



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

D5.2 Updated Dissemination, Exploitation and Community building report. First Sustainability report.

Version 1.1

Documentation Information

Contract Number	101093110
Project Website	https://extract-project.eu/
Contractual Deadline	31.03.2024
Dissemination Level	Public (PU)
Nature	Report (R)
Author	Andrei Costin (BIN) Mariia Gavriushenko (BIN) Janine Marie Gehrig Lux (BSC) Agathe Veillon (SIXSQ)
Contributors	ALL
Reviewer	MATHEMA
Keywords	Report, Dissemination, Exploitation, Sustainability, Community



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	8.03.2024	Initial Draft for Designated Internal Reviewer
V0.2	15.03.2024	Comments from Internal Reviewer
V0.3	22.03.2024	Final Formatting
V1.0	25.03.2024	Ready to Submit
V1.1	22.10.2024	Reviewers' comments (mid-term PR1 review M18) addressed

Table of Contents

1.	Executive Summary	4
2.	Introduction	6
	Structure of the document.....	6
	Aim, Scope, Objectives	7
3.	Updated Communication Plan and Results.....	8
	Branding and Project Identity	8
	Project Website	9
	Social Media	11
	LinkedIn.....	11
	X (ex-Twitter)	14
	Press Strategy	16
	Promotional Material	18
4.	Updated Dissemination Plan and Results.....	20
	Event Participation	20
	Publications	22
5.	Updated Community Building and Training Activities and Results	24
	Community building.....	24
	Training	25
	Internal Training.....	25
	External training.....	26
6.	Updated Exploitation Plan and Report.....	27
	IPR/exploitation contact point identification and listing	27
	Preliminary Key Exploitable Results (KERs) Identification Steps.....	28
	Preliminary Exploration of Tools for Exploitation of KERs	37
	Preliminary Market Understanding and Innovation Gap Analysis	46
	Suggested New Product.....	46
	Potential Need for the Suggested New Product	46
	Competitors analysis	46
	What are the product's advantages in comparison to the competition?	46
	Potential customers	47
	Main target users	47
	Stakeholders	48
	Market segmentation	49
	Upcoming Exploitation Plan for M16-M36.....	49

7. Initial Sustainability Plan.....51

Results and their Sustainability51

Self-sustainable Community52

8. Evaluation of Performance and KPIs (until M15).....53

9. Conclusion55

10. Acronyms and Abbreviations56

.....56

Annex 1 - EXTRACT Dissemination Register57

Annex 2 - EXTRACT Detailed list of publications60

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.

An initial communication, dissemination, community building, training, and exploitation plan was laid out in M6, in the form of *D5.1 Dissemination, Exploitation and Community Building Report*. This deliverable, *D5.2 Updated Dissemination, Exploitation and Community building report, first sustainability report*, presents a detailed overview of the activities undertaken from M1-M15 and provides an update on intended work to be carried out from M16-M36, including an initial sustainability report.

From M1-M15, the EXTRACT project consortium actively raised awareness of the project and reached target audiences through the tools and channels defined in D5.1. The first year of the project was devoted to creating a recognizable brand around the project and the tools and channels that would help access potential stakeholders.

The project is currently on track, meeting the key project indicators (KPIs) defined for its communication activities, and in some cases, it has surpassed the initial indicators set out. Branding and project identity, a project website, social media channels, a press strategy and promotional materials served as the basis of this strategy. Updated and engaging communication channels have been an important part of this strategy. Highlights so far at M15 include 1500 website sessions and 224 LinkedIn followers. Overall, these and other indicators reflect positively on the initial communication strategy and have allowed the communication and dissemination team, led by the Barcelona Supercomputing Center (BSC) to identify activities to continue to pursue and areas of improvement. [Chapter 3](#) provided detailed information on the results of the communication plan and plans for M16-M36.

As the project has developed, the consortium's dissemination activities have increased, as has interest in the project and its technologies and solutions. [Chapter 4](#) provides information on the 16 participations in events, one keynote speech and four publications undertaken by M15. Although currently a bit behind with respect to the KPIs defined, the WP5 team monitoring these activities is convinced that the project is on a good track given that dissemination tends to increase as project technology matures.

Community building and training are two activities that will help to ensure the uptake and application of EXTRACT technologies. The strategy for these activities at M1-M15 has been to identify and reach out to potential stakeholder communities by participating in the larger European artificial intelligence, data, research and industrial communities. As project results mature, more information will be shared with these communities through events and training.

[Chapter 5](#) provides an updated plan for community building and training that reflects on the potential integration of the communities identified in M1-M15. This includes a change in the initial approach set out in D5.1 regarding training. An updated internal training plan addresses the needs of partners, specifically more internal training from M1-M18 (three trainings instead of one in this time period), and an external training plan that builds on the results of the internal training plan (pushing external training past M18 and including recorded material to share with the external community) is a key part of the plan for M16-M36.

[Chapter 6](#) describes an updated and meticulously crafted exploitation report and plan designed to strategically advance the objectives of the EXTRACT project. This includes the identification and cataloging of IPR/exploitation contact points, representing critical interfaces essential for navigating the intricate landscape of intellectual property rights and exploitation endeavors. Furthermore, the plan delineates a systematic approach for the identification of Key Exploitable Results (KERs) at stage M14 and at the end of the project. Also, Chapter 6 includes a rigorous exploration of the market landscape through the lens of Market Understanding and Innovation Gap Analysis, a research-driven process tailored to discerning innovation gaps and potential opportunities inherent in the project's purview. Collectively, these components converge to form a resilient and refined exploitation plan, positioning the EXTRACT project on a well-informed and research-backed trajectory for success in the dynamic domain of extreme data management and analytics.

The Initial Sustainability plan is presented in [Chapter 7](#). The chapter outlines the European Research project's definition of "sustainability" and its criteria, emphasizing the continued delivery of benefits post-financial assistance. Within the context of the EXTRACT project, sustainability is approached through two main strategies. Firstly, the project aims to leverage and promote its complex algorithms by integrating them into a dedicated marketplace for extreme data mining workflows. These algorithms, showcased through two distinct use cases, highlight their ability to analyze vast amounts of data under extreme conditions. The second approach involves building a robust network of collaborations and synergies with public and private bodies, aligning with Task 5.2 Community building activities. The chapter details the initial sustainability plan for EXTRACT, including the promotion of project outcomes, the development of a marketplace, and the establishment of a strong stakeholder ecosystem. Further actions involve brainstorming sessions, publication on research platforms, and collaboration with partners to ensure sustainable exploitation. The chapter emphasizes the dual focus on leveraging EXTRACT results and building a supportive community to ensure the project's long-term impact.

[Chapter 8](#) presents the evaluation of performance and KPIs at M15.

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

This deliverable *D5.2* presents a first dissemination, exploitation, sustainability, community building report and updated strategies of the EXTRACT project. This deliverable provides a description of the activities performed until M15 and details dissemination, community building, and exploitation plans for M15 - M36. The report monitors the initial strategy laid out in *D5.1 Dissemination, Exploitation and Community Building Report* to ensure that the project results are reaching 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 eight chapters. [Chapter 2](#) contextualizes the initial dissemination, exploitation and community building 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](#), Updated Communication Plan and Results, which describes the project's four main communication channels and their effectiveness from M1-M15 and plans for M16-M36. [Chapter 4](#), Updated Dissemination Plan and Results, details the dissemination activities pursued by the consortium and the next steps as project results mature. [Chapter 5](#), Updated Community Building and Training Plan and Report, analyzes the activities undertaken to identify and build a community around the project and suggests concrete activities for ensuring that project results reach this community through training, event participations, and collaboration in the European Artificial Intelligence (AI), data, research and industrial communities. [Chapter 6](#), Updated Exploitation Plan and Report, presents an exploitation report and updated plan for the EXTRACT project, strategically advancing its objectives. This includes identifying critical interfaces for intellectual property rights, systematically approaching Key Exploitable Results, and conducting a thorough market analysis, culminating in an exploitation plan for M15-M36. [Chapter 7](#), Initial Sustainability Plan, defines sustainability in European Research projects and outlines EXTRACT's dual approach to ensure longevity. It focuses on promoting complex algorithms through a dedicated marketplace and building a strong stakeholder ecosystem through community building activities, emphasizing collaborative efforts for sustainable

impact. [Chapter 8](#), Evaluation of Performance and KPIs (until M15), presents activities, KPIs, measures and status. The deliverable D5.2 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). The previous deliverable D5.1 set a strategy for communicating, disseminating and exploiting the outcomes of the EXTRACT project throughout the entire project. This deliverable reports on the status of these activities from M1-M15 and calibrates and proposes a plan for the activities for M16-M35.

D5.1 was submitted at M6, and served as an initial framework for communication, dissemination, training and exploitation. 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 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 D5.2 is to revisit D5.1 and ensure that the project aligns with the initially defined goals, and modify the strategy if needed. D5.2 objectives include a) providing an overview of the status of the initial strategy defined in D5.1; b) reflecting on the success of the strategy by measuring project results against key project indicators; c) revisiting project activities in light of the status of the project; d) defining next steps for M16-M36 to ensure the KPIs are reached.

3. Updated Communication Plan and Results

This section presents the status of the EXTRACT project communication activities from M1-M15. It reports on the status of the initial communication plan, which was prepared as part of the *D5.1 Dissemination, exploitation and community building report*, submitted at M6.

Throughout the first 15 months of the project, the consortium has worked to raise 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. The communication goal has focused on increasing the project's visibility by highlighting its goals, innovative aspects and scientific value and by introducing the partners and their unique contributions.

Emphasis has been placed on the potential impact of the project by introducing the application of the extreme data mining technology in two use cases. These use cases demonstrate the potential for the project to address key societal challenges by harnessing extreme data via a distributed data-mining platform that exploits the leading technologies of our time to provide individuals with high-quality actionable knowledge that will improve crisis management and advance science.

As part of the communication strategy, five main activities were defined to ensure communication with diverse audiences: 1) [branding and project identity](#) 2) [project website](#), 3) [social media](#), 4) [press strategy](#) and 5) [promotional material](#). Great efforts have been made by the consortium to ensure that communication materials and tools are updated regularly with information ranging from more technical, to more focused on the general public.

Branding and Project Identity

Branding and the creation of a project identity have been an important first step for creating visibility and recognition for the project. To make a memorable and representative brand, the WP5 dissemination team subcontracted a design company to define and create a common graphic identity.

The brand guide, logos, and templates have been made available to all consortium members via the internal repository. All materials created include 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 have consistently applied the project brand in their communication and dissemination activities.

KPI

The key project indicator set for this activity was one PowerPoint presentation, 1 word template, and 1 poster template. This KPI has been met.

Project Website

The project website is <https://extract-project.eu>. It is maintained by partner Barcelona Supercomputing Center (BSC) which ensures that the information on the website is up-to-date and representative of the consortium's work. The website is the main public communication channel of the project and it provides information about the project (partners, equality and diversity, branding), its objectives, use cases, project results (publications and eventually deliverables and demos), news and events.

An internal editorial calendar has been created to encourage partners to provide technical contributions to their activities so that they can be posted on the [news section](#) of the website. One partner contributes per month. The texts are intended to address the technical aspects of the project in an accessible way and serve as an opportunity for each partner to explain their work, progress and any preliminary findings. In addition to providing insights into the work of each partner, monthly news pieces provide fresh content that can be shared across the other project communication channels, thus increasing the potential for engagement. All in all, 20 news pieces have been published in the first fifteen months of the project.

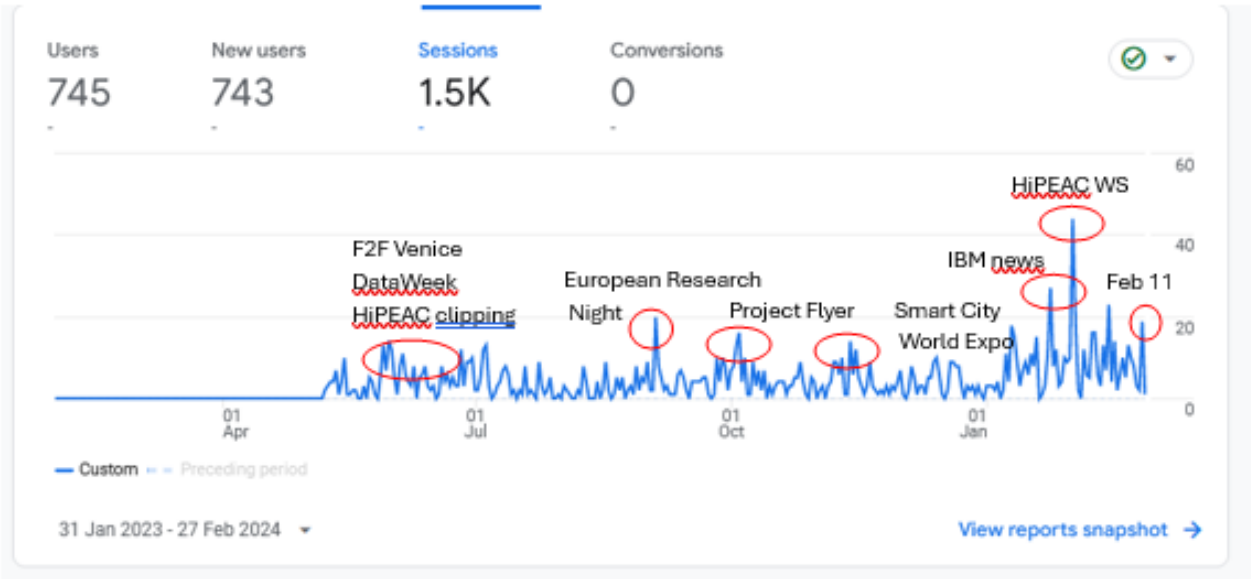
The [event](#) and [publication](#) pages are also important sources of information for project target audiences. The event section lists past and upcoming events to encourage visitors to learn about the types of events the consortium participates in and where there could be possible opportunities for collaboration. The publication page provides information on the open-access publications produced by the project in addition to the slides presented at events. Currently, 13 publications or presentations are available on the website. The dissemination team shares teasers from these texts via social media to draw attention to the website and its full content.

Monthly news pieces, event updates, publications and presentations ensure that the website remains consistently updated and relevant to a variety of audiences. To measure the value of these updates, Google Analytics is used to monitor visits to the website. D5.1 set a key project indicator (KPI) of 1000 per year (total of 3000 sessions by the end of the project). As of M12, there had been 960 sessions, and by M15 there were 1,500 sessions (Figure 1). While the first-year KPI was only partially met, there has been clear growth between M13-M15, suggesting that this KPI is on track to be met by the end of the project.

The peaks in website sessions in Figure 1 were found to loosely coincide with the dates of events or news pieces on the website. An especially large peak can be seen toward the end of M13, when EXTRACT participated in the [CompContinuum Workshop at HiPEAC 2024](#). This jump in website sessions suggests that keeping the website updated with events and news pieces is very important for attracting interest in the website. The consortium will continue to post monthly or bi-monthly updates based on this result.

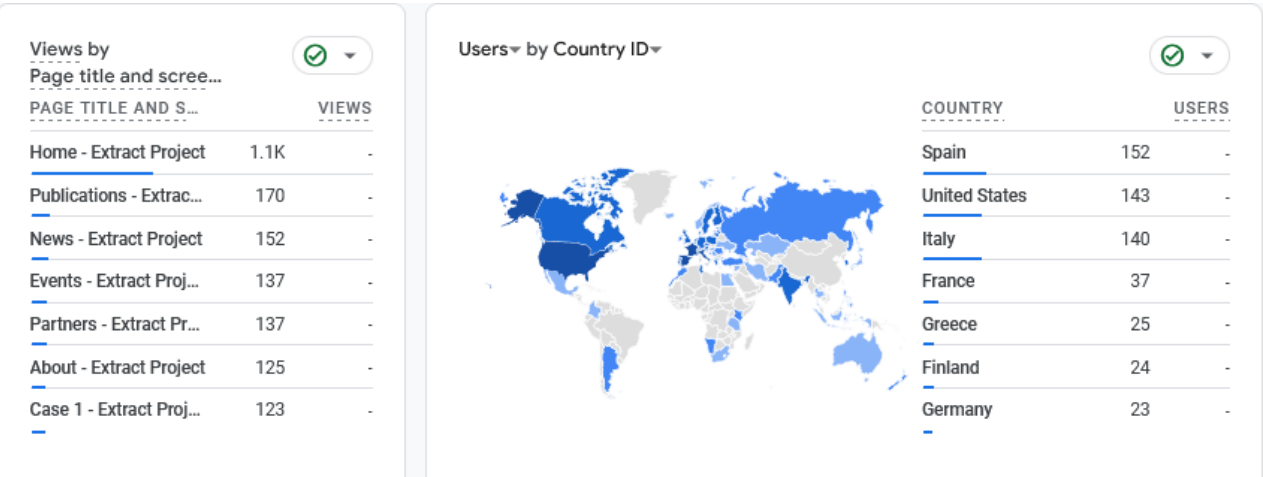
Google Analytics is also used to analyze what web pages are most visited. From M1-M15, website users mostly visited the EXTRACT project home page, followed by the publications page and newspage (Figure 2). This is expected given that the project has been directing users to these main pages via social media. In the next project period, a greater focus will be put on directing users to the publications, news and use case web pages to improve engagement and a deeper understanding of the project results.

Figure 1: EXTRACT website sessions recorded by Google Analytics M1-M15, pin-pointed to some main events and achievements.



Website users by country number also paint an interesting picture. While it is expected that many visitors would come from Spain, Italy and France because the project has three Spanish partners and the use cases take place in France and Italy, it is surprising that the users from the US hold the second-ranking. This bodes well for the project’s applicability and shows that the innovative technology being produced in the project is valuable and competitive outside of Europe.

Figure 2: Ranking of page views and users by country recorded via Google Analytics at M15.



Starting in M16, the project website will also include three project videos, demos produced by project partners and training materials. The use case pages will also be updated with new multimedia material prepared by the use case task leaders to help bolster engagement and show the applicability of the technology in the use cases.

KPI

The key project indicator set for this activity is 1000 sessions per year, 3000 sessions by the end of the project. This KPI is on track to be met with 960 sessions in year 1, and 1500 sessions by M15.

Social Media

The project has two social media accounts, 1 [LinkedIn](#) and 1 [X](#) (ex-Twitter) account. Social media is used to raise awareness of the project's activities and act to attract visits to the project website. Social media also plays a role in expanding communication efforts to broader audiences and engaging with other players in the political and industrial sphere.



EXTRACT EU Project



@EXTRACT_EU_proj

Each social media channel is used to access different target audiences and convey different messages. LinkedIn is used to connect with industry and researchers. The X (ex-Twitter) is used to engage with a broader audience, and focuses on announcing partner participation in events and announcing updates to the project webpage. The shorter word count of this channel yields it to brief and less technical announcements.

For both channels, we follow hashtags [#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 and ADRAe), AIONdemand, and the Big Data Value Association (BDVA), among others. Moreover, multiple posts a week, including posts like [#Flashback Friday](#), keep the LinkedIn channel very active and offer an engaging way to highlight past presentations and event participations.

The EXTRACT project also participates in social media campaigns that are focused on 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.

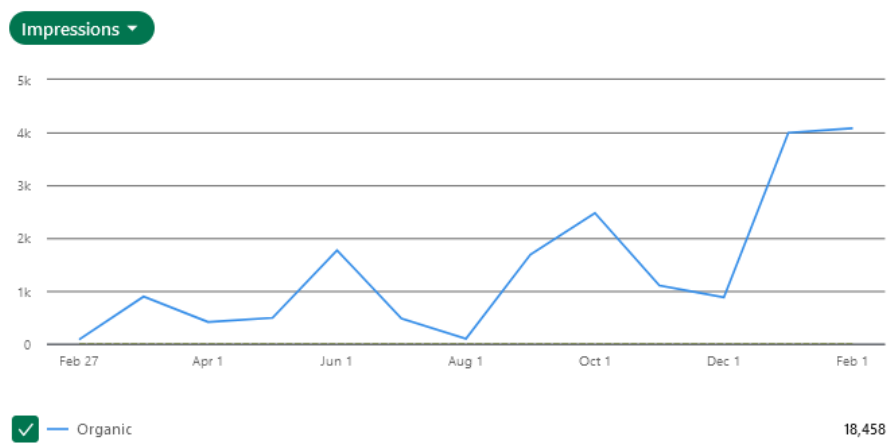
LinkedIn

With 224 LinkedIn followers at M15, we can say that this communication channel has been very successful for the project. Through this channel, we share information on updates on the website, engage with the EU AI and research ecosystem and industry and we highlight the accomplishments and participations of consortium members.

The project posts news and information about the project's participation in different events. The goal is to engage with industry-related stakeholders and academics who have similar interests. LinkedIn activity is monitored via LinkedIn Analytics, which shows general information on followers, visitors, visitor demographics and more.

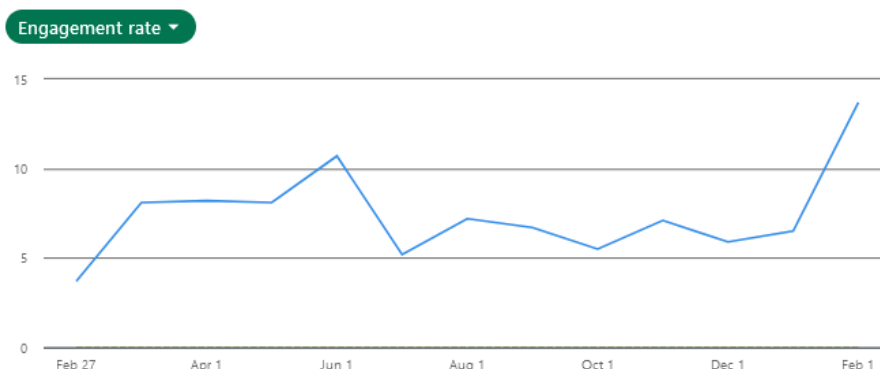
Figure 3 demonstrates the evolution in the amount of impressions (number of views) in the first fifteen months of the project. Overall, there have been 18,458 impressions on the channel. This suggests that LinkedIn is an important channel and that the material we are posting there is of interest to our audience.

Figure 3: LinkedIn impressions recorded by LinkedIn analytics from M1-M15



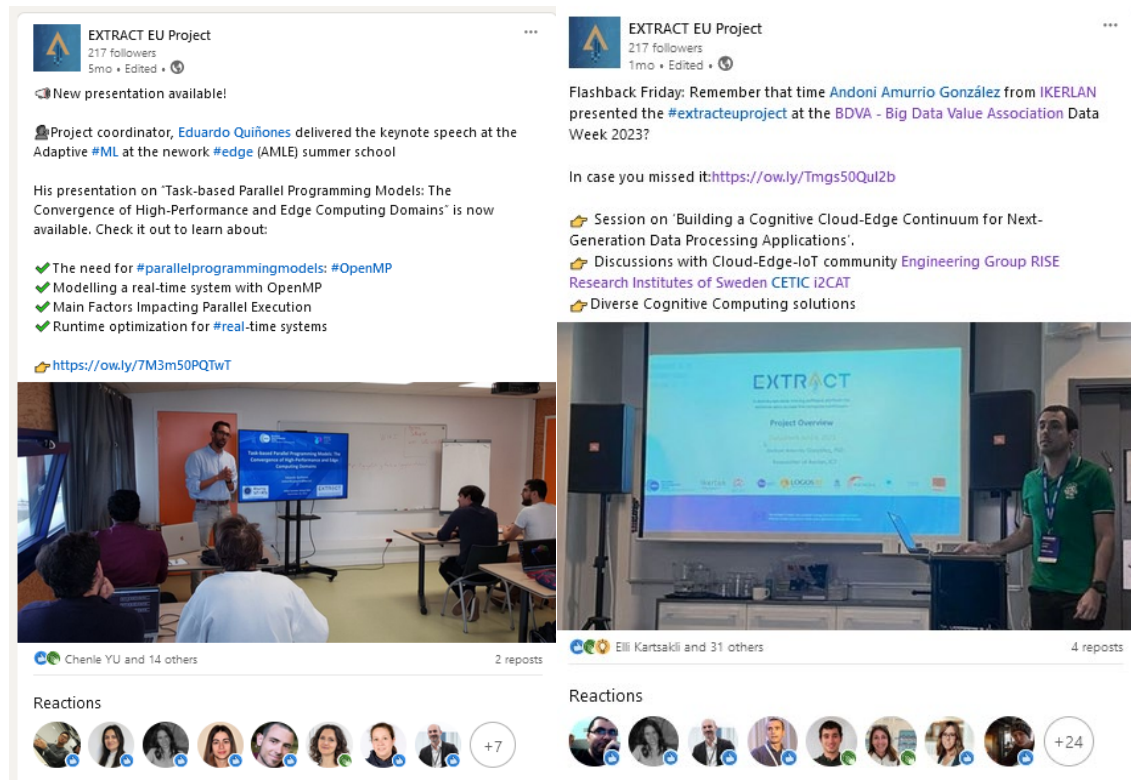
In addition to the continuously increasing number of impressions, the engagement rate (interaction with content, e.g. shares, comments, likes, clicks) also reflects positively on the success of the LinkedIn strategy. The average engagement rate is 5%, and this value has remained fairly constant since the beginning of the project. Given that a 2% engagement rate is considered to be good in the context of social media [marketing](#), we believe our strategy is working well and will continue to heavily use LinkedIn for sharing information about the project.

Figure 4: LinkedIn engagement rate as recorded by LinkedIn analytics from M1-M15.



Some examples of especially successful content on LinkedIn include posts announcing links to presentation slides and Flashback Friday posts summarizing past event participations (Figure 5).

Figure 5: Examples of posts with high impression amounts (1021 and 824 impressions respectively).



LinkedIn analytics also provides the WP5 dissemination team with professional and demographic information on our audiences. This information allows us to monitor and tailor our message. At M15, our followers are primarily from research, higher education and IT services and consulting (Figure 6). This result suggests our message is getting to the audiences we intended. In M16-M36, we will work to expand our followers to more industry-related players. We will do this by sharing information on project results as they become available.

Figure 6: LinkedIn follower percentage by type of industry.

Follower demographics

Industry

Research Services - 45 (18.5%)

IT Services and IT Consulting - 30 (12.3%)

Higher Education - 26 (10.7%)

Software Development - 10 (4.1%)

Data Security Software Products - 10 (4.1%)

Government Administration - 8 (3.3%)

Computer and Network Security - 7 (2.9%)

Technology, Information and Internet - 5 (2.1%)

Business Consulting and Services - 5 (2.1%)

IT System Custom Software Development - 4 (1.6%)

X (ex-Twitter)

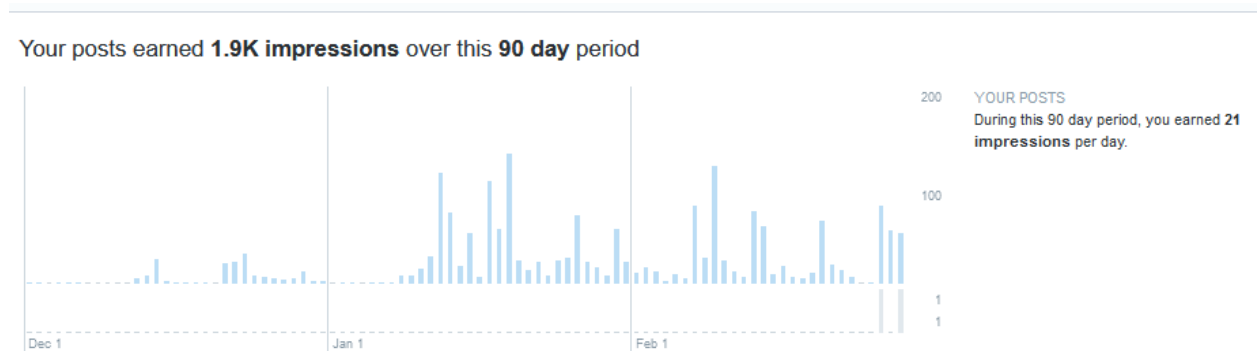
X (ex-Twitter) is the other social media channel the project uses. With 368 million users, this channel grants the project access to many users around the world. The project uses this channel to reach the general public, policymakers, industry and more. The Twitter account has been active since the beginning of the project and has had moderate success. At M15, the project has 60 followers.

This rather low number may be due to a variety of issues not contemplated in D5.1. For one, since the project's inception, this channel has changed ownership, business and operating models, leading to chaotic changes and polarizing agendas, and most recently has been rebranded from Twitter to X. This may have pushed some organisations and individuals to use the channel less, completely quit the rebranded X platform. Additionally, the project does not currently seem to cater to the type of content that tends to do well on Twitter, specifically real-time very short alerts and commentary on trending issues.

Currently, the WP5 dissemination team tweets several times a week about project updates, participations and additions to the project website. While the engagement by followers is positive, it does not seem to be the kind of content that encourages follower growth. From M16 on, WP5 will try to engage more with other X users and to piggy-back on partner institutional content to increase visibility. Information about the use cases and their impact on society will also be emphasized. Despite these activities, it is worth noting that this may not be the best channel for the purposes of this project.

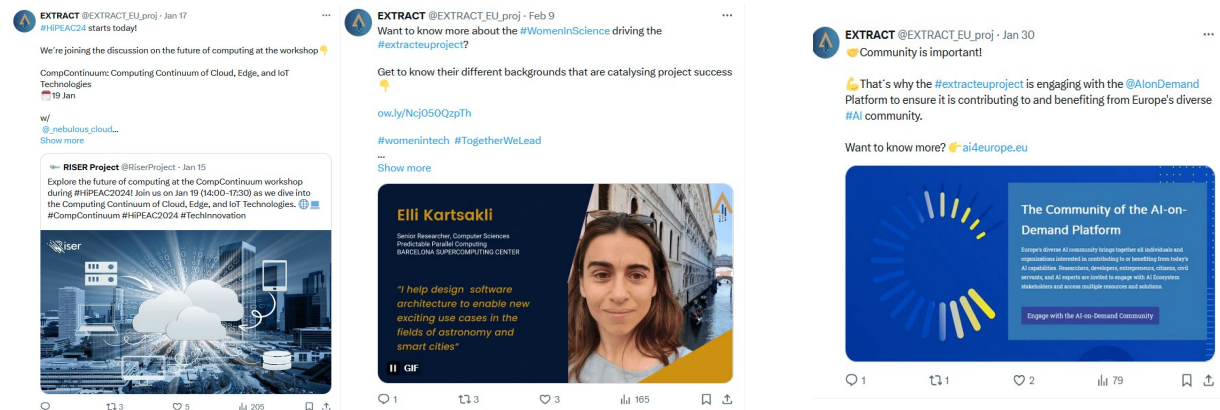
X analytics had been used to measure the success of communication strategy, however fewer analytical tools are available and analytical comparisons are limited to three months of historical, which affects our analysis of success over time.

Figure 7: X Analytics Providing Snapshot of M12-M15 impressions



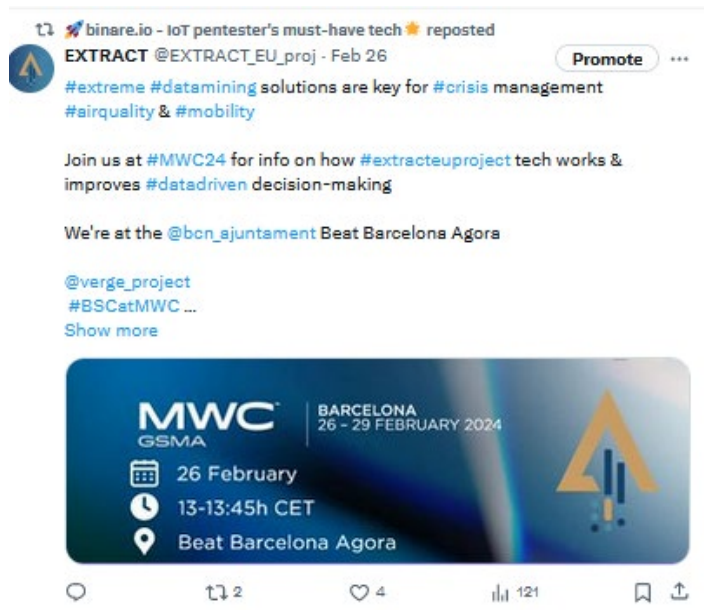
From M12-M15, the X channel of the EXTRACT project received 1,900 impressions. The peaks in impression numbers mainly corresponded to targeted social media campaigns. For example, peaks in January correspond to tweets related to EXTRACT participation in the HiPEAC 2024 CompContinuum Workshop. Several tweets were prepared to build up to the workshop, held on 19 January 2024. Another peak in February 2024 corresponds to the Women and Girls in Science Day campaign launched on 11 February 2024.

Figure 8: Examples of X posts with high follower engagement. Left image, repost of 2024 HiPEAC workshop with fellow EU-funded project, RISER. Top right, Women and Girls in Science day post, Bottom, Post about participation in AionDemand platform



Several project partners also use X to share information about the EXTRACT project. Taking advantage of institutional X networks helps the project take advantage of the reach of partners' networks.

Figure 9: Example of a repost of an EXTRACT tweet by partner BINARÉ.



The WP5 dissemination team will continue to monitor the analytics of both social media channels and adapt as needed to ensure relevant and updated information reaches our target audiences.

KPI

The key project indicator set for this activity is 250 LinkedIn followers and 150 X followers by the end of the project. This KPI is on track with 224 LinkedIn followers and 60 Twitter followers at M15.

Press Strategy

A press strategy was contemplated as part of the communication strategy as a means of sharing EXTRACT project highlights and news with the technical media. As part of this strategy, three press releases are anticipated. A press release announcing the launch of the project, a second one in year two will highlight milestones in the project, and one in the last year sharing project final results. Press releases are an effective way of communicating information about the EXTRACT project to the technical media and their corresponding audiences. Moreover they can be used as reference and foundational documents for project partners to build on.

The first project press release was launched on [3 March 2023](#). It received attention from prominent technical media and presented the project objectives, partners and the two use cases.

Figure 10: First EXTRACT press release, launched 3 March 2023 to the technical media

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

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



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

Another aspect of the press strategy includes establishing relationships with the technical media so that they grant the project space to publish their news pieces in their channels. Partners have actively sought out these technical media outlets and provided them with interesting pieces that are tailored to their specific audiences. Partners BSC, SIXSQ and URV have all appeared in the HiPEAC technical magazine. Other partners have used their institutional channels and their networks to publish information on the project.

Figure 11: Examples of press clippings feature EXTRACT partners BSC and SIXQ

FEB. 10, 2023

'Distributing computations across the continuum opens up huge possibilities' – Elli Kartsakli

Elli Kartsakli is a senior researcher at Barcelona Supercomputing Center (BSC). A recipient of a prestigious Ramón y Cajal grant from the Spanish Government, Elli has been in active in a number of European Union-funded projects, including ELASTIC and CLASS, and is now coordinating two projects funded by the Spanish government, PROXIMITY and AIR-URBAN. We caught up with Elli to find out more about creating a viable compute continuum and powering safer, cleaner mobility.



Industry focus



With the number of connected devices constantly growing, and the data that these devices generate climbing into the zettabytes, orchestration and distribution are becoming increasingly complex problems. In this article, Agathe Veillon (SixSq) explains how a number of European projects are using SixSq's Nuvla technology to manage IoT-edge-cloud orchestration.

IoT-edge-cloud
Building a simple, secure, and future-proof infrastructure

External references to the project, counted as press clippings in this project, add up to 15. This is a satisfactory number of press clippings, and WP5 will continue to build relationships with other technical media and encourages partners to do the same to expand the reach of the project in these channels.

The key project indicator set for this activity three press releases by the end of the project and 25 press clippings. One out of three press releases have been published and 15 press clippings mentioning the project. This KPI is on track to be met.

KPI

The key project indicator set for this activity three press releases by the end of the project and 25 press clippings. One out of three press releases have been published and 15 press clippings mention the project. This KPI is on track to be met.

Promotional Material

Communication and dissemination material has been prepared to help partners to quickly, uniformly and properly (correct acknowledgement and branding) produce and distribute materials that will help them raise awareness of the project and its goals in an attractive way.

- **Flyer:** A tri-fold flyer providing overview information about the project, its objectives and the use cases have been created and used in various events including ([2023 IEEE International Conference on Cyber Security and Resilience](#), [CSC Supercomputing Innovation Summit](#), [2023 Hardware.io Security Conference and Training event](#), [Smart City Expo World Congress](#), [SC23](#), and [the Italian Ministry of Culture Networking Event](#)). The print quality flyer is available for download by the consortium via the internal repository. A downloadable version is also available online at the Branding section of the website.

Figure 12: Side 1 of EXTRACT Flyer



- **Poster:** A general overview roll-up was developed to be used by all partners. The roll-up includes an overview of project objectives, partners and the use cases and includes a QR link to the project website. Partners are encouraged to use the roll-up in their own communication and dissemination activities. Partner Binare printed the roll-up (to be used in Q3 2024 for the "Binare EXTRACT Student Hackathon", initially planned for Q2 2024 but postponed for logistics reasons) and the BSC used the poster for the general public "[Research Night](#)" event. A technical poster was prepared by BSC providing an overview of the project technology. It was submitted to the 2024 ACM womenCOURAGE conference and will be made publicly available if it is accepted.

Figure 13: Project roll-up



- **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 helps transmit the project's objectives and key messages consistently.
- **Videos:** Promotional videos explaining the project were identified in D5.1 as an effective way to presenting the project attractively and concisely. Two videos are currently under production and are expected to be ready by mid-March. These two 1.5 minute video clips will be used as an introduction to the project technology and its application to the two use cases. A wrap-up video will be produced at M34-36 with a greater technical focus for industrial and research audiences.
 - In M14, partner Metropolitan City of Venice released a [video](#) demonstrating the difficulties that individuals could face in a crisis situation when they need to get to safety. Narrow streets, human

obstacles and poor mobility were portrayed via a simulation conducted with the help of retired policemen and women.

- Factsheet/infographic: A factsheet and or infographic will be prepared at the end of the project (M36) to provide a concise and visually appealing overview of the project's results.

KPI

The KPI for this activity is 2 posters, 3 videos, 1 factsheet/infographic. At M15 2/2 posters are available, 1/3 videos, 0/1 factsheet/infographic. This KPI is on track to be met.

4. Updated Dissemination Plan and Results

The project dissemination plan goes beyond communicating the goals and objectives of the project and works to share and advance project results to actors outside the project. The consortium has been using the above-mentioned communication tools to engage with different stakeholders who can use EXTRACT technology, for example, research and innovation organisations across Europe, industry, technology and service providers.

Although the first year of the project mostly focused on implementing and monitoring the communication strategy and its goal of raising awareness of the project and its goals, steps have also been taken to follow through with the dissemination strategy. The dissemination strategy largely centres on sharing project results through academic, industrial and European AI ecosystem events and publications. A series of meetings were also held between the project actors and interested stakeholders and the project and its fellow EU-funded cluster siblings to discuss the potential impact of project results and future collaborations. Results were also shared in the form of a keynote and workshop. More information on these activities will be discussed in greater detail in [Section 5: Updated Community Building and Training Activities and Results](#).

From M1-M15, the consortium participated in 16 events and published 4 publications. These activities serve to set the foundation for the project and to share the research direction and innovative facets of the project. M16-M24 will ramp up these activities. As results become more mature, they will increasingly be shared with communities that could most benefit from them.

In the last phase of the project, from M25-M36, the WP5 team will actively encourage the uptake of technologies to ensure project sustainability. Among other activities, tutorials, industry events and specific partnerships will be pursued within the EXTRACT community and beyond to ensure the broad dissemination and application of project results.

Event Participation

Events represent a key dissemination channel. High-level peer-reviewed conferences in particular allow the consortium to present the latest updates of the projects and effectively involve different audiences in technological discussions early on. A list of

potential dissemination events was prepared as part of D5.1 and the WP5 team is continuously looking for opportunities that will allow the consortium to access its target audiences.

From M1-M15, the consortium participated in 16 events with academic, research and industrial audiences. The events and their details can be found publicly on the [project website](#) and internally in the dissemination register ([Annex 1](#)).

Figure 14: Event participations on the EXTRACT website.



EXTRACT at the MWC24

The 2024 Mobile World Congress (MWC) was held in Barcelona from 26-29 February. Partner Barcelona Supercomputing Center (BSC) attended and

Location: Barcelona
Date: February 26, 2024



HiPEAC 2024: Workshop on CompContinuum

Partner Barcelona Supercomputing Center will take part in the HiPEAC 2024 Workshop: "CompContinuum: Computing Continuum of Cloud, Edge, and IoT"

Location: Munich
Date: January 19, 2024



9th Int'l WS on Serverless Computing

Partner URV participated in the 9th International Workshop on Serverless Computing (WoSC9). This workshop brought together researchers and practitioners to

Location: Bologna, Italy
Date: December 11, 2023



URV presents at MIDDLEWARE 2023

BSC partner Universitat Rovira i Virgili (URV) presented its paper, "Glider: Serverless Ephemeral Stateful Near-Data Computation" on 13 December 2023.

Location: Bologna, Italy
Date: December 11, 2023

Figure 15: Right, URV partner Daniel Pons presents at the 25th ACM/IFIP International Middleware Conference , Left, BSC partner presents at the 2023 Smart City Expo World Congress.

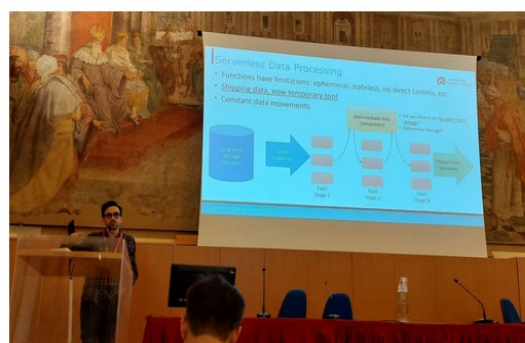
Sharing EXTRACT's vision with Smart City Attendees

Date: November 10, 2023



Partner URV presents "Glider" storage service at Middleware Conference

Date: December 20, 2023



Keynote talks are an important source of credibility for project partners, and they help focus the attention of specialized audiences on the work being done in the project. For this reason the project expects to participate in two keynote talks. One keynote was given by project coordinator, Eduardo Quiñones at the [Adaptive Machine Learning at the Network Edge Summer School 2023](#). His talk entitled, "Task-based Parallel Programming Models: The Convergence of High-Performance and Edge Computing Domains", provided an important look into parallel programming, modelling real-time systems, and runtime optimization.

Figure 16: Eduard Quiñones delivering the keynote speech at the AMLE Summer School

EXTRACT Coordinator Gives Keynote at AMLE Summer School

Date: September 28, 2023



Partners asked to inform the WP5 dissemination of upcoming events so that mini campaigns can be launched to highlight the consortium's involvement in these events. News pieces summarizing the event are published afterward with a link to the presentations. Participations are logged in a dissemination register (see [Annex 1](#)) to facilitate their reporting in the participant portal.

Initial brainstorming is underway to identify the location and logistics of participating in an industrial booth and setting the stage for the final event.

The key project indicator set for this activity 2 keynote talks, 1 industrial booth, 1 final event with more than 50 attendees. This KPI is on track to be met 1/2 key notes, 0/1 industrial booths, 0/1 final event.

KPI

The key project indicators set for this activity are 2 key note talks, 1 industrial booth, 1 final event with more than 50 attendees. This KPI is on track to be met 1/2 key notes, 0/1 industrial booths, 0/1 final event.

Publications

Peer-reviewed publications are an important way for the consortium to disseminate its findings and ensuring that EXTRACT technical results are shared openly and as widely

as possible. WP5 prepared a set of publication guidelines to inform the consortium of the OpenScience requirements and to define a transparent and agreed upon publication procedure. Partners have been informed that all EXTRACT publications should include the following sentence:

The research leading to these results has received funding from the European Union's Horizon Europe Programme under the EXTRACT project, grant agreement number 101093110.

From M1-M15 four publications were published. At M15, one additional publication has been accepted and two publications are in the pipeline. Table 1 provides an overview of the publications. Additional details can be found in [Annex 2](#). **The publications have been updated in the SyGMA portal as requested by the reviewers and updated accordingly on the project website.**

Table 1: EXTRACT publications in M1-M15

Publication type	Title	Authors	Title of journal or equivalent	Date
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	August 2023
Conference proceedings	Glider: Serverless Ephemeral Near-Data Computation	Barcelona-Pons, Daniel, García-López, Pedro and Metzler, Bernard	Middleware '23: 24th International Middleware Conference	Nov. 2023
Conference proceedings	Scaling a Variant Calling Genomics Pipeline with FaaS	Arjona, Aitor, Gabriel-Atienza Arnau, Lanuza-Orna Sara, Roca-Canal, Xavier	Middleware '23: 24th International Middleware Conference	Dec. 2023
Journal	MLLess: Achieving cost efficiency in serverless machine learning training	Gimeno Sarroca, Pablo, Sánchez-Artigas, Marc	Journal of Parallel and Distributed Computing	Jan. 2024
Just accepted Conference proceedings	Dataplug: Unlocking extreme data analytics with on-the-fly dynamic partitioning of unstructured data	Arjona Pérez, Aitor , Garcia Lopez, Pedro and Barcelona-Pons, Daniel	24th IEEE/ACM International Symposium on Cluster, Cloud and Internet Computing	May 2024

The key project indicator set for this activity is four publications per year. At M15, we're on track to meet this KPI with 3 publications in year 1, 4 total at M15 and 2 in the pipeline.

KPI

The key project indicator set for this activity is four publications per year. At M15, we are on track to meet this KPI with 3 publications in year 1, 4 total at M15 and 2 in the pipeline.

5. Updated Community Building and Training Activities and Results

Community building and training are two crucial activities that help to ensure the uptake and application of EXTRACT technologies beyond the life of the project. These activities depend heavily on identifying interested and potentially interested targets whose needs can be addressed thanks to project results. In terms of community building, this means ensuring that potential stakeholders are aware of the project and its potential benefits, and that they are kept updated on new results and their applications. For training, this goes a step further by providing potential users with the knowledge and tools necessary for using EXTRACT technology.

From M1-M15, the focus of the community building and training activities has been to identify potential stakeholder communities who could benefit from the project results and to create relationships with them. Successful examples of EXTRACT's initial activities in this regard include contributions to collaborative platforms such as the ADR Awareness Centre, the AIONDemand Platform, meetings with its HORIZON-CL4-2002 DATA-01-05 cluster, the HiPEAC Workshop on CompContinuum, and targeted event participations that build interest around specific aspects of project technology. As the project technology matures, a greater emphasis will also be placed on trainings, as defined in D5.1. From M1-M15, the focus has been placed on internal trainings, with external trainings envisioned to play a greater role from M24 on.

Community building

Community building focuses on providing stakeholders with a deeper understanding of the project results, and by converting interested audiences into project stakeholders. The WP5 team, together with the entire consortium, has worked to identify and grow the EXTRACT community by using the communication and dissemination channels described in Chapters 3 and 4. Synergies and collaboration have been pursued 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) to find areas of overlap and additional channels to share information about the project and its results.

Since M1, the EXTRACT project has participated in two Big Data Value Association (BDVA) events ('Get to Know' event and Data Week 23), it is a collaboration of the Data Space Support Centre, and its profile is accessible via the AIONDemand Platform. Its participation in these events offered the opportunity to see what other projects are underway and to identify potential overlaps. Additionally, steps have been taken to work with other projects in the HORIZON-CL4-2002 DATA-01-05 cluster to amplify EXTRACT messages and to identify common issues that could be exploited to take advantage of each other's networks. Along these lines, the EXTRACT project participated in a lively exchange of technologies and knowledge together with six other

European-funded projects working to pursue a more efficient and impactful compute continuum as part of the 2024 HiPEAC Conference. The workshop, "[CompContinuum: Computing Continuum of Cloud, Edge, and IoT Technologies](#)" was held on 19 January 2024.

Figure 17: HiPEAC 2024 Workshop: "CompContinuum: Computing Continuum of Cloud, Edge, and IoT Technologies" on 19 January 2024.



Activities tailored to specific audiences have also helped initiate relationships with specific communities where project technology could be used. For example, Observatoire de Paris and BSC have been working within the astronomy and adaptive optic communities to highlight the potential benefits of the project in these fields, while BINARE has brought attention to the EXTRACT project at security-focused events. Participation in industrial events has also granted access to smart city and telecommunication industrial communities who have also been made aware of the opportunities project technology could offer for their extreme data mining needs. Partner Venice has also taken steps to engage the general public through its recorded simulation of emergency movements engaging retired policemen and women volunteers.

While M1-M15 has provided an initial introduction to the communities and a means of understanding their needs and requirements, M16-M36 will delve deeper into creating lasting relationships. Synergies between T5.2 with T5.4 on Sustainability will be pursued to advance this activity.

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.

Internal Training

Internal training helps to connect different project partners and ensure that the project technology addresses the needs of the end users and provides innovative solutions that can be used across domains.

Given the diverse background of the consortium, internal training is important for aligning goals and to ensure a common understanding of project technology. Weekly technical meetings are held by the consortium to ensure that user requirements are clear, and that the technology being developed is appropriate to user needs. Dedicated internal training on specific technologies has also been held remotely to provide additional training on specific technologies and use-case requirements.

Moreover, two in-person training sessions were held during the project's second and third in-person meetings, held at the site of the project use cases, (Venice, Italy and Nançay, France) to provide a better understanding of the infrastructure requirements of the use-cases. The internal training plan has deviated from the initial plan proposed in D5.1, based on the actual needs of the project. Instead of one internal training per year, two were held during M1-M15, with another planned for June 2024.

The June 2024 internal training will take the form of a two-day Hackathon, hosted in Barcelona. It will be open to all project partners and their institutions. One focus of the hackathon will be to work in close collaboration with the use cases to ensure their needs are met by the technical work packages. It will also include a session presenting their technologies and uses, which will be recorded and used for external training in the future.

KPI

The key project indicator set for this activity is 3 internal trainings, including a Hackathon. This KPI is on track with 2/3 internal trainings having taken place in M1-M15 and a third training planned for June 2024.

External training

Transferring the knowledge gained from internal training to interested stakeholders is a key aspect of the external training activity. In M1-M15, the consortium had not planned to hold any external training. External training were expected to begin in M18 with an external Hackathon hosted in Barcelona.

However, this initial plan has been adapted based on the project's needs. For one, project coordinator, Eduardo Quiñones participated in a summer school mobility course, in which he gave a [tutorial](#) on EXTRACT project technology. This opportunity arose as an occasion to combine training on extreme data and AI in cities with a practitioner target audience looking to apply this information to mobility in cities.

Moreover, the consortium decided to change the 2024 Hackathon in Barcelona to an internal versus external training. Partners believe that doing so will help advance the maturity of project results, and allow for more useful external training from M18 on. The 2024 internal Hackathon will include a session that will allow partners to present their individual technologies and uses.

This session will be recorded, edited and used for future external training. Depending on the outcome, these recordings could then be offered as a resource in EU/ AI ecosystem platforms (ADRAe, AI-on-Demand, cluster initiatives). Further external training opportunities in the form of a Hackathon, workshop and a tutorial will be offered once the technology is more developed (from M18 on), so that it is possible to show its value to potential users. Overall, the consortium is confident that it is on track to meet its goals for external training. Sharing external training with its emerging

community will be key to ensuring a core following that can benefit from project results.

KPI

The key project indicator set for this activity is 1 workshop, 1 tutorial and 1 Hackathon. This KPI is on track to being met with 0/1 workshop, 1/1 tutorial and 0/1 Hackathon.

6. Updated Exploitation Plan and Report

This section presents the status of the EXTRACT project exploitation activities from M1-M15. It reports on the status of the initial exploitation plan, which was prepared as part of D5.1 Dissemination, exploitation and community building report, submitted at M6.

The definition of 'exploitation' within the Horizon Rules for Participation is articulated as follows: "Every beneficiary is required, up to four years after the period specified in [GA] Article 3, to implement measures to ensure the 'exploitation' of its outcomes. This includes (a) utilizing them in additional research activities, (b) developing, generating, or marketing a product or process, (c) establishing and offering a service, or (d) incorporating them into standardization activities."

Exploitation activities encompass actions like utilizing project results in research activities extending beyond the project's scope, creating and marketing products or processes, delivering services, etc. The primary objective of the exploitation plan is to harness results for societal, scientific, financial, and various other purposes.

Project partners are the initial entities tasked with exploiting project results, either independently or by facilitating exploitation by others (e.g., by making results accessible under open licenses). This can be achieved through actions such as innovation management, copyright management, data management plans, and stakeholder/user engagement. Common tools employed in this regard include patent applications and publications, establishing spin-off or start-up companies, licensing practices (open or copyleft), and utilizing results for academic purposes (PhD, post-PhD, follow-up projects).

IPR/exploitation contact point identification and listing

As EXTRACT endeavours to harness and commercialize its technological innovations, a critical facet involves the identification and listing of Intellectual Property Rights (IPR) alongside the establishment of dedicated Exploitation Contact Points. This strategic initiative aims to safeguard and leverage the intellectual assets developed throughout the project. The meticulous identification of IPR encompasses patentable technologies, unique algorithms, and proprietary methodologies emerging from the project's research endeavours. Concurrently, the project established dedicated Exploitation Contact Points, acting as liaisons between the project consortium and external stakeholders interested in leveraging or licensing EXTRACT's technological

advancements. This proactive approach not only ensures the protection of intellectual property but also streamlines the process for interested parties to explore collaborative opportunities, fostering a dynamic ecosystem around EXTRACT's innovative solutions. Extract exploitation contacts for consortium partners presented in Table 2.

Table 2: EXTRACT exploitation contacts for consortium partners.

Partners	Contact	Comment
BSC	Alba Abia Hernandez <alba.abiahernandez@bsc.es>	
IKL	Jesus Ruano <jmruano@ikerlan.es>	
URV	Àngels Luque <angels.luque@fundacio.urv.cat> Daniel <daniel.barcelona@urv.cat> Pedro <pedro.garcia@urv.cat>	Angels Luque responsible for managing IPR, funding, spin-offs, etc. Daniel and Pedro responsible for project-related questions, like specific KERs
OBS	SRCV (Héloïse Risac)	
CNRS		
UP		
LOGOS	Raffaella Santucci <raffaella.santucci@logos-ri.eu>	
VEN	Enrico Gavagnin <enrico.gavagnin@comune.venezia.it>	
SIX	Agathe Veillon <agathe.veillon@sixsq.com>	
BNR	Andrei Costin <andrei.costin@binare.io>	
IBM	Erez Hadad <erezh@il.ibm.com>	
MATH	Stefano Cuomo <stefano.cuomo@mathema.com>	

Preliminary Key Exploitable Results (KERs) Identification Steps

In the pursuit of maximizing the impact and value derived from the EXTRACT project, a crucial aspect is the systematic identification of Key Exploitable Results (KERs). These KERs represent the tangible outcomes and innovations that hold significant potential for commercialization and widespread utilization. The process of identifying these results involves a meticulous examination of the project's achievements, breakthroughs, and technological advancements. This includes scrutinizing enhanced data infrastructures, novel MRLAI & big-data frameworks, data-driven orchestration mechanisms, and integrated computing technologies developed within EXTRACT. By pinpointing these Key Exploitable Results, the project aims to not only showcase its technological prowess but also pave the way for strategic commercialization efforts that align with the broader goals of addressing extreme data challenges across diverse sectors. Table 3 presents KERs at stage M14 and Table 4 presents potential KERs at the end of the project.

It is important to note that the below KERs represent a vision from the current project and market condition standpoints, nevertheless it is important to keep in mind that

both markets and end-users need shift or change which may affect in turn the actual KERs evolution throughout the project and even after the project completion.

Table 3 of KERs has been updated to include descriptive information about the value of the results as requested by the reviewers. The same is being updated in the M18 technical report. More details about the data management related to the datasets mentioned below can be found in the latest update of D6.2 Data Management Plan.

Table 3: EXTRACT KERs **identified or in-planning** at stage of M14.

Partner	KER name	KER type (data, software, hardware, service documents, other IPR)	TRL <u>now</u> (estimated)	TRL <u>end of the project</u> (estimated / desired)	MRL <u>end of the project</u>	KER Description and added value for the project
BSC	MARL	software	TRL3	TRL6	MRL4	The 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	COMPSSs	software	TRL4	TRL6	MRL4	COMPSSs 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.
BSC	Multi-cluster Solution	software	TRL2	TRL6	MRL4	A multi-cluster solution connects and manages multiple Kubernetes clusters across different locations or infrastructures. It helps balance workloads between clusters, improving resource use and system reliability.

IKL	K8s Monitoring tool	software	TRL2	TRL4		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		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		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.
ODP CNRS UP	Use Case A: Event database	dataset	N/A	N/A		Use case A provides real-time high resolution event classification. This will replace the visual inspection currently in place (typical event duration is less than 100ms and the observation durations are several hours). The events are fed into a database which will be published openly.
ODP CNRS UP	Use case A: ML model coefficient	dataset	N/A	N/A		The ML algorithms are openly published, as well as the modelled coefficients obtained through training, allowing for reuse and traceability of processing.
ODP CNRS UP	Use case C: Orchestration strategy ML model coefficient	dataset	N/A	N/A		The ML model coefficients of the orchestration strategy allow the reuse of the orchestration model.

ODP CNRS UP	Use Case C: orchestration model	document / knowledge -base	N/A	N/A		The orchestration model documentation describes how the workflows are orchestrated, and how to configure extra components (e.g., new workflow tasks).
ODP CNRS UP	Use Case D: dynamic imager	software	N/A	N/A		The dynamic imager software will implement interferometric imaging capabilities (currently only static imaging is available), enabling video reconstruction techniques.
ODP CNRS UP	Use case D: ML model coefficients	dataset	N/A	N/A		The dynamic imager modelled ML coefficients obtained through training, allowing for reuse and traceability of processing.
LRI	UDT	software	TRL2	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.
VEN	USE PER case	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.

SIX	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.
BIN	Compute Continuum Security Analysis	service	TRL1	TRL4	MRL2	This component is a consulting + training service type of KER planned by BIN. It leverages the BIN Cybersecurity Toolkit for Compute Continuum Software (see above) and couples it with deep-tech know-how within BIN 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 follow "secure by default" paradigm and methodologies.
BIN	Compute Continuum Security Toolset	software	TRL1	TRL3	MRL3	This KER includes all the essential tools and safeguards needed to secure the entire compute continuum 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	TRL4	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.

Table 4: **(Additional)** EXTRACT KERs **envisioned** for the end of the project (NOTE: not final/committal, and as seen from the current M15 perspective)

Partner	KER name	KER type (data, software, hardware, service documents, other IPR)	TRL <u>now</u> (estimated)	TRL <u>end of the</u> <u>project</u> (estimated / desired)	MRL <u>end of the</u> <u>project</u>
BSC	MARL	software	TRL3	TRL6	-
BSC	COMPSs-k8	software	TRL4	TRL6	-
BSC	Multi-cluster Solution	software	TRL2	TRL6	-
BSC	Scheduling Solution	software	TRL1	TRL6	-
BSC	HPC-PER Infrastructure	hardware	TRL4	TRL6	-
IKL	K8s Monitoring tool	Software	TRL2	TRL4	
URV					
VEN	USE case PER	dataset	N/A	N/A	N/A
IBM	Advanced Model Serving on K8s	Software	TRL1	TRL3	MRL1
MATH	Data Security and Privacy Solutions	Software	TRL3	TRL6	TBD

BSC is actively advancing research to seamlessly integrate HPC technologies into the continuum, specifically targeting robust support for handling extreme data. The strategy involves enhancing BSC's HPC frameworks by incorporating essential connectors to facilitate the development of data mining workflows. This innovation aims to significantly broaden the applicability of BSC frameworks, catering to a diverse range of use cases with extreme data requirements. An integral aspect of BSC's exploitation plan involves the intentional integration of EXTRACT's outcomes with those of other pertinent BSC projects, notably the EU initiatives CLASS, ELASTIC, and eFlows4HPC. The overarching goal is to accumulate knowledge that contributes to comprehensive continuum solutions, marking a significant step forward in addressing the challenges posed by extreme data scenarios.

IKL is actively engaged in the development of the Edge-to-Cloud Continuum Monitoring Platform, named Extract monitoring module. This platform builds upon IKERLAN's existing background, incorporating new functionalities such as monitoring CPUs, power consumption, memory, etc. The project aims to create a distributed data infrastructure for securely collecting information across the entire computing continuum. The exploitation plan involves providing added value to customers by automating the control of systems within the edge-to-cloud continuum, optimizing deployment through efficient orchestration. IKL's solution differentiates itself in the market by combining all components into a single solution within the European software ecosystem. The identified exploiters and potential user companies, including LKS, SIXSq, Lantek, Ederlan, Mondragon Assembly, Goizper, Capital Energy, Orona, and CAF, highlight the broad applicability of the developed asset. The protection strategy involves software registration or copyright, with a planned launch in 2026 and ongoing development efforts until 2025. IKL as an affiliated cooperative within the MONDRAGON company, is committed to leveraging the insights acquired from this

project to assist industrial firms in optimizing the utilization of their analytics platforms and workflows, particularly those with demanding data requirements.

LRI, a non-profit research organization, actively contributes to the EXTRACT project, focusing on urban complexity as a Dynamic System of Socio-Technical Systems. The organization aims to leverage EXTRACT outcomes to reinforce its position in security, safety, and resilience at the Italian and European levels, fostering new agreements with municipalities. LRI is dedicated to enhancing its Urban Digital Twin prototype for integration into evolving Smart City Control Rooms in various Italian cities. Adopting Open Innovation 2.0 and the Quintuple Helix Model, LRI aligns innovation strategies with societal needs, emphasizing collaboration with academia, industry, government, civil society, and environmental concerns. In terms of exploitation, LRI explores the potential of Urban Digital Twin technology to address urban complexities, collaborating with local governments, urban planners, and communities for informed policy decisions and improved urban management. The organization's strategic plan involves integrating open data platforms, developing public-private partnerships, and actively pursuing opportunities offered by FESR funds and a network of public and private institutions. This collaborative effort underscores LRI's commitment to harnessing resources and institutional support for innovation within the PER Use Case and the EXTRACT project.

The URV Cloud and Distributed Systems Research Lab actively collaborates with various global companies and research centres, fostering technology exploitation. Leading the CloudStars open-source Mobility network and collaborating with industrial clusters and innovation hubs, URV maintains strong links in bioinformatics with entities like the UK-Health Security Agency and the European Molecular Biology Lab. URV plans to leverage this extensive network for exploitation targets. Moreover, having secured Spanish National funding to create a startup on Cloud Analytic technologies, URV anticipates benefiting from and exploiting EXTRACT results across diverse domains, including bioinformatics. In managing intellectual property, URV emphasizes open-source software development, intending to protect it with widely recognized licenses such as Apache 2, enabling commercial exploitation. Internally, URV plans to utilize the results for academic purposes, license them as open-source, and consider the possibility of establishing a spin-off or startup based on the outcomes.

IBM's exploitation plan for EXTRACT revolves around two key assets. The first is a model serving framework, with the goal of influencing future versions of IBM model serving offerings, both in off-prem (public clouds) and on-prem (private data centers). The second asset is Sky Store, a research prototype developed in collaboration with UC Berkeley. IBM aims to impact its future cross-cloud offerings by leveraging Sky Store, which functions as a global object-storage backbone available across different clouds and premises, featuring efficient policy-based caching and data transport. Regarding the legal aspect, IBM emphasizes the need for complete freedom in accessing its own results during and post-project, and any exploitation agreement must align with this policy.

Binare Oy (BIN, BNR), a deep tech cybersecurity spin-off from the University of Jyväskylä in Central Finland, envisions a two-fold exploitation strategy. First, is to develop a highly specialized know-how and knowledge related to the security needs, requirements and approaches related to the Compute Continuum as one of the upcoming new-wave computing paradigms. All this know-how and knowledge-base subsequently will be offered on a Service-based model (e.g., expert services, highly-specialized consultant services) related to security architectures for Compute Continuum solutions and deployments as well as security best-practices and

approaches for Compute Continuum applications. Second, is to leverage the expertise and development into a set of software tools that can assist Compute Continuum users and organizations to either secure (e.g., configurations and configuration-checks) or to assess the security of existing deployments for compliance and internal/external needs of the organizations developing or using Compute Continuum solutions.

Mathema (MATH) is an ICT company active since 1987 whose main mission is to support very large public & private organisations in their digital transformation and innovation. In particular, the company ideate, design, fast prototypes/MVP and deliver innovative and non-standard secure complex information systems based on Big Data, Blockchain, VR/AR tool for Command and Control and exploration, Artificial Intelligence based tools for mining of large data sets. Mathema has been acknowledged as PMI Innovativa (Innovation Performing SME) according to a recent Italian law (d.l. 3/2015 "Investment Compact"). The know-how matured along its lifetime will allow to develop custom-oriented solutions for AI, such as the simulator developed on-purpose for the PER and cyber-security solutions. The main asset that MATH is developing for market exploitation purposes is related to advanced solutions for data security and data privacy. MATH is currently developing a Python-based framework for data security/privacy/integrity grounded on advanced techniques such as Differential Privacy, Full Homomorphic Encryption and MultiParty Computation that can be of high interest to companies/institutions exchanging great volumes of data with high security requirements.

SixSq (SIX) is a Swiss software company specialized in edge computing, which develops the Nuvla technology to manage edge-to-cloud infrastructures. Nuvla is composed of two products:

- Nuvla.io, a Software-as-a-Service platform hosted on the cloud, or on premise. The platform allows remote management and orchestration of edge devices running NuvlaEdge. Each NuvlaEdge is uniquely identified on the platform, and can be managed individually or as part of a fleet. Management functions include the deployment and update of edge apps, the monitoring of the edge device (status, resource consumption, etc.) as well as data management basic functionalities.
- NuvlaEdge, a containerised piece of software which runs on edge devices. It allows communication with Nuvla.io, and remote management of the edge device (apps deployment, device monitoring, etc).

Nuvla.io and NuvlaEdge are used in the EXTRACT project for data management features, namely the data catalogue. Even though Nuvla.io and NuvlaEdge are in production (TRL9), SIX is developing from scratch new features on top of the data catalogue, to enhance edge-to-cloud data management. This development work is mapped in the table above with a lower TRL. It will be of high interest not only to the EXTRACT project, but also to companies or European institutions with large datasets who seek to improve on a qualitative and quantitative basis their data analysis, such as the European Space Agency.

The Personalized Evacuation Routing (PER) System, developed by the Metropolitan City of Venice as part of the EXTRACT project, holds significant potential for broader dissemination and commercialization. This report outlines how the innovative approach adopted by the city can serve as a catalyst for introducing the PER System to a wider audience, including potential commercial partners like MyTraffic. Moreover, it identifies key stakeholders in the Veneto region with a vested interest in digital innovation. Utilizing the PER System for Wider Adoption involves collaborating with local media outlets to promote its benefits in enhancing public safety during emergencies,

partnering with universities and research institutes to conduct workshops and seminars on its technological advancements, and hosting live demonstrations in high-traffic areas to showcase its effectiveness to the general public. Commercialization opportunities include forming a strategic alliance with MyTraffic and Veneto Innovazione SPA to integrate the PER System into their traffic management solutions and collaborating on marketing campaigns to promote it to transportation authorities and urban planners. The PER System presents an opportunity for the Metropolitan City of Venice to enhance public safety and stimulate economic growth through strategic partnerships and collaborations with key stakeholders in the Veneto region, leveraging the support of public and private entities interested in digital innovation to achieve widespread adoption and commercial success.

Preliminary Exploration of Tools for Exploitation of KERs

As EXTRACT progresses towards its exploitation and commercialization phase, strategic collaboration with funding and exploitation agencies at both the EU and country levels emerges as a pivotal element of success. The project recognizes the importance of aligning its objectives with the priorities and initiatives of funding agencies, both within the European Union and the individual countries involved. By actively identifying and engaging with these agencies, EXTRACT aims to tap into potential funding opportunities, leverage existing programs, and establish partnerships that support the widespread deployment and utilization of its innovative solutions. This dual-level approach ensures a comprehensive strategy, allowing EXTRACT to navigate the diverse landscape of funding and exploitation opportunities, thereby maximizing its impact on a broader scale.

Below are presented tables with potential funding opportunities divided by the countries.

Table 5: Funding opportunities in Finland.

Name	Description
Business Finland	Companies and research groups work with BF and receive funding for research, development, and innovation activities. Business Finland works with the top innovative companies and research units in Finland. Research to Business funding is intended for public research organizations for projects in which research groups and researchers aim to develop their research into new businesses and to commercialize their research results.
The Riitta ja Jorma J. Takasen Foundation	The founders of Scanfil plc, Riitta and Jorma J. Takanen, established the foundation in 2005 to promote the development of know-how in the technological industry, and to support international competitiveness. The foundation awards grants for students enabling them to participate in foreign exchange and research work, such as doctoral research. In addition, the foundation endorses research innovations and commercialisation.

Table 6: Funding opportunities in Italy.

Name	Description
Cassa Depositi e Prestiti	CDP is a unique group in the Italian landscape, combining financial and industrial capabilities. CDP uses its skills to serve Italy and the other countries we operate in, promoting growth and employment, supporting innovation, and the competitiveness of businesses, infrastructure, and the country. From start-ups to large industrial groups, CDP Group helps make Italian enterprises more competitive, supporting the country's growth. CDP's integrated offering of financing, equity instruments, and advisory

	services helps enterprises achieve their innovation, growth, and international expansion goals.
Confindustria Nazionale	Confindustria is the main association representing manufacturing and service companies in Italy, with a voluntary membership of more than 150,000 companies of all sizes, employing a total of 5,383,286 people. The association's activities aim to guarantee the central importance of companies, the drivers of Italy's economic, social, and civil development. By representing companies and their values at institutions of all levels, Confindustria contributes to social well-being and progress, and from this standpoint guarantees increasingly diversified, efficient, and modern services. Confindustria's added value stems from its network: a headquarters in Rome, 216 local and trade associations, and since 1958 a delegation in Brussels, which has seen its strategic importance grow over the years and is now a benchmark for the entire Italian economic system at the European Union. Confindustria offers several chances for funding mainly articulated at the local (regional) level.
Regione Toscana POR FESR	The Tuscany Region offers funding opportunities for businesses through the POR-FESR calls to strengthen investments in the Tuscan entrepreneurial system by encouraging the introduction of advanced technologies to promote ecological transition and cooperation among small, medium, and large enterprises and research organizations. The resources available through these two calls of the Regional Program (Pr) of the European Regional Development Fund (Fesr) 2021-2027, approximately 125 million euros, will enable the development of investments totaling around 400 million euros, allowing many businesses to be competitive and promote innovation and creativity.

The measure indeed provides significant resources for research and development projects to stimulate the Tuscan productive system to make a major effort in terms of innovation.

Table 7:Funding opportunities in Switzerland.

Name	Description
Innosuisse Assisted Patent Search	Swiss SMEs receiving funding from the Swiss Innovation Agency can benefit from IPR management services, in cooperation with the Swiss Federal Institute of Intellectual Property. The two bodies mainly offer expertise in patent search and patent landscape analysis. For more advanced projects, additional IPR management tools are offered, such as registration of a trademark, advanced patent search, and development of an IP strategy, and licensing.
Innosuisse Start-Up coaching	Innosuisse helps startup founders launch their businesses thanks to startup coaching programs, focusing on the different stages of business creation.
Swiss Accelerator	For existing startups/SMEs, the Swiss Accelerator frequently publishes calls for tenders around highly innovative projects, intending to quickly go to the market.

Table 8:Funding opportunities in Spain.

Name	Description
Ayudas para Redes de I+D+i	What is of interest is not the call itself, but the networks created in these calls, which aim to promote collaborations between research groups and private entities by providing them with mechanisms for technology transfer and knowledge valorisation.

MERCAT-PoC	Combines a financing operation in the form of a convertible equity loan or equity granted by Financial Instruments for Innovative Enterprises (IFEM) and aid in the form of a guarantee granted by the Department of Research and Universities.
ACCIÓ - Programa Indústria del Coneixement, Modalitat C. Innovadors	The purpose of the grants is to finance projects for the recruitment of a person with an entrepreneurial profile who will carry out the activities necessary for the transfer to the productive sector of technologies or knowledge generated in the public research system, through the creation of a spin-off or the exploitation of any form of intellectual or industrial property with licensing or co-development agreements.
ENISA	Loans are designed to boost Spanish small and medium-sized enterprises with viable and innovative projects.
CDTI Projectes d'innovació: Línia Directa d'Expansió (LICA)	Loan to promote innovation in certain Spanish regions that are more disadvantaged and/or experiencing particular economic difficulties, through the improvement of the capacities of companies that propose technologically innovative investment projects that facilitate their growth.
CDTI Projectes de R+D	Lending to R&D projects developed by companies and aimed at the creation and significant improvement of production processes, products, or services. They may include both industrial research and experimental development activities. There is no restriction as to the sector or technology to be developed.
CDTI Projectes d'innovació: Línia Directa d'Innovació (LIC)	Aid to support projects of an applied nature, very close to the market, with medium/low technological risk and short payback periods, which improve the competitiveness of the company through the incorporation of emerging technologies in the sector.
CDTI - EUROSTARS	Eurostars is aimed at generating R&D projects carried out by small consortia and led by innovative SMEs. These projects must represent a break with the technical state of the art and a

	commercial challenge that will allow these companies to make a significant qualitative leap in their market position.
ICF - Venture Tech	ICF Venture Tech is a venture capital fund aimed at investing in technology and/or innovation projects with high growth potential in their early stages.
ELKARTEK	Non-refundable grants to support the implementation of Collaborative Research, carried out by the Research, Development, and Innovation Entities integrated in the Basque Science, Technology, and Innovation Network of the Basque Country, in the areas of specialization framed in the Science, Technology, and Innovation Plan Euskadi 2030.
HAZITEK COMPETITIVO / HAZITEK ESTRATÉGICO	The Hazitek program, regulated by the Department of Economic Development, Sustainability and Environment (SPRI) is composed of non-refundable grants, to provide the necessary resources for Industrial Research or Experimental Development. This program has been designed to promote Research and Development projects for SMEs, large companies, and associations in the Basque region.
Knowledge Industry grants for 2023. Modality B. Producte Grants. Agency for Management of University and Research Grants	The grants are intended for innovation projects that have completed the conceptualization of a transfer plan and have identified a candidate for a product or service to be developed. They should have a proposal for value generation based on the proposed technology and/or knowledge, following the validations carried out.
Knowledge Industry grants for 2023. Modality C. Innovadors. Agency for Management of University and Research Grants	The purpose of the grants is to fund projects for hiring individuals with an entrepreneurial profile who will carry out the necessary activities for transferring technologies or knowledge generated within the public research system to the productive sector. This can be done through the creation of a spin-off or the exploitation of any form of intellectual or industrial property through licensing or co-development agreements.

The Collider: deep tech venture builder	Mobile World Capital Barcelona's dedicated tech transfer program supports scientific and innovative projects in their transition from the lab to the market. Through its multiple initiatives, The Collider works with ecosystem agents to strengthen technology transfer mechanisms in Catalonia, Spain, and Europe, facilitating the consolidation of scientific-technological startups.
Prueba de Concepto del Ministerio	its purpose is to promote and accelerate the transfer of knowledge and results generated in research projects of the State Plan, through the modalities of "Excellence" or "Knowledge Generation" R&D projects, of "Research Challenges" R&D projects and of oriented fundamental research projects.
CDTI (for companies)	Public business entity, dependent on the Ministry of Science, Innovation and Universities, that promotes innovation and technological development of Spanish companies. It is the entity that channels requests for help and support for R&D&I projects of Spanish companies at the national and international levels. Consist of loans partially repayable.
ACCIÓ-START-UP CAPITAL (for companies)	Startup Capital is a direct grant of up to €100,000 for recently created technological startups. This grant is aimed at startups that must be constituted at the time of publication of the call in the DOGC and must have a maximum of 18 months of life at the moment of request. The grant is a non-repayable public award and does not require a guarantee.
ACCIÓ- NUCLIS (for companies) Business enterprise R&D nuclis	Through these grants you can receive up to 250,000 euros in non-returnable funds to develop a different project and tackle the challenges of the technology and green transformation.

Table 9:Funding opportunities in EU.

Name	Description
Eurostars	Eurostars is the largest international funding program that supports research and development projects led by small and medium-sized enterprises (SMEs). The grant scheme focuses on promoting international collaboration among partners/SMEs from different Eurostars participating countries to develop innovative products, processes, or services with high market potential.
EIC Accelerator	It supports startups and high-risk and high-potential SMEs which are willing to develop and commercialise new products, services, and business models that drive economic growth and shape new markets or disrupt existing ones in Europe and worldwide. EIC Accelerator is open to innovations in any field of technology and supports the later stages of technology development (>TRL5) as well as scale-up. Grant funding of up to €2.5 million and equity investment of up to €15 million are available to support the development and scaling-up of European high-risk innovations!
EIC Pathfinder	The overall objective of the EIC Pathfinder for advanced research is to develop the scientific basis to underpin breakthrough technologies. It provides support for the earliest stages of scientific, technological, or deep-tech research and development. EIC Pathfinder supports early-stage development of future technologies like activities at low Technology Readiness Levels 1-4 and up to proof of concept. Pathfinder projects can also receive additional funding for testing the innovation potential of their research outputs.

EIC Transition

EIC Transition Open is an open funding scheme for projects in any field of science or technology as well as challenge-driven funding on specific strategic fields. The majority of funding will be granted through open calls with no predefined thematic priorities.

EIC Transition projects should address both technology and market/business development, possibly including iterative learning processes based on early customer or user feedback. Grants of up to EUR 2.5 million are available to validate and demonstrate technology in application-relevant environments and develop market readiness.

Who can apply ranges far - SMEs, spin-offs, start-ups, research organizations, and universities) or small consortia (two partners) or consortia of three to five different independent legal entities established in at least three different eligible countries.

Throughout the M15-M36 timeline, each project partner will dedicate focused efforts to delve deeply into the specific instruments outlined earlier. This involves a comprehensive exploration of each tool, with a specific emphasis on identifying at least one potential exploitation or commercialization instrument, contingent upon qualification and relevance. The partners will diligently undertake preliminary work in connection with these instruments and subsequently document their findings in D5.3. This proactive approach aims to lay the groundwork for potential joint or non-joint exploitation and commercialization initiatives, thereby ensuring preparedness for strategic endeavours in the later stages of the project.

Moreover, the exploitation plan includes applying and completing the EU Innovation Radar questionnaires. As suggested by the PO and the reviewers during the mid-term PR1 review meeting and in the review report, the EU Innovation Radar mid-term questionnaires will be completed latest during M22 (Oct 2024) and this activity will be led by the BIN (T4.3 lead and Exploitation lead) in collaboration with all the partners. The questionnaire's fulfilment will be in line with the exploitation and KER activities performed so far as well as planned till the end of the project.

Preliminary Market Understanding and Innovation Gap Analysis

Suggested New Product

Data-mining platform software with computing continuum / extreme scale data processing and machine learning, with flexible security integrated at multiple layers: in the cloud, at the edge, for OT, IoT deployments, and in application domains such as, for example, smart cities.

Potential Need for the Suggested New Product

The important aspect to consider in this question is the huge amount of data generated daily and its complexity fueled by the recent advances in Cyber-Physical Systems, the Internet of Things (IoT), and Artificial Intelligence (AI): "According to the World Economic Forum, 463 exabytes of data will be daily generated worldwide by 2025 [Wefor], whereas the European big data and business analytics market is projected to reach \$105.82 billion by 2027 [Allma]."

Competitors analysis

Competitors in the compute continuum market will be analyzed to get a better understanding of the market and how competitive it is. This is a critical element in creating a successful exploitation plan.

Arm: Arm Neoverse Platform. Neoverse N1 for cloud-to-edge computing.

Intel: Intel® Virtualization Technology (Intel® VT); Intel VT, including Intel® Virtualization Technology® for Directed I/O (Intel® VT-d); Intel® Trusted Execution Technology (Intel® TXT); Intel Web APIs. These are the technologies that support the computing continuum.

The computing continuum is a new notion formed by the recent increase of connected devices and systems. The notion is more in a research phase. Recent innovation. Therefore, the competition is very low as of this writing (Mar 2024).

No direct and/or immediate competitors for mature or complete "Compute Continuum with ML/DL" were identified as of now (Mar 2024). The identified competitors seem to be indirect as they develop technologies that support the computing continuum technologically. Nevertheless, such technology providers could be the first movers in the market, therefore EXTRACT project will monitor those closely.

What are the product's advantages in comparison to the competition?

Extreme data has a sophisticated nature as it simultaneously exhibits several of the following characteristics to an extensive degree: (1) volume, speed (including the need for real-time guarantees), and variety; (2) complexity and diversity; (3) dispersity of

data sources; and (4) existence of sparse, missing, insufficient data, often with extreme variability. Current solutions start to fail when coping with data of such complexity. "The reason behind this technological gap is that current programming practices and computing technologies are optimized to cope with some of these data characteristics independently and uniformly, rather than holistically addressing them." The extract platform aims to address the extreme data characteristics holistically.

The main competitive advantage of the platform is its uniqueness as no direct competitors are identified as of now.

Potential customers

The European Cloud Edge IoT Continuum for business and research Use Cases:

"Increased density of edge and cloud facilities allowing ubiquitous connectivity will enable a wide range of innovative use cases based on ultra-low latency infrastructure. The use cases underpinning the **European Cloud, Edge & IoT Continuum** initiative tackle key societal challenges, including healthcare, carbon emissions, safety, privacy, resource optimisation in storage, network and computing centers, synchronization and Digital Twins, energy, and environmental footprint.

Some application areas and examples include **smart agriculture** (e.g. precision farming, data autonomy of vehicle swarms, CO₂-neutral intelligent farming, smart robots), **renewable energy production centers** (e.g. reliability and security of power grids), **smart homes** (sensors and robots for remote monitoring, enablement of domestic services), **factory robotics** (e.g. reducing CO₂ in production, predictive maintenance), **logistics and transport** (object recognition, energy, and process optimisation in movements, predictive maintenance of railway networks, smart ports, last-mile delivery optimisation), **healthcare** (e.g. remote real-time medical diagnostics with machine learning and artificial intelligence-based analytics). Particular attention will be paid to the impact of CEI across various sectors, with benefits including energy efficiency, flexibility, and security."

Examples:

[Enhancing Real-Time Connectivity and Agile Decision-Making in Automotive Manufacturing with Multi-Access Edge Computing](#)

[Innovating Quality Management: Pioneering Cloud-Edge-IoT and Digital Twin Solutions in Manufacturing](#)

[The Power of Cloud-Edge-IoT and 5G for Advanced Intralogistics Management](#)

AIOTI Computing Continuum Use Cases:

- Cloud-based medical imaging
- Cloud-based visual inspection in production
- Cloud-based control of automated guided vehicles
- Cloud-based control of production via optical wireless communication
- Protecting data within smart cities

Main target users

The targeted audience includes computer and computational scientists, instrument scientists, experts in ML/AI and digital twins, system architects, as well as end users.

Stakeholders

Several stakeholders can be identified:

Internal Stakeholders:

- **Research & Development:** Data scientists, AI researchers, software engineers, and project management teams spearheading the platform's development and implementation.
- **IT Infrastructure & Security:** IT professionals ensure seamless integration with existing systems, optimal performance, and robust security measures.
- **Technical Advisors:** Experts in fields such as machine learning, natural language processing, and data science provide guidance and expertise to ensure the technical viability and integrity of the project.

External Stakeholders:

- **Policymakers and regulators:** They might be interested in the platform's capabilities for ensuring data privacy and security, especially when dealing with sensitive data.
- **Data Users:** Businesses, researchers, and individuals seeking to harness insights from extreme data mining for various applications.
- **Open-Source Community:** Developers and users contributing to and benefiting from the open-source nature of the platform.

Additional Stakeholders:

- **Investors & Funding Partners:** Providing financial resources to support the project's development and ongoing operations.
- **Media & Public:** Journalists and the general public are interested in the ethical and responsible use of data and its potential impact on society.
- **Academic Institutions:** Universities and research institutions contribute knowledge, resources, and expertise to the project, potentially including faculty members, students, and affiliated researchers.

Potential/Future stakeholders:

- [H-Cloud](#) - The European Commission initiative that aims to consolidate and grow the Cloud Computing research and innovation community in Europe.
- [NGIoT](#) - The NGIoT works to lower the barrier to the adoption and development of IoT-empowered solutions, by supporting business models, innovation, and skills.
- [EUCloudEdgeIOT](#) - The organization bridges the gap between researchers, businesses, and users to develop the Cloud-Edge-IoT (CEI) landscape.
- [Eclipse Foundation](#) - The Eclipse Foundation facilitates a robust ecosystem for open-source software collaboration and innovation, fostering a synergistic environment for individuals and organizations worldwide.

- [Institute of Communication and Computer Systems](#) - The Institute of Communications and Computer Systems (ICCS) is an independent research organization founded in 1989 at the National Technical University of Athens (NTUA).
- [EOSC](#) - Federated and open multi-disciplinary environment with open and seamless services for storage, management, analysis and re-use of research data, across borders and scientific disciplines.

Market segmentation

Size of IoT device market:

[According to the latest available data, there are approximately 15.14 billion connected IoT devices in 2023. The number of active IoT devices is expected to double by 2030.](#)
→ The data generated from IoT devices will be enormous.

- The IoT market is segmented based on the industry/application and regions.
- [According to the IoT-Analytics](#), IoT Market segments include Retail, Healthcare, Energy, Mobility, Cities, Manufacturing, Public & Services, and Agriculture & Military.
- Industries such as telecom, healthcare, government, energy, manufacturing, space, and critical infrastructure are significant consumers of data management solutions that offer **security at all layers** since they have/operate many IoT devices in their networks.

[According to Eurostat](#), countries with high enterprise IoT and HPC usage involve Austria, Slovenia, Sweden, and Finland.

[Data creation by region:](#)

- US
- Germany
- UK
- China
- Canada
- France

Upcoming Exploitation Plan for M16-M36

As per the reviewer's comments in the review report, the M16-M36 Exploitation Plan has been updated to extend the duration of community building and training.

Months M16-M18: Technical Asset Identification and Integration

- Finalizing individual technical assets and Key Exploitable Results (KERs).
- Develop plans to integrate existing tools listed in Table 1 to support extreme data.

- Conduct a preliminary assessment of the Technology Readiness Level (TRL) of each identified asset.

Months M19-M21: Market Understanding and Innovation Gap Analysis

- Deepen the understanding of the project market and exploitation context.
- Conduct a thorough market watch to identify innovation gaps and potential opportunities.

Months M22-M36: Training Activities and Community Building (CB) as a basis for Sustainability Activities

- Define a training and community building plan together with the sustainability task for months M22-M36
 - Set the foundation and tentative plan for webinars and collaboration with similar projects (HiPEAC 2025 M25, ADRA-hosted Birds of a Feather (BoF) session M26, use-case specific tutorials M30, etc)
- Participate in the Big Data Value Association European Big Data Value Forum with other projects funded under other projects in the HORIZON-CL4-2002 DATA-01-05 (session and booth) M22
- Host a EXTRACT Tech-to-Market Hackathon Challenge: Driving Innovation for Sustainability M22
 - Bootstrap initial seeds of community building (mixing tech and business)

Months M25-M27: Exploitation Plan Finalization and collaboration with sustainability on training and CB

- Finalize individual exploitation plans for technical assets and KERs.
- Ensure alignment with market needs, regulatory compliance, and project objectives.
- Host HiPEAC 2025 Compute Continuum Workshop
- Participate in ADRA-hosted BoF with projects in cluster call

Months M28-M30: IPR Management and Business Modeling

- Develop strategies for Intellectual Property Rights (IPR) management.
- Refine the business model, emphasizing how the EXTRACT platform addresses identified market needs.
- Collaborate with legal experts to ensure compliance with the management of knowledge and IPR.
- Define webinar topics on use cases tailored to end-users

Months M31-M33: Market Penetration and Deepening Collaboration

- Execute the go-to-market plan to penetrate the market with the EXTRACT platform.
- Collaborate with consortium members and stakeholders for joint exploitation opportunities.
- Identify potential business models for the exploitation and sustainability of the EXTRACT ecosystem.
- Collaborate with end-users and EU Data Community to share results

Months M34-M36: Sustainability Planning and Final Evaluation

- Develop a long-term sustainability plan for the EXTRACT technology.
- Collaborate with private and public bodies to shape the technology for present and future needs.

- Establish a feedback loop for insights from users and stakeholders.
- Iteratively improve the EXTRACT platform based on feedback and market demands.
- Conduct a final evaluation of the exploitation and commercialization efforts.
- Generate a comprehensive report detailing activities performed, achievements, and recommendations for sustained success.
- Prepare for the transition of the project outcomes into sustained commercialization beyond the project's lifecycle.

7. Initial Sustainability Plan

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 main approaches have been foreseen to ensure the sustainability of the project. The first is to leverage and to promote EXTRACT results, which shall be packaged and integrated into a dedicated marketplace for extreme data mining workflows. The second will work together with Task 5.2 Community building activities to build a strong network of collaborations and synergies with public and private bodies and related solutions.

This Initial Sustainability Plan corresponds to the Task 5.4 Sustainability of the EXTRACT project, which will officially start in April, 2024. This first iteration will describe the first ideas and actions to be investigated, and eventually put in place to ensure the sustainability of the project. A second iteration will be issued at the end of the project, to provide an update of the Plan, and an estimate of sustainable outcomes.

Results and their Sustainability

First, EXTRACT will leverage relevant findings and outcomes to ensure its sustainability.

The main results of the project are complex algorithms designed to perform data analysis in extreme conditions. Within the course of the project, these complex algorithms will be applied to two distinct use cases that are extremely relevant when it comes to the sustainable vision of the project.

The two use cases tackle needs that have not been addressed by the market and provide commercial illustrations in two different fields. These two examples will both demonstrate the ability of the algorithm to quickly analyze vast amounts of data that are periodically updated. The behavior of the algorithm and its ability to produce

relevant metrics will be key in demonstrating the added value of the product and convincing public and private bodies to use it after the end of the project.

Furthermore, the setups used for each of the use cases can easily be replicated in other configurations. One of the objectives of the project is to develop a dedicated marketplace for these algorithms. Relying on the expertise of the different partners and on the existing marketplace developed by SixSq which sells paying and free edge apps, the algorithms will be packaged in a way (Kubernetes) they can easily be deployed from this platform on edge devices and then remotely managed. Being able to demonstrate that these products can easily be integrated into an existing infrastructure and managed will also reinforce the sustainable aspect of the project.

Several actions will take place in Task 5.4 Sustainability to produce the most relevant recommendations so that these two aspects are leveraged. Brainstorming sessions will be organized with the exploitation contacts identified in Task 5.3 Exploitation, to better shape the development and exploitation of EXTRACT outcomes and marketplace.

The possible publication of the results in research platforms such as Horizon Europe or Euresearch will be investigated, as well as more business-oriented platforms. The use of promotion tools such as the Horizon Results Booster programme will also be assessed, as well as the potential application to a private booster programme such as an incubator, in cooperation with Task 5.3 Exploitation.

Self-sustainable Community

Second, EXTRACT will rely on a strong ecosystem of stakeholders, both from the public and private sectors, to establish collaborations and synergies. This will ensure that the EXTRACT solution is, from a business perspective, relevant today and in the future.

Task 5.2 Community building plays a major role with regards to this, in shaping a community which will help the project grow and maybe tackle new use cases. Internal training and events are already happening. External events, webinars and workshops will complement these to widen the horizon of the project and do some networking.

It is also foreseen to rely on the partners' network to build a sustainable ecosystem around the project. The EXTRACT Consortium is made of both public research institutions and private companies, which both participate in the European research ecosystem in their area of expertise.

As a first step, a table has been created for partners to list the other R&D activities they are working on. The completion of this table will enable the identification of potential companies or research institutes of interest that are already in the network of EXTRACT partners.

Then, Task 5.2 Community building and Task 5.4 Sustainability will join forces to organize external events and to promote the EXTRACT project and results.

Finally, the project will make sure that the results, and achievements of the project are acceptable by the business world; mainly for what concerns IPR management, promotion of the results through Open Source communities.

Procedures for joint exploitation agreement will also be assessed to ensure the sustainable exploitation of the project's outcomes.

8. Evaluation of Performance and KPIs (until M15)

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	Met: 2/2 posters On track: 1/3 videos On track: 0/1 infographic
Website	Visitor statistics	Average 1000 sessions per year	On track: 960/1000 sessions Yr 1 (at M12) On track: 1500 total (at M15)
Social media	Number of Twitter followers and LinkedIn group members	At least 150 Twitter 250 LinkedIn	On track: 60/150 Twitter Ahead of track: 224/250 LinkedIn
Press strategy	Number of press releases/clippings	One press release /year At least 25 press clippings	On track: 1/3 PR On track: 15/25 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	On track: 1/2 key notes 0/1 organized events 0/1 booths in industrial event 16 event participations
Publications	Papers in conferences and journals	At least 4 publications per year	On track: 3/4 publications Yr1 On track: 4/4 by M15
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: 1/1 tutorial On track: 2/3 internal training course 0/1 workshop 0/1 external Hackathon

Final event	Organization of a final Project event	Significant attendance 50 or more	N/A
-------------	---------------------------------------	-----------------------------------	-----

9. Conclusion

The EXTRACT project, as detailed in this deliverable D5.2, has made significant strides in its journey towards dissemination, exploitation, sustainability, and community building. Since its initiation on January 1, 2023, the dedicated efforts of the Work Package 5 team, led by partners Binaré, Barcelona Supercomputing Center, and Sixq, have been instrumental in achieving multifaceted objectives.

The report thoroughly documents activities conducted up to month 15, presenting a roadmap for the subsequent months (M15-M36) with updated strategies and a focus on sustainability. The project has been unwavering in its commitment to raising awareness and ensuring longevity beyond its completion.

From M1-M15, the consortium successfully cultivated a recognizable project brand and engaged with target audiences through strategic channels, meeting and surpassing communication KPIs. Noteworthy accomplishments include a robust online presence, with 1500 website visitors and 224 LinkedIn followers at M15.

As the project has evolved, dissemination activities, including participation in events, a keynote speech, and publications, have increased. Despite minor deviations from predefined KPIs, the conviction remains that the project is on a solid track, mirroring the natural progression of technology maturation.

Community building and training, vital for the uptake of EXTRACT technologies, have been strategically approached, involving participation in European communities and adapting training plans to address evolving needs.

The exploitation plan, detailed in the report, is meticulously crafted, focusing on critical aspects such as IPR/exploitation contact points, Key Exploitable Results (KERs) identification, and a comprehensive exploration of the market. This strategic positioning enhances the project's potential for success in the challenging realm of extreme data management and analytics.

The sustainability plan, as outlined, emphasizes a dual strategy. Firstly, the integration of complex algorithms into a dedicated marketplace demonstrates a commitment to ongoing benefits. Secondly, building collaborations with public and private entities aligns with the goal of fostering a supportive community, ensuring the long-term impact of EXTRACT results.

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

Annex 1 - EXTRACT Dissemination Register

EXTRACT Dissemination Register (activities)					
Nº	Type of activity	Date	Details	Comment	
1	Non-scientific publication	15/01/2023	Distributing computations across the continuum opens up huge possibilities	BSC-Kartsakli	Elli
2	Project LinkedIn account launched	15/01/2023	EXTRACT EU Project	ALL	
3	Project Twitter account launched	15/01/2023	@EXTRACT_EU_proj	ALL	
4	Non-scientific publication	25/1/2023	The URV obtains two million euros to promote four international projects on cloud computing and artificial	URV	
5	Non-scientific publication	02/10/2023	'Distributing computations across the continuum opens up huge possibilities	BSC-Kartsakli	Elli
6	Presentation at a Workshop	23/02/2023	Get-to-know" introductory workshop and welcome day to HE Data projects (WP2021-22)	BSC-Eduardo Quiñones presented	
7	Non-scientific publication	26/02/2023	Binare Platform for (I)IoT/Embedded Security, Electronics Today Feb 2023	Binaré 1 page infographic	
8	Project website launch	28/02/2023	https://extract-project.eu/	ALL	
9	Press release	03/03/2023	Horizon Europe project EXTRACT kicks off with a holistic approach to extreme data across the compute continuum	ALL	
10	Non-scientific publication	03/06/2023	EXTRACT; A distributed data-mining software platform for extreme data across the compute continuum	SIXSQ	
11	Dissemination Meeting	03/08/2023	Presentation of EU projects to T-Systems	URV	
12	Dissemination Meeting	15/3/2023	Presentation of EU projects to Arsys	URV	
13	Non-scientific publication	18/3/2023	Artificial intelligence to save lives	URV	
14	Dissemination Meeting	21/3/2023	Presentation of EU projects to Telefonica Research	URV	
15	Non-scientific publication	22/03/2023	BSC Leads Multi-Partner Consortium in Innovative Data Mining Project Across Compute Continuum	BSC	
16	Non-scientific publication	20/04/2023	Run Scalable GPU Workloads on IBM Cloud VPC with Ray and CUDA	IBM Hadad)	(Erez

17	Non-scientific publication	20/04/2023	Get VMs with CUDA and other features on IBM Cloud VPC – Easily!	IBM Hadad) (Erez)
18	Dissemination Meeting	05/09/2023	EXTRACT-TASKA: orchestration, data mining and decision making in radio astronomy	ObsParis (Baptiste Cecconi)
19	Dissemination Meeting	05/12/2023	Online Meeting with Xartec Salut to present the projects	URV
20	Panelist in a conference (Participation + Presentation)	14/06/2023	Building a Cognitive Cloud-Edge Continuum for Next-Generation Data Processing Applications	IKERLAN (TBD)
21	Summer School participation (Community Building)	27/06/2023	CSC Summer School in High-Performance Computing 2023	Binaré (two members participate + disseminate EXTRACT)
22	Non-scientific publication	28/06/2023	IoT-Edge-Cloud: building a simple, secure, and future proof infrastructure	SIXSQ
23	Non-scientific publication	30/06/2023	EXTRACT EU Project: A distributed data-mining software platform for EXtREme data ACross the compute conTinum	IKERLAN
24	Tutorial Mobility course	07/11/2023	"La Digitalizació de La Mobilitat"	BSC- Edu
25	Flyer	30/07/2023	Available online	BSC, all
26	Presentation at a conference	08/02/2023	Multi-party Computation for Privacy and Security in Machine Learning: A Practical Review	MATHEMA presentation
27	Participation in a conference	08/02/2023	2023 IEEE International Conference on Cyber Security and Resilience	LOGOS RI co-sponsor BINARE provided merchandise with EXTRACT logo
28	Poster presentation	28/08/2023-30/09/2023	Nit de la Recerca Europea	BSC, all
29	Non-scientific publication	28/08/2023	EXTRACT: A distributed data-mining software platform for EXtREme dAta Across the Compute conTinum	LOGOS RI
30	Keynote Speech	20/09/2023	AMLE Summer School	BSC
31	Participation in Summit	10/04/2023	CSC's Supercomputing Innovation Summit 2023 (distribute promotional materials)	Binare
32	Participation in industrial conf.	1/11/2023	https://hardware.io/netherlands-2023/ (cybersecurity x IoT x edge x embedded x hardware)	Binare

33	Presentation at industry event (Smart city)	11/09/2023	Edge Computing for a Cleaner and Safer Mobility	BSC-Elli Kartsakli
34	Participation in a workshop	12/11/2023	WoSC 9-9th International Workshop on Serverless Computing	URV
35	Presentation at a conference	13/12/2023	25th ACM/IFIP International Middleware Conference	URV
36	Non-scientific publication	15/01/2024	Taming a universe of data	URV - Daniel Pons, Enrique Molina Giménez
37	Participation in a workshop	19/01/2024	HiPEAC 2024: Workshop on CompContinuum	BSC- Xavier Palomo
38	Presentation at an industry event (MWC)	26/02/2024	https://extract-project.eu/mwc24-opens-door-to-industry/	BSC-Elli Kartsakli, Alba Cañete Garrucho
39	Non-scientific publication	26/02/2024	http://x4hpc.cat/mwc24/custom/themes/100x100/assets/download/Catalogue-BSC-MWC-Agenda.pdf	BSC

Annex 2 - EXTRACT Detailed list of publications

Detailed list of publications

Type	Title	Authors	Title of the Journal or equipment	Date	Link to publication	Link to repository	Publisher	Place of publication	ISBN/ISSN	DOI
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	Aug 2023	https://ieeexplore.ieee.org/document/10224826	https://iris.polito.it/retrieve/handle/11583/2982820/682640	IEEE	Venice, Italy	979 835 031 171 6	https://doi.org/10.1109/CSR5750.6.2023.10224826
Conference proceedings	Glider: Serverless Ephemeral Stateful Near-Data Computation	Barcelona-Pons, Daniel; García-López, Pedro and Metzler, Bernard	Middleware '23: 24th International Middleware Conference	Nov2023	https://dl.acm.org/doi/10.1145/3590140.3629119	https://bdrop.bs.ces/index.php/apps/onlyoffice/2518133?filePath=Extract%2F4.%20WPs%2FWP5%2FEXTRACT_Dissemination%20Register.xlsx	Association for Computing Machinery	New York, New York	979 840 070 177 1	https://doi.org/10.1145/3590140.3629119
Conference proceedings	Scaling a Variant Calling Genomics Pipeline with FaaS	Arjona, Aitor, Gabriel-Atienza Arnau, Lanuza-Orna Sara, Roca-Canal, Xavier	Middleware '23: 24th International Middleware Conference	Dec2023	https://dl.acm.org/doi/10.1145/3631295	https://arxiv.org/html/2312.07090v1	Association for Computing Machinery	New York, New York	979 -8-400 7-045 5-0	https://doi.org/10.1145/3631295.3631403
Journal	MLLess: Achieving cost efficiency in serverless machine learning training	Gimeno Sarroca, Pablo, Sánchez-Artigas, Marc	Journal of Parallel and Distributed Computing	Jan 2024	https://www.sciencedirect.com/science/article/pii/S074373152300134X?via%3Dihub	https://arxiv.org/abs/2206.05786	Elsevier Inc.	Amsterdam	074 3-731 5	https://doi.org/10.1016/j.jpdc.2023.104764

ihub#qsp0010						
Conference proceedings	Dataplug: Unlocking extreme data analytics with on-the-fly dynamic partitioning of unstructured data	Arjona Pérez, Aitor , Garcia Lopez, Pedro and Barcelon a-Pons, Daniel	24th IEEE/A CM International Symposium on Cluster, Cloud and Internet Computing	May2024	Accepted, pending publication May 2024, CCGRID 2024 (ccgrid-conference.org)	Pending publishing