ARCHIVER

ARCHIVING AND PRESERVATION FOR RESEARCH ENVIRONMENTS

Project Abstract: Phase 1

Contractor Rhea Consortium

Guidance

This project abstract is a summary to clearly explain the R&D to be carried out in Phase 1 and the objectives to be met. It shall include:

- short description of the company or consortium
- a description of the proposed solution and how it meets the R&D objectives identified in the Contract
- the R&D for Phase 1 (what specifically will be done)
- how the objectives will be achieved (technical and organizational measures to carry out the R&D and achieve the stated objectives)
- any steps that will be taken to prepare for commercialization of the solution after the PCP.

This project abstract will be communicated to the European Commission, and information that is not indicated as being confidential will be provided to interested parties such as early adopters. Any confidential information must be marked as such.

Consortium 4 participants

Participant Number	Participant organization name	Short name	Country
1 (Coordinator)	RHEA SYSTEM S.P.A.	RHEA	IT
2 (Member)	DEDAGROUP S.P.A.	DEDAGROUP	IT
3 (as subcontractor)	GTT SWITZERLAND SARL	GTT	CH

Document Information

Project Number	IT-4531/IT/PCP	**Acronym**	ARCHIVER Project
Full title	Archiving and preservation for research environments		
Project url	https://www.archiver-project.eu/		
Document url	Not Available		
Archiver Project Office	Marion Devouassoux (CERN)		

Deliverable	**Number**	D1.1	**Title**	Archiver Project Abstract
Work Package	**Number**	WP11000	**Title**	Coordination & Project Reporting

Date of delivery	**Contractual**	30/06/2020	**Actual**	30/06/2020
Status	**Version**	1.0	**Title**	D1.1 ARCHIVER Project Abstract
Type	○ X Report ○ Demonstrator ○ Prototype ○ Pilot ○ Other			

Dissemination level

PU	Public	
PP	Restricted to other programme participants (including the Commission Services)	
RE	Restricted to a group specified by the consortium (including the Commission Services)	
CO	Confidential, only for members of the consortium and Buyers Group (including the Commission Services)	X

Abstract

This project abstract is a summary to clearly explain the Research & Development to be carried out in Phase 1 and the objectives to be met.

Keywords

Hybrid Cloud, Science Cloud, Identity Management, Data Management, Data Archiving, Long Term Preservation, Science Heritage, OAIS, Service Catalogue, Business Model, Maturity Matrix, FAIR.

Statement of originality

Not Applicable

Authors List

Authors	Name	Contact information
RHEA SYSTEMS S.p.A.	Iolanda Maggio	i.maggio@rheagroup.com
DEDAGROUP S.p,A	Roberta Svanetti	roberta.svanetti@dedagroup.it

History of changes

Version number	Issue Date	Revised by	Reason of change
1.0	30/06/2020	Iolanda Maggio	First Issue

Table of contents

Abbreviations

ARCHIVER	Archiving and Preservation for Research Environment
AARC	Authentication and Authorisation for Research Collaborations
CSF	Critical Success Factor
EO	Earth Observation
EOSC	European Open Science Cloud
ESA	European Space Agency
OAIS	Open Archival Information System
OGC	Open Geospatial Consortium
PAYG	Pay-as-you-go
QoS	Quality of Service
R&D	Research and Development
SAML	Security Assertion Markup Language
SLA	Service Level Agreement
TPM	Third Party Mission
WBS	Work Breakdown Structure
WP	Work Package

Introduction

The project abstract objective is to summarize to clearly explain the R&D to be carried out in Phase-1 with the solution and the objectives to be met.

Consortium description

To address the Archiver objectives, the Consortium (called Consortium 4) brings together the following participants:

Table 1:

Participant Number	Participant organization name	Short name	Country
1 (Coordinator)	RHEA SYSTEM S.P.A.	RHEA	IT
2 (Member)	DEDAGROUP S.P.A.	DEDAGROUP	IT
3 (as subcontractor)	GTT SWITZERLAND SARL	GTT	CH

The Consortium is composed of RHEA as Lead Tenderer, DEDAGROUP as Consortium member, and GTT as RHEA's subcontractor. The figure below shows the hierarchy of the Consortium and the scope WPs assigned to each company.

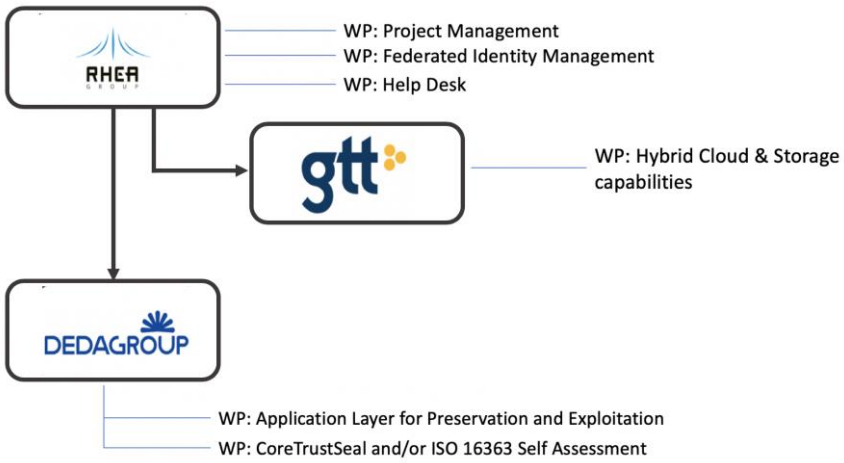

Figure 1 - Consortium 4 diagram

In the following table the key characteristics, skills and expertise of the components of the Consortium are described:

Table 1-1:

Short name	Key characteristics, skills and expertise
RHEA	RHEA is an experienced leader of large frame contracts in Space and Defence sectors, including procurements and integrations of IaaS and PaaS services for facilitating access to and exploiting of

	EO data in a cloud environment. RHEA has developed a solid relationship with cloud specialists and supports spin-in of its technologies for Space & Defence. RHEA will act as the Lead tenderer for ARCHIVER and will bring specialists and experts to support data management and data preservation tasks and to manage the Federated Identity Management solution. RHEA Group is a privately-owned professional engineering and solutions company, providing bespoke engineering solutions, system development and security services for space, military, government and other critical national infrastructure organizations. Since its creation in 1992, RHEA has built a reputation as a trusted partner, developing tailored solutions that help drive organizational and cultural initiatives, leading to sustainable added value for its customers. Headquartered in Belgium, RHEA Group employees over 500 people and has offices in Belgium, UK, Czech Republic, Italy, France, Germany, Spain, Switzerland, the Netherlands and Canada, and works at clients' premises throughout Europe and North America. RHEA is ISO 9001 and ISO 27001 certified. In the context of ESA Heritage Space Programme, RHEA covers the stewardship (i.e. preservation, discovery, access and exploitation) of its space science heritage data, theoretically for an unlimited time. These activities address not only ESA EO data holdings but also ESA data acquired from space in other scientific domains. The EO Data Preservation System has the main objective of providing the required infrastructure and services to assure ESA and Third Party Missions (TPM) EO Data Records and Associated Knowledge preservation and accessibility, and to support the cooperation activities with national and international organizations in the data preservation domain. The implemented and supported services include dataset and information preservation, Persistent Identifier assignment, Heritage Software exploitation, Metadata management (OGC standard, PREMIS and Dublin Core), Provenance and context management and a Data Management and Stewardship Maturity Matrix self-assessment. Best practices and guidelines on preservation are produced and maintained up-to-date. The infrastructure for Data and Information long term preservation is based on the open source Archivematica and AtoM tools also proposed in the ARCHIVER project.
DEDAGROUP	DEDAGROUP has many years of experience and strong skills in archiving and long-term data stewardship processes gained with a multidisciplinary team in many industries (energy, oil & gas, engineering & constructions, earth observation, digital humanities, life science, manufacturing, public services, financial services). DEDAGROUP is able to apply various best practices in a sustainable and scalable service architecture, considering the peculiarities of each sector. DEDAGROUP has an organizational structure, namely "Digital Knowledge Life", dedicated to the provision of data archiving and long-term preservation services for any type of digital objects on premise and cloud modalities, including IaaS and PaaS modes from certified data centres. With revenues of €247 million in 2018, a staff of over 1,700 and more than 3,600 customers, Dedagroup is a major aggregator of Italian excellence in software and solutions as a service and a natural partner to companies, financial institutions and public services in developing their IT and digital strategies. Founded in 2000 and based in Turin, the Group has enjoyed constant growth. In addition to its over 20 offices in Italy, it also operates in Switzerland, France, Germany, the UK, the USA, Mexico and China. Deda.Cloud is the cloud managed service provider to companies and organisations that use innovative technologies to develop products and services and constantly improve their processes. A division of Dedagroup S.p.A., it specialises in cloud strategy and is organised to work in synergy with the Group's other companies and business units: Dedagroup Business Solutions, Dedagroup Public Services, Dedagroup Stealth, Dedagroup Wiz, Derga Consulting and Piteco. For over 20 years DEDAGROUP has been the engine of innovation at the service of several customers, putting their experience in process management and technological competence, skills developed also thanks to the close collaboration with the other specialized business units within the group and with the best technological partners worldwide. DEDAGROUP's mission is to transform data into actions, enabling digital business through the design of the technological and information bases that complete the services and application solutions, generating high-performance, safe and oriented solutions to customers' business results. For more than 10 years DEDAGROUP has been providing cloud services aimed in particular at the adoption of clouds and hybrid solutions, Business-continuity and data protection, Cyber and IT security, Managed Services and Integration, management and licensing of business services. All these services comply with the standards required by the ISO27001 certification and associated guidelines. DEDAGROUP has been implementing data archiving and preservation solutions for over 15 years, which includes several successful installations in government, in cultural heritage, in scientific domains and other regulated industries. DEDAGROUP has many years of experience providing consulting and IT services in the data archiving domain and strong skills in archiving and long-term data stewardship processes gained with a multidisciplinary team with concrete experiences in many

	industries (energy, oil & gas, engineering & constructions, earth observation, libraries, historical archives, digital humanities, life science, manufacturing, public services, financial services, ...). They have the capability to apply various best practices in a sustainable and scalable XaaS architecture, considering the peculiarities of each sector in terms of business processes needs, challenges, compliance constraints and innovation opportunities. DEDAGROUP has an organizational structure, namely "Digital Knowledge Life", dedicated to the provision of data archiving and long-term preservation services for any type of digital objects (like data, raw data, video, audio, images, databases, documents, virtual machines, databases, ...) in both on-premise and cloud deployments, including IaaS and PaaS modes from certified data centre. Open source software is privileged, and enterprise platforms are integrated depending on customer guidelines and IT strategies.
GTT	GTT connects people across organizations, around the world and to every application in the cloud. Clients benefit from an outstanding service experience built on core values of simplicity, speed and agility. GTT operates a global Tier 1 internet network and provides a comprehensive suite of cloud networking services. From financial services trading firms to manufacturers and government, GTT is committed to providing our clients with the services, reach and capabilities that improve communication, productivity and efficiency across their organisations. GTT connects people across organizations, around the world and to every application in the cloud. Our clients benefit from an outstanding service experience built on our core values of simplicity, speed and agility. GTT owns and operates a global Tier 1 internet network and provides a comprehensive suite of cloud networking services. GTT offers wide area networking, internet, managed services, transport & infrastructure, and voice, all designed to meet our clients' unique needs. Take advantage of GTT's software-defined wide area networking, a managed service with the broadest range of access options, to gain visibility and control across your WAN. We deliver services in over 100 countries across six continents, ensuring that we are everywhere you do business. GTT has made available in his own Amsterdam data centres a dedicated Storage Infrastructure up to 8PB for Multi-Mission Satellite Data for ESA. This solution is based on disk with different tiering. Depending on the customer requirements and legal compliance, GTT is used to deploy different storage technologies for his customers, typically high-end storage for databases and low-end storage for backup or cold data. Typical deployment is done with Netapp cluster (two controllers) and shelf disks to serve local compute deployed within the same data centre. This type of deployment is standard and is currently in operations for dozens of customers. Customers who have other requirements like disaster recovery, second copy of their primary site, offsite backup can use the standard GTT offering to replicate their data to another GTT data centre across Europe. The replication is done using different technologies. At storage level the replication is done via the Netapp controllers using the Netapp snapmirror utilities. Snapmirror will replicate at block level all the changes from primary to secondary site. Snapmirror is mainly used for Disaster recovery purpose. GTT is also providing backup using the storage capability using the Netapp Snapvault utilities. Snapvault is mainly used for backup and is using the Snapmirror capability to generate the backup and version history based on snapshots. GTT also provides snapshot functionality to allow end user to benefit the previous version and quick retrieval of file without need to engage IT department. All these utilities can be customised to meet customer business requirements. GTT also deploys tape storage infrastructures for long term preservation. Tape storage is mainly used to backup and archive data that need long retention and based on Quantum or HP. Initial copies are done on disks and for customer requiring longer retention, the data is replicate into tapes. Other customer requirements mandate for these tapes to be also located in a second offsite data centre to avoid having all data in the same data centre for disaster recovery purposes or to comply with their internal or external audit requirements. The offsite copy is either done using the GTT backbone and duplicate the data into the tape library located in another GTT datacentre or using GTT partners that only do tapes handling and have facilities made specially to handle the tapes. GTT allows customised policies and setup to allow customer to meet their business requirements.

Excellence of the proposed solution

This paragraph describes the proposed solution and how it meets the R&D objectives identified in the Contract.

Ambition

The Consortium ambition is to contribute to the R&D activities for designing, implementation and deploying of a complete and secure service model to fully cover the stewardship lifecycle of multiple and heterogeneous research data exceeding the occurring lacks in these concepts, involving organisation, processes and technologies. The Consortium will consolidate the achieved outcomes creating relevant Best Practices.

The Consortium has long and strong experience and competence in Data archiving and Long Term Preservation, specifically in:

- Space (Science and Earth Observation);
- Life Science (Pharmaceutical, Biotech, Clinical Research Organization);
- Libraries and Historical Archives.

The Consortium expertise is focused on:

- Scientific and Space data;
- Clinical data and Records;
- Confidential, Reserved and Unstructured data;
- Historical and Unique data.

The solution is therefore composed of open source applications and a dedicated hybrid cloud in order to grant the whole integrated and fully protected data stewardship and curation lifecycle. The solution assures full adherence to the EOSC guidelines and specifications and readiness for any kind of future Early Adopters.

Solution description

The proposed solution is in line with the vision and scientific scope of the ARCHIVER project and addresses the challenges set forth by the ARCHIVER project applying a methodology gained through extensive and proven experiences in R&D processes, long-term secure data archiving and preservation, cloud infrastructures and services management in several industries and European agencies with heterogeneous datasets, needs and requirements.
The Consortium (RHEA, DEDAGROUP and GTT) is fully confident of its complementary key competences and experiences allowing to develop innovative and robust services covering all ARCHIVER requirements for Organizational Change Management and Governance, Service Business and Cost Model, Service Processes, Technical Solutions Design and Implementation. The solution addresses business requirements raised by the Buyer Group under strong regulations (Energy and Life Sciences domain) and supports testing and validations of IT infrastructures, applications, services and change control processes.

Figure 2: Solution Layers

The high level architecture components of the solution (Figure 2) are the following:

- Secure Service Portal (Identity Access Management, Access Layer Interface, Validation and Pre-ingestion services), implemented by RHEA;
- Existing and mature Open Source platforms for data archiving, preservation, reporting and access/discovery (Archivematica/AToM/JasperSoft), integrated and configured by DEDAGROUP;
- Readiness XaaS services, provided by all partners;
- Cloud connect product for integration with proposed robust and scalable managed Hybrid Cloud, provided by GTT.

The complete and detailed solution design and architecture will be provided as deliverable of Phase-1 as requested.

The proposed foundation will accelerate the R&D prototyping phases in order to incrementally provide the Buyers with a robust and scalable service model covering all ARCHIVER's objectives and Use Cases.

The Service Portal recognizes and processes user requests for authentication, and orchestrates the required services that are dispatched or exposed according to the Services Catalogue. The proposed solution includes and integrates unified authentication services based on SAML 2.0.

The solution will be implemented in compliance with applicable guidelines and standards for data archiving and preservation. The implemented architecture will be based on open standards and robust and scalable technologies (the baseline), enabling a secure and efficient interaction between data producers, service providers and service consumers.

The solution is integrated with GTT Cloud services and infrastructure. GTT manages several Data Centres in different EU countries able to host the infrastructure. One of them is located in Geneva and already connected via direct fibre to CERN. The employed storage solution features a so-called "Active Archive" strategy offering a hybrid storage architecture employing different types of tier storage including cloud, high performance flash storage for the front-end and tape libraries for long-term data archiving: overall, this providing high performances and much lower running costs than only local magnetic-disk-based architectures.

Service Innovation

The proposed service includes the following (not exhaustive) list of innovations:

- Predefined use cases for specific target markets to accelerate services setup and mapping with specific regulations of service quality management, data integrity management, long-term preservation and dissemination to achieve appropriate configurations;
- Petabyte scale data management solution, portable to any cloud, integrating private and community clouds;
- Unified dashboard to better manage hybrid resources, including private or community clouds;
- Unified authentication service integrated.

Alignment with R&D objectives

The paragraph contains a description of the proposed solution and how it meets the R&D objectives identified in the Contract.

To follow, the pillars of the solution that meet the ARCHIVER project R&D objectives:

1. The Consortium with the aforesaid solution will provide R&D services that address the data archiving and long-term preservation needs of multiple scientific domains, implemented under the OAIS reference model and related standards supporting FAIR data principles. The Consortium is designed to maximise the opportunity of innovation and ARCHIVER Project success and minimise the risk of failure in the design and the development and execution of the future phases implementation. The challenges and innovations of ARCHIVER Project requires a well-selected team with expertise and specific and multidisciplinary roles.

2. The solution will extend the already existing OAIS conformant platforms demonstrating a high scalability (at the petabyte scale) of scientific data from multiple research domains, with very high data ingestion rates (10Gbps/day), via commons solutions.
3. The proposed federated identity management services will support authentication and authorization to the archives and, the exploitation of a hybrid deployment model. The proposed implementation follows the AARC Blueprint Architecture.
4. The implemented services will be compliant with the current regulations and legislations (e.g. GDPR) and will be flexible for the future regulations and guidelines (e.g. UE Data Act 2020-2021). The compliance includes the development of transparent and auditable business models considering the following key indicators: archives lifetime, maximum ingestion rates, data volumes, number of copies, data portability and exit strategies.
5. The Consortium will develop a commercialisation plan in line with Early Adopter Programme.
6. The governance layer aims at developing, monitoring and continuous improving the organizational change management initiatives (Education, testing, training, communication) in order to involve, engage and sustain the ARCHIVER Buyers Group organisations and Early adopters.
7. The union of governance layer and Data Archiving and preservation services developed in the ARCHIVER project will enable the creation of trustworthy repositories where researchers can store their data for sharing, a key building block of the research process.
8. The Consortium leveraging on their previous experience will be ready for the inclusion of ARCHIVER services in the context of European Open Science Cloud (EOSC). The Phase-1 Design Phase will early also consider the specification for this future compatibility.

In Phase-1 Design Solution the R&D implementations will be:

- Data archiving and long-term preservation for multiple scientific domains;
- OAIS Functional Model compliance;
- Buyers accessibility to the DEMO platform;
- First draft of initial self-assessment (ISO 16363);
- Initial report on the Total Cost of Services (comprehensive study of all quantitative and qualitative factors significantly impacting the cost that a customer bears when purchasing and running the services resulting from ARCHIVER.

Challenges and objectives mapping

The table below contains the mapping between the solution objectives and the ARCHIVER challenges which will consist of the baseline of the Design Results evaluation.

Table 2:

Contractual Archiver objectives/challenges	**Our solutions objectives**

1. Innovative Archiving & Preservation Services	The Solution includes existing and mature Open Source platforms for data archiving, preservation, reporting and access/discovery (Archivematica/AToM/JasperSoft). The innovation of the solution is represented by predefined use cases for specific target markets to accelerate services setup and mapping with specific regulations of service quality management, data integrity management, long-term preservation and dissemination to achieve appropriate configurations. A further innovation, to create innovative and complex data harvesting between application services and IT services platforms, is the opportunity to introduce of an enterprise service bus (or hub) technology that implements a communication system between mutually interacting software applications in a service-oriented architecture. The hub promotes agility and flexibility with regard to high-level protocol communication between applications and is a core element in the XaaS frame. The primary goal of the high-level protocol communication is enterprise application integration (EAI) of heterogeneous and complex service or application landscapes. The hub maps all the backend exposed endpoints in Archivematica and makes them accessible in an orchestrated modality to applications that requires services as APIs. It allows to manage the APIs ecosystem including for example OpenAPI Specifications. Normally, the hub manages also logging of all the messaging coming from the various applications APIs call. The logs are used with reference to control the flow in its execution and to manage service performances measurements.
2. OAIS, CoreTrustSeal & FAIR	The Consortium has full confidence in the technical solution due to our extensive and proven experience in services management, application, technical architectures and benchmark models' development and management, in addition to usage of Consortium's methodologies and compliance with ISO 9001:2015 Quality Management System. The final implemented solution is aligned with guidelines and applicable standards for data archiving and preservation like OAIS-ISO 14721:2003, ISO16363, PREMIS, ISAD (G), ISAAR (CPF), EAD, METS. The proposed solution maps the OAIS Functional Model, relevant ISO16363 metrics and technical requirements. In particular a Use Cases Toolkit and a dedicated set of microservices already exist for all six OAIS functional entities: 1. Preservation Planning; 2. manual/automatic Ingest (including functional entities for SIP transfer, backlog, appraisal and re-ingest); 3. Data Management; 4. Archival Storage (to manage multiple pipelines and storage management infrastructure integration), 5. Access (DIP dissemination for user usage); 6. Administration for monitoring, configuration and reporting. The Consortium contributes to the FAIR Data Maturity Model Working Group, which aims at creating guidelines, maturity levels and prioritisation to assist the implementation of indicators. They are based on the FAIR principles and aim to formulate measurable aspects of each principle that can be used by evaluation approaches. In the frame of the ARCHIVER Project, the Consortium will provide competence in order to give evidence of the FAIR indicators mapping and the score of the components of the solution.
3. Network & Federated Identity (IAM)	The solution will be based on the AARC Blueprint Architecture guidelines, aiming at providing "in-house" solutions reusing and adapting already available components and taking advantage of internal know-how, or alternatively using external available services/infrastructures (e.g. eduTEAMS). Federated access service will be implemented incrementally following the Buyers Group needs (e.g. machine-to-machine approach or specific protocols for FIM).
4. Business Models	RHEA, DEDAGROUP and GTT are working towards a clear and multi-level business model. The intention of our Consortium is to deliver large-scale XaaS services based on architectures in which open source solutions will be preferred, aiming to cover the variety of Buyer's targets without increasing the final cost of resulting services and minimising the efforts and complexity of the whole project. Our vision is to create a tiered offering, considering business models with varying

	levels of services and corresponding costs (e.g. Free self-service backed by a shared revenue model with Buyers, and/or a subscription based model for premium services, including different SLA and bouquet of managed services etc.). The business model conceived by the Consortium will include in ingress and egress traffic cost to control unforeseen increase of data ingestion and retrieval without increasing the total cost of service. Exit or migration plan will consist in transferring all data being managed to another ARCHIVER solution following the same cost rate in term of traffic. The cost of the receiver will allow massive data ingestion without additional costs. The business model must consider the maturity and ability of the Buyers to purchase the proposed ARCHIVER services and leveraging Early Adopters' subscriptions for extending of "as a Service" culture to lead the evolution in several public organisations' procurement practices. Industry must also adapt to be able to engage in this evolution.

Impacts

The paragraph provides a brief overview of the impacts/benefits that the Consortium aims at obtaining participating in the ARCHIVER PCP procurement.

The ARCHIVER R&D activities will impact in the European Research harmonizing and creating a common best practice for Data Stewardship and Long Term Preservation. The best practice will include the development of new methodology simplifying the future Early Adopter engagement. The project will promote the involvement of the Buyers and future Early Adopters using innovative tools for engagement and training in order to extend as much as possible the usage of the ARCHIVER services.

This PCP has given to the Consortium the opportunity to access and to have visibility to the Buyers institutions and their needs. The fact that we have been working with these institutes will raise awareness more widely. The nature of the PCP is a closed set of buyers, however, in Phase-3, it will be opened-up to early adopters who would also like to try out the innovative ARCHIVER services.

EOSC integration will contribute to enforce the European Infrastructure for Research.

Expected risks

The Consortium has analysed the risks for the Phase-1 Design Solution and the following not-exhaustive list of risks has been highlighted:

- Unavailability of staff;
- Incident affecting multiple users;
- Simultaneous unavailability of the Buyers Group;
- Failure in engaging of traditional researchers;
- Future Archiving initiatives (coming after Phase-3 deployment) not in line with European Regulations;
- Anticipated closure of ARCHIVER contract (Data Repatriation task needed).

Technical and organizational measures

The paragraph describes how the objectives will be achieved (technical and organizational measures to carry out the R&D and achieve the stated objectives).

Each Consortium member is expert in its own domain, with numerous acknowledgments from its customers. All partners are involved as suppliers for the European Space Agency, bringing a demonstrated capability to collaborate.

The ARCHIVER project will be organised in a simple and efficient structure, with clear roles and responsibilities for each Work Package (WP) and a distinct structure for management, including finance and quality. Each partner will dedicate a key person responsible for technical and managerial aspects that will be able to guarantee that all the wealth of knowledge of each company is leveraged to its maximum. This will ensure a single point of contact for each partner, enabling him/her to best coordinate the efforts and ensuring the right staff are involved at the right time, with the best skills and knowledge. In order to guarantee continuity, each key person will also have a deputy.

To perform the design work in a relatively short period the team will deploy senior engineers to ensure the that best possible results are delivered, fulfilling the needs of the Buyers Group with a solution that can be implemented and delivered with minimum risks during Phases 2 and 3. All team members are experienced in iterative and incremental methodologies, ensuring that the proposed design is properly tested and successfully validated.

Based on a strong relationship and a complete collaborative approach between the members, the Consortium has developed and agreed the Work Breakdown Structures (WBS) for all three phases, considering the efforts and the resources to be put in place for the full development of the final solution.

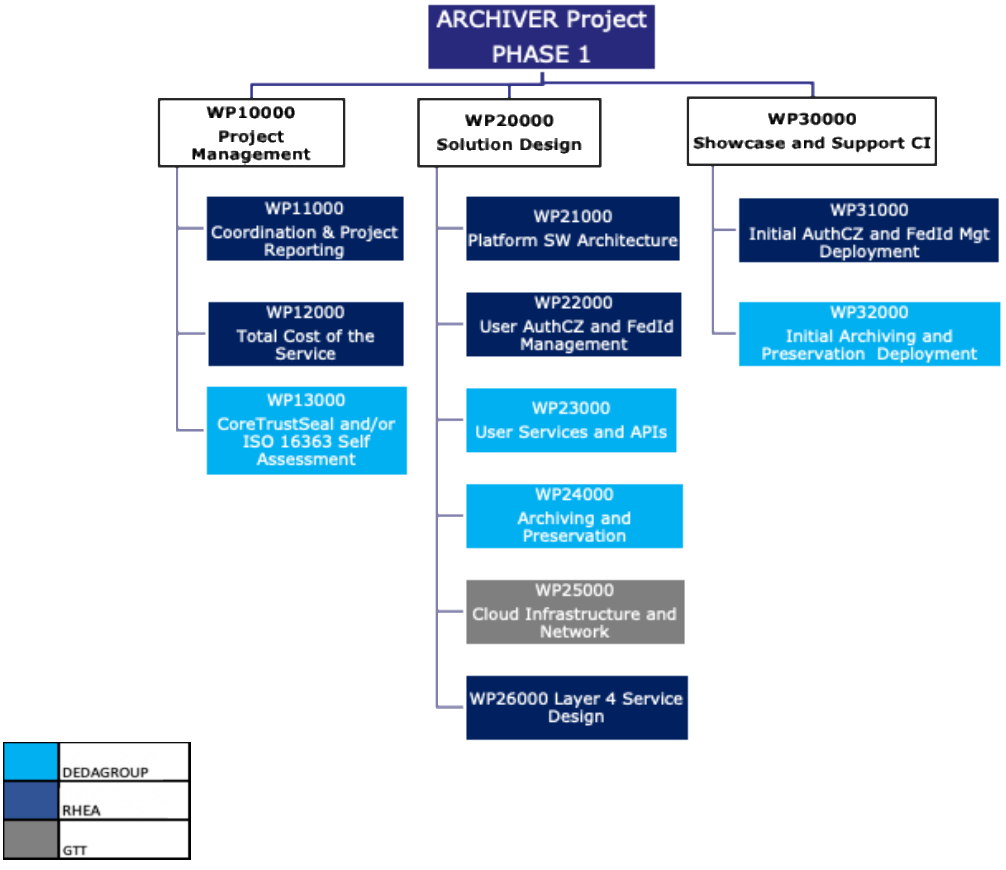

Figure 3 - Phase-1 WBS

The *ARCHIVER Master Work Plan* (Baseline 1.0) has been created to verify that all milestones, deliverables, activities, tasks, resources and responsibilities are managed properly.

The Buyers use cases have been analysed and UML designed; these have been used as input for the WBSs and the *ARCHIVER Master Work Plan*.

In order to better orchestrate all communications and implementations, a specific layer of governance has been considered in the project organisation. Lines of communications have been planned to ensure sharing of knowledge and results between the Consortium's members and the Buyers. Weekly online and face-to-face internal meetings are taking place to ensure tight coordination, with as much opportunity for the Buyers Group to get involved and to provide feedbacks, if and when needed.

			Phase 1	The overall aim is: • elaborate the solution design architecture and determine the approach to be taken to develop the new solutions; • demonstrate the technical, financial and commercial feasibility of the proposed concepts and approach to meet the procurement needs.	The outputs of the Phase are: • a written detailed solution design, a system architecture, and a technical design of all the components; • the deliverables as indicated below												
										Rhea								De
WBS	WBS Title	Activity	Description	Workproduct	CERN Notes on 09 June 2020	Deliverables	Milestone	Due Date	System Engineer	Engineer	Junior Engineer	IT Specialist	Other	System Engineer	Engineer	Jun Eng
WP11000	Coordination & Project Reporting	A1.1	Contract kick-off meeting				M1.1	08-Jun-20	2					2		
WP11000	Coordination & Project Reporting	A1.2	DEL: Finalization of Project abstract (in the format required by the EU for publication)	- description of the R&D efforts so far in the Phase - details of the progress against the Contractor original schedule - description of the R&D to be carried out in the remainder of the Phase and a corresponding schedule - any other relevant technical information	Based on template: Your innovative solution Steps towards commercialisation	D1.1 AUTHORS: IM/RS REVIEWERS: IM APPROVER: IM		30-Jun-20	20					4		
WP11000	Coordination & Project Reporting	A1.3	DEL:Finalization of Declaration of Contractor's Background IP	- description of the R&D efforts so far in the Phase; - details of the progress against the Contractor's original schedule; - description of the R&D to be carried out in the remainder of the Phase and a corresponding schedule; - any other relevant technical information.	Based on template	D1.2 AUTHORS: IM/RS REVIEWERS: IM APPROVER: IM		30-Jun-20	20					4		
WP320000	Initial Archiving and Preservation Deployment	A1.4	Demo Platform	prepare the test and think about the name of the platform: - OAIS Reference Model compliance - Format obsolescence Observatory (PRONOM) -Buyers access via Internet -Ingestion Functionality -Basic Discovery functionality -Verification Storage capability (less than 100 TB).	ARCHIVING workflow: Every 24 hours, at a random time, the BG will PUT 500 files of 2GB to all Contractors, interleaving the Contractors on a file by file basis. If the Contractor supports the functionality, the BG will only send these binary files, providing documentation of byte offsets and type information as to where metadata is embedded. In some cases, the BG expects the Contractor to extract the metadata and archive the file (not applicable to all use cases). As the Contractor might not support the functionality, in all cases it shall be possible to provide an interface by which one sends the metadata in addition to each file. RECALL workflow: Once a week, at a random time, the BG will launch a metadata query which will specify a 10% subset of all the archived data. This query should return a	Workproduct	M1.2	14-Jul-20							

Figure 4 - ARCHIVER Master Work Plan

Moreover, the methodology proposed for internal project management is in line with PMI's PMBOK and the methodology for service management will follow the ITIL Foundation V4.0.

Project Result and Quality Measurements

The paragraph describes the acceptance criteria for the deliverables and milestones produced by the Consortium.

Critical success factors measurement

The ARCHIVER project aims at being compliant to ISO 16363 and the quality of the produced documentation (Deliverables/Work Product) can have impacts on any relevant future auditing activities. For these reasons and according to ISO 9001 criteria, in the ARCHIVER Phase-1 Project scope, have been defined the following monitoring areas, identified as Critical Success Factors (CSF):

- **Planning**: related to project timeline, project status actual vs programmed baseline and impacts on activity accomplishment;

- **Risk Management**: related to open items, risks and emerging critical items, mitigation activities and solving capabilities;
- **Quality**: related to quality levels of produced output;
- **Consortium Team Management & Staffing**: related to the appropriate fits of the project staffing requirements, in terms of number of resources and in terms of skills;
- **Budget**: related to coverage of the budget, planned for the activity.

The indicators related to the above CSFs are the following:

- **Planning**
  - Effective variance on project baseline
  - % approved deliverables on deliverables to be approved
  - % submitted deliverable on the planned deliverables
  - % activities with delay on critical path
- **Risk Management**
  - Risks – Occurrence average probability (occurrence probability and impacts)
  - % risks without mitigation plan
- **Quality**
  - % produced deliverables acceptance
- **Budget**
  - Accounting Variance (ACWP-BCWS).

The collected data are included in a record for internal usage of the Consortium Project Team and it will be used to produce a summary to be included in the periodic status report.

Quality Metrics

In the following table are described objectives, metrics and target values that are applied to measure the ARCHIVER Project quality.

Objectives	Metrics	Target Value	Input elements
Document/deliverable accepted	Num. of rework cycles on the documents internally reviewed and submitted to the customer	Max 3 rework cycles	Master ARCHIVER Workplan
Adherence to standard and methodologies	Num. of internal audit, Num. of observed deviations	0 Major deviations for single internal audit	Internal Audit report
Issue solving time	% of issues solved in the estimated time frame	90%	Rhea Weekly meeting with Buyers

Objectives	Metrics	Target Value	Input elements
Achievements of planned deadlines	% delays due to project documentation reworking	80%	Rhea Weekly meeting with Buyers
Buyer satisfaction	Internal Buyer survey Num. of Buyer critical complaints	0 Critical complaints	Buyer satisfaction report, Rhea Weekly meeting with Buyers

SLA for ARCHIVER Platform Demonstrator

In line with ITIL standards the Service Level Management needs a framework that disciplines provider and customer for which the SLA are the result of an interaction between the stakeholders of the project. Furthermore, the EU Cloud Code of Conduct remarks "Cloud Service Level Agreements (Cloud SLAs) form an important component of the contractual relationship between a cloud service customer and a cloud service provider of a cloud service. Given the global nature of the cloud, SLAs usually span many jurisdictions, with often varying applicable legal requirements, in particular with respect to the protection of the personal data hosted in the cloud service. Furthermore, different cloud services and deployment models will require different approaches to SLAs, adding to the complexity of SLAs".
The Consortium is following the evolution of EU data strategy (Feb. 2020) and the relevant impact in applicable regulations, Cloud services Provider registry and new guidelines for data archiving, protection and accessibility.

Total Quality of Service

In line with the ITIL guidelines, QoS for archiving and preservation service management will be ensured starting from our resource management policy (baseline), part of the Service Management processes and services.

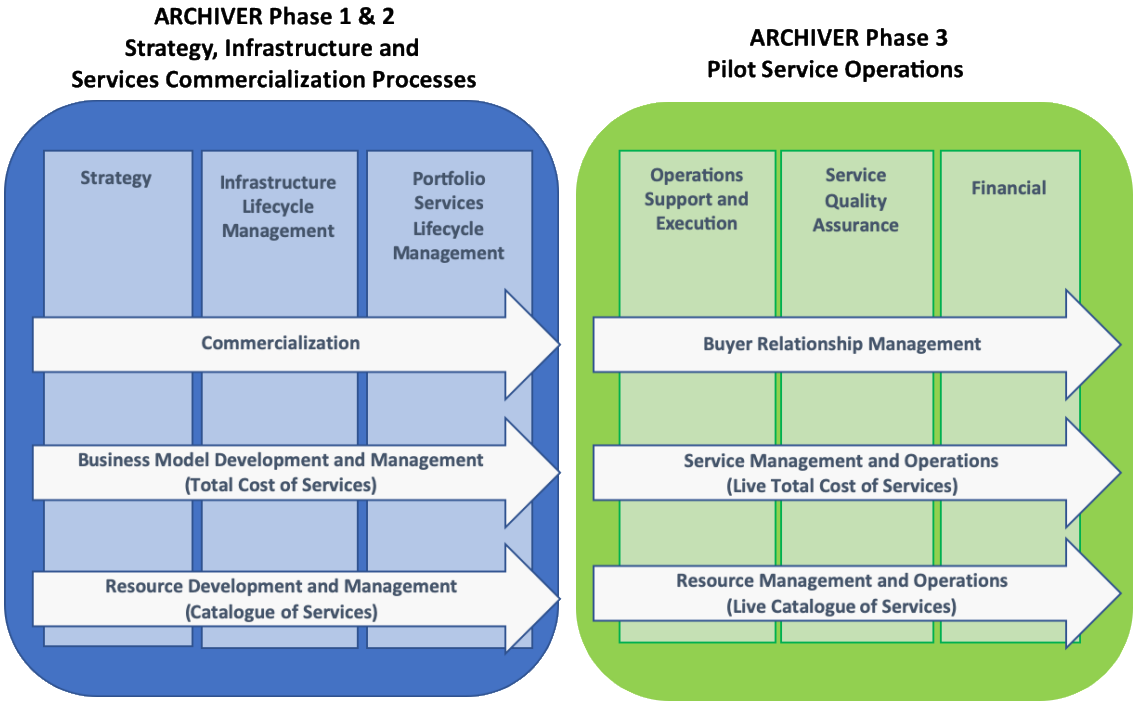

Figure 5 - Total Quality of Service approach

Commercialisation Plan

This paragraph contains the steps that will be taken to prepare for commercialization of the solution after the PCP.

Results of R&D activities will enhance and innovate the foundation to bring a viable service on the Early Adopters' market. Commercial approach and service portfolio will be extended to support the exploitation of the results and the commercialization of the existing Managed Cloud service. The Commercialisation formula might include:

- Pay-as-you go
- Bulk subscription
- Optional managed services (data management, based on performance, data volume and transfer, as well as container management, based on number of underlying cloud resources).

To fulfil the needs of the ARCHIVER's target markets subject to different regulatory constraints (Research centres, and other potential clients), the commercialisation approach will provide the possibility of configuring different service combinations in a flexible and attractive way. This will include the continuous engagement of target markets to collect and enrich new requirements, to contribute on the tests of new services under development and to promote usage of new services in disciplines with ad-hoc ARCHIVER's communication initiatives.

A key factor for the sustainability of new service architectures and the related Commercialisation approach will be the continuous involvement of open source communities in maintaining and further expanding and promoting underlying solutions.

During Phases 2 and 3, the Commercialisation approach will also include the estimation of the time-to-market, the market trends, as well as results from other projects, investments and developments related to ARCHIVER.

The Cloud service is provided following specific rate card depending on the workload deployed providing PAYG model and committing cost model that could be reduced.

References

Not applicable