of their data processed through BINARE's software and applications.

BINARE ensures that its intellectual property rights (IPR) are protected through contractual agreements and licensing mechanisms governing the use of its technologies. When using third-party IPR (including IPR resulting from EXTRACT project), BINARE follows the standard contractual agreements mechanisms to secure clear usage, licensing, ownership, benefits, ownership, and other rights relevant to the use of the specific IPR covered by the agreement.

Overall, BINARE manages intellectual property internally according to its corporate policies and in compliance with applicable national and international intellectual property regulations. Nevertheless, BINARE follows and observes any IPR clauses specified in the Grant Agreement (GA) and Consortium Agreement (CA) of the EXTRACT project, especially to the results that qualify as "Joint IPR" between BINARE and at least another organization/partner of the EXTRACT project.

As a Finnish company, BINARE's employee invention rights are governed by the Act on the Right in Employee Inventions, which stipulates that original rights belong to the inventor (employee), but employers can claim rights to inventions made during employment duties. Employees are entitled to "reasonable remuneration" if the employer acquires the rights, a right that cannot be waived by contract. Key Aspects of Finnish Inventor Law:

- Definition of Employee Invention: The invention must be patentable and result from the employee's work duties or significantly utilize the employer's experience/resources.
- Notification Obligation: Employees must notify the employer in writing without delay about any potentially patentable invention.

- Employer Acquisition: Employers have a right to acquire the invention, in whole or in part, by notifying the employee.
- Reasonable Remuneration: When an employer acquires rights, the inventor is entitled to reasonable compensation, considering the invention's value, the scope of rights transferred, and the employment terms.
- Disputes: Inventor rights to claim remuneration generally expire 10 years after the employer notifies the employee they are acquiring the rights.
- Limitations: The Act does not apply to employees at universities or higher education institutions, which have their own specific rules.
- International Transfer: If an inventor changes employers, any agreement regarding a previous invention is invalid if made more than one year after employment termination.

Last but not least, BINARE is bound to and follows the Finnish IPR Laws and guidelines as follows:

- Patents: Governed by the Patents Act, protecting technical solutions, new inventions, and industrial applications for up to 20 years.
- Trademarks: Protected under the Trademarks Act, safeguarding brand names and logos against unauthorized use.
- Copyrights: Protected under the Copyright Act (1961), applying to literary and artistic works, including performers' rights (commonly 50 years).
- Trade Secrets: Protected under special legislation covering non-public information of business value.
- Utility Models: Often called "short-term patents," these protect smaller technical inventions, notes the PRH.