

ARCHIVER - Data archiving and preservation for research environments

Marion Devouassoux¹, João Fernandes², Bob Jones³, Ignacio Peluaga Lozada⁴, and Jakub Urban⁵

^{1,2,3,4,5}CERN, Information Technology Department, Geneva, Switzerland

Abstract. Over the last decades, several data preservation efforts have been undertaken by the HEP community, as experiments are not repeatable and consequently their data considered unique. ARCHIVER is a European Commission (EC) co-funded Horizon 2020 pre-commercial procurement project procuring R&D combining multiple ICT technologies including data-intensive scalability, network, service interoperability and business models, in a hybrid cloud environment. The results will provide the European Open Science Cloud (EOSC) with archival and preservation services covering the full research lifecycle. The services are co-designed in partnership with four research organisations (CERN, DESY, EMBL-EBI and PIC/IFAE) deploying use cases from Astrophysics, HEP, Life Sciences and Photon-Neutron Sciences creating an innovation ecosystem for specialist data archiving and preservation companies willing to introduce new services capable of supporting the expanding needs of research. The HEP use cases being deployed include the CERN Opendata portal, preserving a second copy of the completed BaBar experiment and the CERN Digital Memory digitising CERN's multimedia archive of the 20th century. In parallel, ARCHIVER has established an Early Adopter programme whereby additional use cases can be incorporated at each of the project phases thereby expanding services to multiple research domains and countries.

1 Introduction

Data has both a value and an associated cost. Research data management makes many promises in terms of capacity, scalability, ease-of-use and security. The stewardship of research data involves not only all data-related tasks during the active lifetime of a project itself but also, packaging the data, keeping its intellectual property safe and associated information for later re-use once the project has finished. The period during which research data remains valuable can stretch into decades.

Data Preservation has a long record in the High Energy Physics community, driven by the Data Preservation for High Energy Physics group (DPHEP)[1] over the last 20 years. More recently, CERN has announced a new open data policy for scientific experiments at the Large Hadron Collider (LHC) that will make scientific research more reproducible, accessible, and collaborative[2].

In spite of these efforts, there is a critical gap in offering services that are standards-based, cost-effective for long-term archiving and preservation[3]. This is particularly evident when in presence of high sustained data ingest rates (10 Gbps) and research data volume of multiple petabytes.

In this context, using the EC Pre-Commercial Procurement (PCP)[4] instrument under Horizon 2020[5], with a budget of 3.4M euros for R&D competitive tendering, the ARCHIVER project's goal is to introduce significant improvements to fulfil these data management promises in a multi-disciplinary environment, allowing each research group to retain ownership of their data whilst leveraging best practices, standards and economies of scale.

Acting as a collective of procurers identified as the Buyers Group, CERN[6], EMBL-EBI[7], DESY[8] and PIC[9] commit funds, research datasets and testing effort to create an innovation ecosystem for specialist ICT companies active in archiving and digital preservation willing to co-develop end-to-end archival and preservation services for data generated in the context of scientific research supporting the IT requirements of European scientists.

2 ARCHIVER approach

As an innovation procurement project, ARCHIVER started with a preparation phase to identify the current gaps in preservation services offered to the public sector, assimilate lessons learned from past initiatives and consolidate future opportunities for the ARCHIVER resulting services.

In order to stimulate an open dialogue about the intended R&D scope, ARCHIVER organised an Open Market Consultation (OMC) with a series of activities for both demand and supply sides. The dialogue within the community, end-users and potential tenderers of the project took place between April and June 2019 as announced in a Prior Information Notice[10] published in the Official Journal of the European Union. These activities pursued a three-fold objective:

- Inform the research community, supply-side and the demand-side of the project and the upcoming call for R&D tender;
- Assess the requirements for digital preservation and archiving services of the research community at large and specifically in the context of the EOSC;
- Improve the mutual understanding of the R&D challenges across procurer organisations, future early adopters organisations and industry in order to verify the innovation potential and feasibility of the project.

The preparation phase also resulted in the publication of an analysis of the current state-of-the-art establishing the technological baseline and assessing the current solutions available in the market[11].

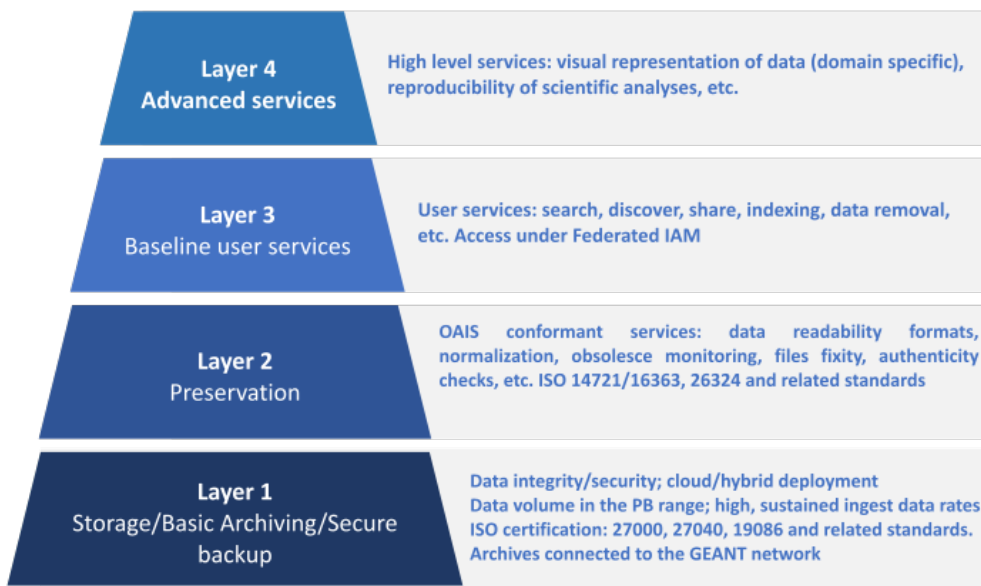


Fig. 1. R&D challenge high level architecture resulting from the ARCHIVER preparation phase.

These findings depicted in Figure 1 are reflected in the R&D tender launched on the 29th of January 2020, specifically in the selection and award criteria of the functional specifications, in order to ensure that the selected bids were capable of meeting ARCHIVER's R&D challenge.

2.1 R&D execution

ARCHIVER is following an implementation on three phases with multiple competing consortia:

- **Phase 1 - Solution Design:** Contractors develop designs including architecture and technical components. The activity during this phase has produced the results to be taken into account in the selection process that allows a Contractor to proceed to the subsequent project phase.
- **Phase 2 - Prototype Development:** Selected Contractors from the Design Phase are building prototypes based on the designed solutions and making them available to the buyers group for testing purposes.
- **Phase 3 - Pilot Deployment:** Selected Contractors will deploy expanded prototype services. These services will potentially be exposed to end-users and early adopters[12], in order to determine if they are suitable for their needs.

To prepare the R&D execution, the Buyers Group have provided a set of use cases in the form of deployments scenarios[13] drawn from their scientific activities, covering multiple scientific domains.

These base technical requirements are derived from four ESRIs landmark research infrastructures[14], namely the Cherenkov Telescope Array (CTA), ELIXIR[15], HL-LHC[16], and European X-Ray Free-Electron Laser Facility (XFEL)[17].

The use cases have demanding archiving and data preservation requirements in combination with significant volumes of data (several tens of PBs, in some scenarios doubling each year) sustained high data ingest rates (ranging from 1-10 GB/s) and long retention periods (up to 25 years).

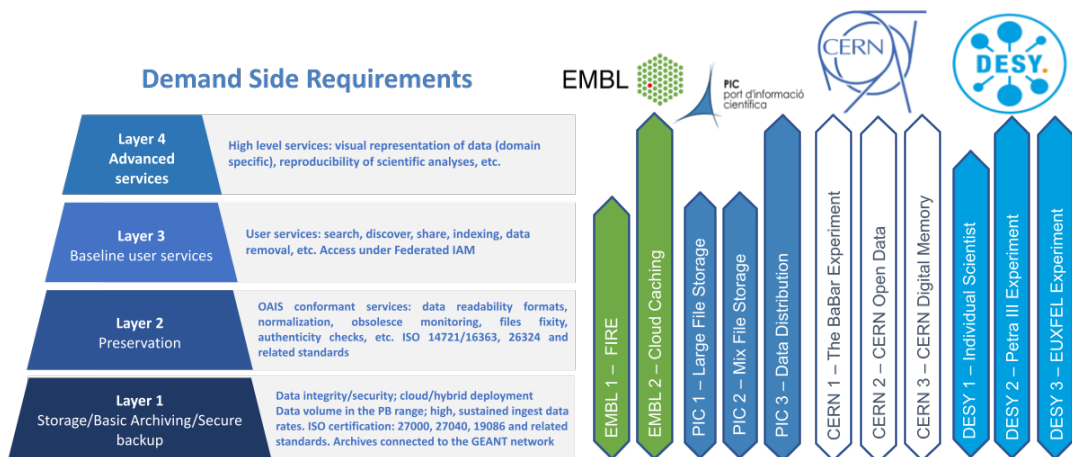


Fig. 2. Mapping of the R&D challenge layers to the initial set of deployment scenarios.

The requirements that encapsulate the R&D challenge (Figure 2) for the resulting services of ARCHIVER can be summarised as follows:

- Support for tens of PBs of scientific data over the years and sustained data ingest rates capabilities from 1-10 GB/s.
- Follow best practices as foreseen in standards such as ISO 27000/27040/19086 for storage and CoreTrustSeal[18] for trusted digital repositories.
- Support of open standards for data preservation (such as PREMIS[19], METS[20] and Bagit[21]) and aligned with Cloud portability frameworks such as the SWIPO Code of Conduct[22] to demonstrate exit plans as concrete measures to prevent vendor lock-ins.
- Open, well established APIs (OpenAPI[23] and Rest[24]) for workflow integration with existing IT infrastructures.
- Implement organizational and technical measures adopting a “Privacy by Default and by Design” under European legislation such as GDPR[25] and Free Flow of Data[26].
- Promote innovative business models, taking into account preservation planning and the cost of long-term data archiving, compatible with public research sector procurement cycles and individual research grants.
- Support reproducibility of scientific results, independently of the underlying infrastructure.

The ARCHIVER R&D tender closed on 28th April 2020. The preparation phase counted 35 companies participating in the OMC, the majority of which are Small and Medium Enterprises (SMEs), as well as many public organisations in need of innovative data archiving and preservation services. During the submission period, the R&D tender was downloaded 147 times in 29 countries translating into 15 R&D bids involving 43 organisations and companies. The ARCHIVER team, in particular the CERN Procurement Service had to adapt the submission period in sake of the COVID-19 outbreak. Given the situation, the result obtained was impressive as the period ranging from R&D tender publication to contract signature was less than 5 months, being able to limit further delays.

2.1.1 Design phase

The Design phase started with a remote Kick-off event held on 8 June 2020 where the selected consortia presented their proposals for technical designs. The consortia selected, led by companies headquartered in Europe, represent a good mix of capabilities and expertise: European SMEs specialists in data preservation services, European infrastructure providers, American hyperscalers and public funded organisations supporting both proprietary and FOSS[27] components.

A total of 5 consortia[28] were selected to produce designs (Figure 3).



Fig. 3. Consortia participating in the ARCHIVER Design Phase.

The ARCHIVER Project Management Team engaged in a set of R&D monitoring activities during the Design Phase (Figure 4). These included regular R&D assessment reports describing the main strengths vs. weaknesses of the proposed solution, R&D prioritization and recommendations were shared with the selected consortia.

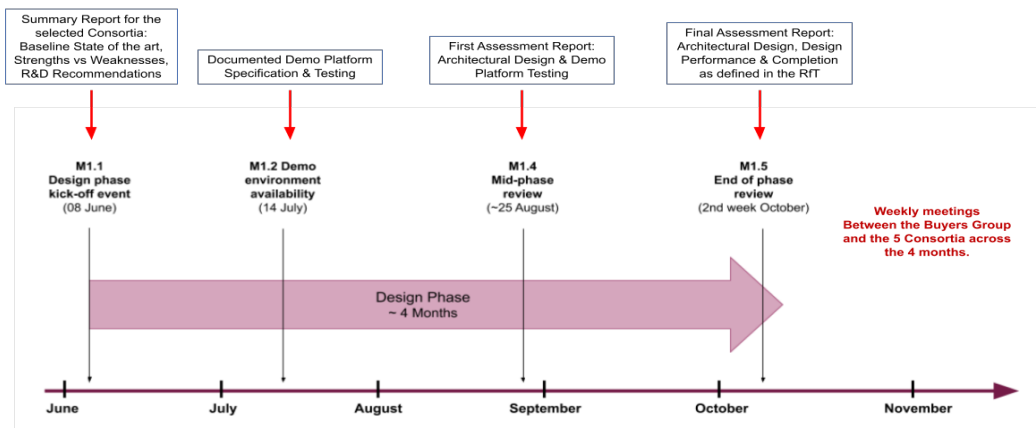


Fig. 4. Design Phase timeline and R&D monitoring checkpoints.

In addition, the ARCHIVER team requested access to demo systems of the five existing platforms in order to establish a reference point in the current state of the art, compare it with the resulting architectures proposed and consequently assess the effort needed from the starting point for the R&D to be performed during the project.

2.1.2 Prototype phase

The three most promising designs (Figure 5) were selected to proceed to Prototype phase via the organisation of a call-off (mini-competition). The call-off period ensures the transition between project phases described in section 2.1. The assessment took place in November 2020, executed by the ARCHIVER Technical Review Board (TRB) composed by technical representatives of each of the partner research organisations, in two steps: assessments of the Design phase defining the eligibility to proceed to Prototype and consequent evaluation of the Prototype R&D bid of each of the qualified consortia. The three awarded consortia participated in the Prototype Phase virtual Kick-off event held on the 7th December 2020.



Fig. 5. Selected Consortia for the Prototype Phase.

The consortium led by T-Systems (Figure 6) has produced a gap analysis for OSS Archivemata with an R&D plan to make innovative integrations between the data management layer based on Onedata[29], Archivemata, Flowable[30] and OpenFaaS[31] to manage workflows and metadata extraction at scale. The components and R&D performed by the consortium will be made available with OSS licensing.

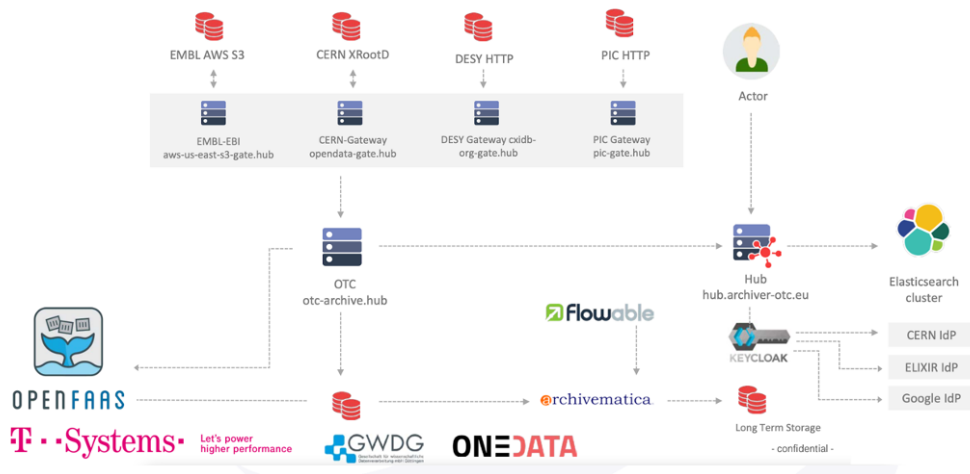


Fig. 6. Prototype architecture of the T-Systems consortium.

Libnova (Figure 7) is supporting a prototype based on proprietary LibSAFE SaaS using infrastructure provided by AWS[32] and Voxility GmbH[33] with all capabilities in the UI available via API endpoints and active monitoring to control data loss.

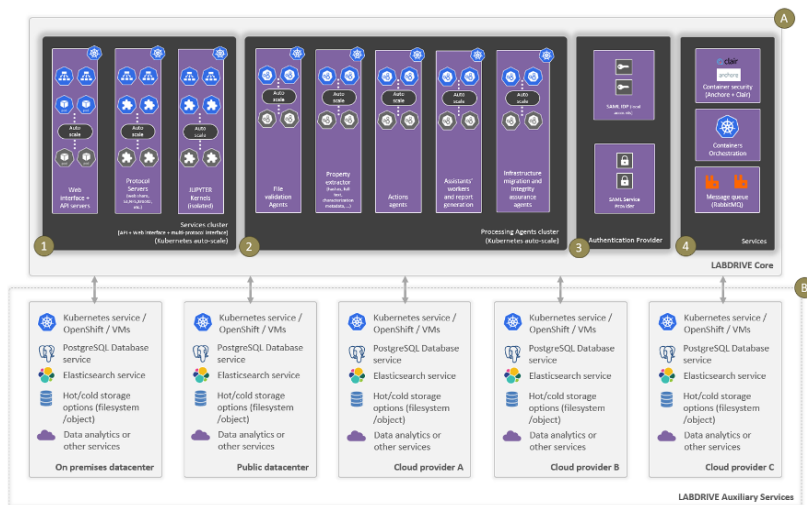


Fig. 7. Prototype architecture of the Libnova consortium.

The architecture is composed by two groups:

- Group A (Core components) as a set of software components running inside Kubernetes containers. The number of containers running in parallel of each class can be adjusted manually or automatically by the platform based on service demand to ensure full scalability.
- Group B (Auxiliary services) based on Core services. When running in “on-prem” mode, organizations will need to provide them for the platform to work. When running “as a service”, the service provider (LIBNOVA in principle) will provide them.

In Arkivum (Figure 8), the overall approach is composed by Kubernetes orchestration together with micro-services architecture implementing Archivemata workflows as an option, scaling to multi-petabyte volumes of billions of objects.

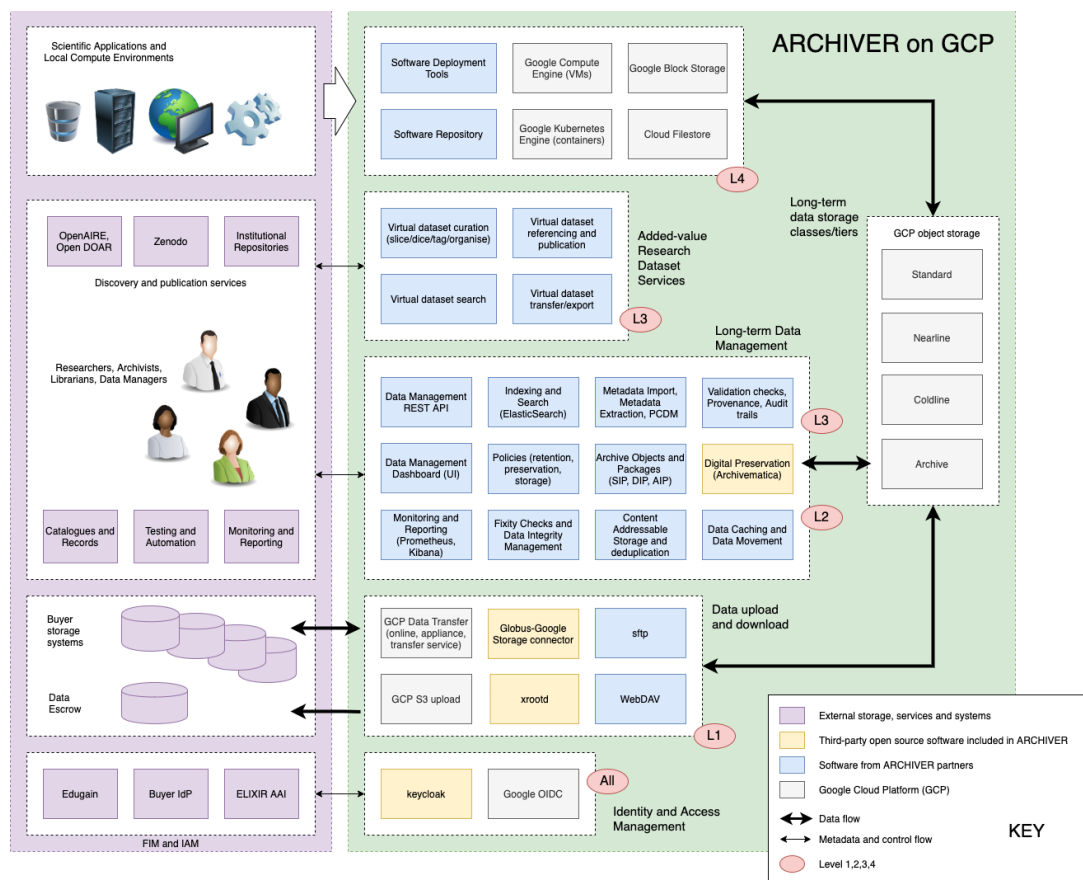


Fig. 8. Prototype architecture of the Arkivum consortium.

It consists of a service-oriented SaaS stack deployed on Google Cloud Platform (GCP). It promises to address all four layers of ARCHIVER described above. The solution can also be deployed on-premise or in a hybrid cloud configuration.

Globally, the ARCHIVER solutions are pushing significantly the current state of the art. If mapped against the R&D challenge Layers depicted in Figure 1, the main areas of development are described in the table below.

Table 1. Main areas of development

Area	ARCHIVER R&D
Data Archiving & Preservation Scalability (Layer 1/2) tens of PBs	Support for tens of PBs of scientific experimental data volume with growth over the years (Layer 1). Demonstration of support of multiple tenancies, with data isolated where each organization can be a top-level tenant in the offered solutions. Subtenants of a given top level tenant are able to share datasets amongst themselves. Sustained data ingest rates capabilities from 1-10 GB/s. Development of new architectures to manage metadata at scale (e.g., automated indexing of several PB tens of scientific data while maintaining interoperability, serverless functions for scalable metadata extraction, etc.), including the development of capabilities for handling unstructured or missing metadata. Concrete improvement of current Open-Source solutions for data preservation (i.e., Archivematica) at scale.

FAIR (Layer 2/3/4)	<p>Total FAIR approach by “Design & by Default” for the long-term and not only the outset: High level of redundancy, long-term planning for monitoring of data integrity on top of the infrastructure layer capability to detect unwanted changes such as file corruption or loss.</p> <p>Access and permission management against repositories and various collections supporting Federated Identity and Access Management.</p> <p>Richer API set for system integration: basically, everything in a GUI supported via programmatic APIs. Support of best practices (DPC RAM[34], CoreTrustSeal).</p> <p>Support of tools for search, look up or filter potential datasets rapidly, to access dataset metadata and decide on its relevance.</p> <p>Demonstrators of full reproducibility of services (initial examples are database services and/or software distribution services) on top of the resulting supported data archives. Ability to run additional scientific analyses independently of on-prem infrastructure in stages. Container Orchestration Engines at scale for automated, infrastructure independent deployments.</p> <p>Fast information tagging and indexing for PB of data (easy and broader search, as a strategy to promote open data access). Support for OSS components and vendor independent standards and interfaces (such as PREMIS, METS and Bagit).</p> <p>Demonstration during the Prototype and Pilot of exit strategies (cloud to on-premises, and cloud to cloud) to prevent vendor lock-ins. Data linking and discovery of defined new datasets; Support of incremental archiving.</p>
-----------------------	---

2.1.3 Pilot phase

Two consortia will be selected to deploy expanded prototype services during the pilot phase scheduled from October 2021 to June 2022. These pilots will be pre-production services to be tested in aspects such as performance, scalability and robustness. These services will be exposed to end-users and early adopters, in order to determine if the resulting services are suitable for their needs.

As part of the project activity, ARCHIVER is promoting the development of sustainable business models for the resulting services with increased emphasis during the Pilot Phase, requesting the selected Contractors to provide total cost of services (TCS) and commercialisation plans for their solutions. This aspect is fundamental in order to provide a clear costing perspective to organisations that will purchase the resulting services at the end of the project.

An additional aspect that will be prioritised during the pilot phase is the definition of a clear path for certification of the ARCHIVER resulting services as trustworthy digital repositories against schemes such as CoreTrustSeal that is a candidate for inclusion in the EOSC rules of participation[35]. Certification can be a clear benefit for the resulting services as repositories that can demonstrate to their future users that an independent authority has evaluated them and endorsed their trustworthiness.

A similar call-off period as the one organised in the previous phase, taking place in September 2021 will ensure the transition from three Prototypes to two Pilots.

2.2 R&D testing and validation

In order to monitor the R&D progress during the Prototype and Pilot phases a testing environment has been created. The environment is managed by a unified portal able to track issues and report results. This portal resides in a cloud-based JIRA instance[36].

The organisations forming the Buyers Group are promoting a single portal approach in order to simplify monitoring of the testing progress and ensure an overall agile test management. To make the approach effective, both Buyers and Contractors must make use of the JIRA instance and keep it updated when tests are executed, or issues (i.e bugs) are found.

The testing method is staged, satisfying a set of requirements:

- Full Alignment with the requirements defined in the R&D tender.
- Allow both manual and automated test deployment and execution.
- Provide a central results repository, supporting user/group role authorisation to ensure results confidentiality.
- Support collaborative work via tasks and issues definition, tracking progress following the Agile principles[37].
- Contribute to the assessment of Contractors deliverables (such as the Prototype, Pilot systems), associated components and documentation.

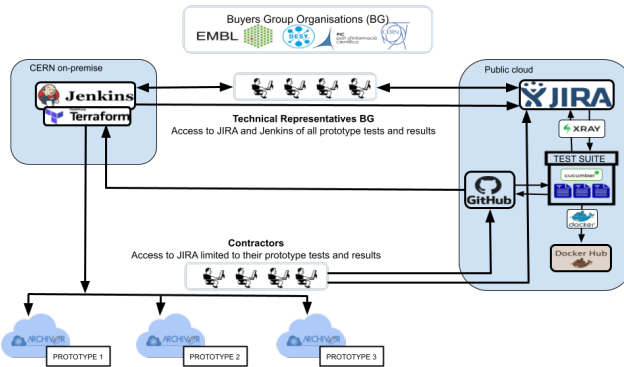


Fig. 9. Continuous Testing Environment for ARCHIVER’s testing activities.

The environment for continuous testing during project execution is composed by a heterogeneous set of components:

- **JIRA** - Provides the access portal for test management, tracking and overall agile project management, run in a cloud-based instance.
- **Jenkins** - Open-source automation server for continuous integration (CI). It runs on the CERN cloud, specifically on OpenShift PaaS.
- **Terraform** - Open-source IaaS provisioning tool, to be used to provision virtual machines on the contractor’s cloud for the automated testing workflow, in order to run the tests on those resources if needed to avoid overloading the Jenkins server.
- **XRAY** - JIRA test management plugin for test execution, tracking and reporting results.
- **Cucumber** - Coding component to empower Behavior-Driven testing, based on Gherkin Syntax. It is used with the TEST SUITE framework described below, to allow test cases to become readable for any non-technical user, or track FAILED/PASSED steps, in order to quickly identify problems. It can also be integrated through the XRAY plugin to manage and organise complex test workflows.
- **TEST SUITE** - Combination of third party tools and a set of scripts (standalone, independent scripts or Cucumber based), conforming a framework to ease the deployment of automated tests. More broadly, any third-party toolkit or script can be considered for inclusion in the test suite. To mitigate potential limitations of this server, the EOSC Test Suite[38] or a similar approach relying on automation tools can be used.
- **Docker** - Open-source tool that provides an abstraction layer allowing to run applications on heterogeneous environments, making use of isolated containers on which Docker images conform the OS.
- **Docker Hub** - Official and online registry hosting public and private Docker images.
- **GitHub** - Public hosting service for software development and version control using Git[39].

In addition, a variety of testing approaches and workflows is covered by the environment, to provide a greater degree of flexibility for test execution. Three testing workflows are supported in the environment depicted in Figure 9:

- **Manual test:** Unmanaged manual execution of a test case with optional manual result reporting. Workflow supports Web interface, CLI, programmatic API or SDK.
- **Automated test:** test deployment and execution process as well as reporting of results is automated. The automated process fetches the required scripts from a GitHub repository (certain tests may use a Docker[40] image). Once the run completes, results are automatically reported back to a centralized reporting system.
- **Independent test:** testing workflow with no reporting in JIRA. Workflow is used to provide the necessary degree of freedom to run tests as many times as necessary for developing and/or debugging purposes.

Tests are defined by the technical representatives of the Buyers Group organisations. The assessment of the tests is defined with four states: Passed, Passed with an issue, Failed with possible workaround and Failed. The specific criteria are defined for each test as depicted in Figure 10.

The testing activity is organised in sprints of two weeks. Sprints are short time-bound periods with a predefined set of work. ARCHIVER sprints are organised as follows:

1. Contractors inform buyers when functionality is ready to be tested in their platform.
2. Buyers prioritize testing that functionality and enter it into the backlog.
3. Two-week sprints are organised, driven by prioritized entries in the backlog.
4. A sprint review/retrospective is organised, providing feedback to the next sprint.

- E.g. 1.4.3, “Webhook”
 - **Passed:** The Archive notifies the user with a webhook call that an AIP is ready, enabling the user to fetch the data.
 - **Passed with an issue:** The Archive can notify the user with a real time notification, but without using the web callback as defined by the user.
 - **Failed with a workaround:** The status of an AIP creation is not notified to the user, but the Archive allows the user to pull some meta-information about the on-going - or finished - process.
 - **Failed:** The Archive does not provide the user with an easy mechanism to know the status of an archival process.

Fig. 10. Example of a test assessment in ARCHIVER.

In terms of code management, the ARCHIVER team is maintaining one private repository per contractor, all hosted on a public GitHub organisation[41]. Representatives of Buyer Group organisations have access to all Contractor repositories while Contractors have access only to their own repository. Docker Hub is used as the official and online registry hosting public and private Docker images.

In what concerns large scale tests, a peak test period was agreed and scheduled with the Contractors, preserving the peak of intense testing activity. This scheduling also allows a resource envelope defined to allow staging infrastructure resources, avoids having resources ready too soon, and consequently wasted. In addition, it encourages Contractors to provide elastic archives, minimizing idle resources.

3 Sustainability of the resulting services after the project

The ARCHIVER project has created an Early Adopter programme[42] whereby additional use cases and procurer organisations can be incorporated at each of the project phases. A set of additional use cases, expanding further the set of supported scientific domains, have been collected from 11 Early Adopter[43] publicly funded research organisations, external to the ARCHIVER consortium.

In addition, ARCHIVER is establishing a sustainability plan for the resulting services after the end of the project. Currently, three scenarios are being considered:

- Scenario (I): PPI (Public Procurement of Innovation) - this is a follow up EC funded project after the PCP, where the group of organisations forming the Buyers Group would act as «launching customers» of the innovative PCP solutions. The emphasis of the activity is on extended conformance testing and verification of the resulting ARCHIVER PCP services.
- Scenario (II): Negotiated Procedure - This scenario is valid for both Buyer and Early adopter organisations, for limited volumes only; to carry out further testing of new services after the PCP ends.
- Scenario (III): Proposed Collaboration with EOSC-Future - ARCHIVER has proposed to sign a collaboration agreement to the recently EC funded EOSC FUTURE project during Q2 2021. The agreement would include the following activities such as Identifying a set of commonly agreed additional use cases expanding the research communities, and requesting the ARCHIVER contractors to estimate work/costs to support the additional use cases as part of the pilot phase tender call-off process. The ultimate objective is to implement a model focussed on procuring services beyond the lifetime of ARCHIVER for a broader set of use users in the EOSC[44] context, including TB volumes.

4 Conclusions

The R&D challenge of digital archiving goes beyond simply storing data: keeps intellectual control of data and associated products for decades, rendering research outputs reusable. To this end, ARCHIVER is acting as a template to commoditise archiving and preservation in the research domains. The ARCHIVER consortium has now strong evidence there will be suitable solutions developed as a result of the project. Significant innovation has already been achieved with the selected prototypes: scalability, API support, cost-effectiveness data workflows.

The ARCHIVER consortium is fully committed: CERN, EMBL-EBI, DESY and PIC are allocating significant effort assessing, testing and ingesting petabyte experimental data resulting in a significant push of current state-of-the-art. The interaction between the technical representatives of the Research Performing Organisations (RPOs) part of the ARCHIVER Buyers Group and the contractors, is generating good progress when compared to other project formats, including project dissemination actions.

Finally, ARCHIVER is a sustainable model for the resulting services, verified and aligned with the EOSC Rules of Participation[45], beyond the project lifetime.

References

1. Data Preservation in High Energy Physics: <https://dphep.web.cern.ch>
2. CERN announces new open data policy in support of open science: <https://ep-news.web.cern.ch/cerns-new-open-data-policy-supports-open-science>
3. J. Fernandes, B. Jones, M. Devouassoux, D. Buyse, R. Van Calenberg, V. Motsenitalis, J. Urban, *ARCHIVER D2.1- State of the Art, Community Requirements and OMC Results*, Zenodo (2020)
4. Pre-Commercial Procurement: <https://ec.europa.eu/digital-single-market/en/pre-commercial-procurement>
5. Horizon 2020 - The EU Framework Programme for Research and Innovation: <https://ec.europa.eu/programmes/horizon2020>
6. The European Organization for Nuclear Research: <https://home.cern>
7. The European Bioinformatics Institute: <https://www.ebi.ac.uk>
8. German Electron Synchrotron: <https://www.desy.de>
9. The Port d'Informació Científica: <https://www.pic.es>
10. ARCHIVER Prior Information Notice: <https://ted.europa.eu/udl?uri=TED:NOTICE:87071-2019:PDF:EN:HTML>
11. J. Fernandes, B. Jones, M. Devouassoux, D. Buyse, R. Van Calenberg, V. Motsenitalis, J. Urban, *ARCHIVER D2.1- State of the Art, Community Requirements and OMC Results*, Zenodo (2020)
12. ARCHIVER | Early Adopters Programme: <https://www.archiver-project.eu/early-adopters-programme>
13. Deployment Scenarios | Technical Summaries: <https://www.archiver-project.eu/deployment-scenarios>
14. ESFRI Projects & Landmarks: <http://roadmap2018.esfri.eu/media/1049/roadmap18-part3.pdf>
15. A distributed infrastructure for life-science information: <https://elixir-europe.org>
16. High-Luminosity Large Hadron Collider project: <https://hilumilhc.web.cern.ch/content/hl-lhc-project>
17. European X-Ray Free-Electron Laser Facility: <https://www.xfel.eu>
18. Core Trustworthy Data Repositories: <https://www.coretrustseal.org>
19. PREservation Metadata Implementation Strategies: <https://www.loc.gov/standards/premis/v3/premis-3-0-final.pdf>
20. Metadata Encoding and Transmission Standard: <https://www.loc.gov/standards/mets>
21. The BagIt File Packaging Format (V1.0): <https://tools.ietf.org/html/rfc8493>
22. SWIPO Code of Conduct: <https://scope-europe.eu/en/projects/swipo-code-of-conduct>
23. The OpenAPI Specification: <https://swagger.io/specification>
24. REpresentational State Transfer, Rest: <https://restfulapi.net>
25. General Data Protection Regulation: <https://gdpr.eu>
26. Free flow of non-personal data: <https://ec.europa.eu/digital-single-market/en/free-flow-non-personal-data>
27. J. Feller, ed. *Perspectives on Free and Open Source Software*, MIT (2005)
28. Design Phase Award | Archiver-project: <https://www.archiver-project.eu/design-phase-award>
29. Onedata: <https://onedata.org>
30. Flowable: <https://flowable.com>
31. OpenFaaS: <https://www.openfaas.com>
32. Amazon Web Services (AWS) | Cloud Computing Services: <https://aws.amazon.com>
33. Voxility GmbH: <https://international.eco.de/mitglieder/voxility-gmbh>
34. DPC Rapid Assessment Model: <https://www.dpconline.org/digipres/dpc-ram>
35. Directorate-General for Research and Innovation, EOSC Executive Board, *EOSC Rules of Participation, Report from the EOSC Executive Board Rules of Participation Working Group*, EU Publications (2021)
36. ARCHIVER Project Jira Instance: <https://archiver.aflasian.net>
37. Principles behind the Agile Manifesto: <https://agilemanifesto.org/principles.html>
38. EOSC Test Suite: <https://eosc-testsuite.readthedocs.io/en/latest>
39. Git | free and open source distributed version control system: <https://git-scm.com>
40. Docker: <https://www.docker.com>
41. ARCHIVER Project GitHub organisation: <https://github.com/archiver-project>
42. ARCHIVER Early Adopters programme: <https://www.archiver-project.eu/early-adopters-programme>
43. ARCHIVER Early Adopters Use Cases: <https://www.archiver-project.eu/early-adopters-use-cases>
44. The European Open Science Cloud: <https://eosc-portal.eu>
45. Directorate-General for Research and Innovation, EOSC Executive Board, *EOSC Rules of Participation, Report from the EOSC Executive Board Rules of Participation Working Group*, EU Publications (2021)