

Recommendations Consultation

Long Term Data Preservation Task Force

October 2023

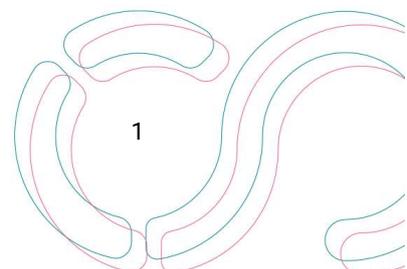
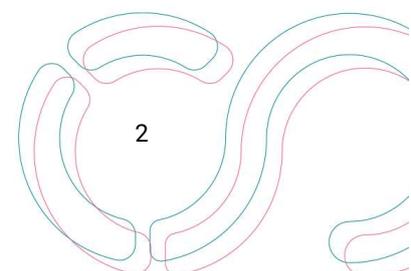


Table of contents

TABLE OF CONTENTS	2
INTRODUCTION	3
STRUCTURE AND READING OF THE DOCUMENT	4
A VISION FOR OPTIMAL PRESERVATION OF FAIR DIGITAL OBJECTS WITHIN EOSC	5
FULL OVERVIEW OF ASSERTIONS AND RECOMMENDATIONS FOR OPTIMAL PRESERVATION OUTCOMES	5
DIGITAL OBJECTS, RESEARCH PROJECTS, REPOSITORY AND (META)DATA SERVICES.....	5
CURATION & PRESERVATION LEVELS	7
FAIR + TIME	8
PRESERVATION SYSTEMS, ACTIONS & OUTCOMES.....	8
TRUSTWORTHY DIGITAL REPOSITORIES (TDR) AND OTHER (META)DATA SERVICES.....	9
STANDARDS, ASSESSMENT, CERTIFICATION & ALIGNMENT	11
OUTCOMES, JUDGEMENT AND GATEKEEPING	11
RETENTION, APPRAISAL & RE-APPRAISAL.....	12
TRANSPARENCY OF SERVICES, ARTEFACTS AND LEVELS OF CARE FOR OBJECTS.....	13
GENERIC VERSUS DOMAIN, DISCIPLINE AND OBJECT TYPE SPECIFIC ISSUES.....	13
ROLES AND RESPONSIBILITIES	14
FINANCE & FUNDERS.....	16
NETWORK OF TDRS	17
ADDRESSING PRESERVATION AT EUROPEAN, NATIONAL AND INSTITUTIONAL LEVEL	18
RECOMMENDATIONS AT EUROPEAN LEVEL	18
RECOMMENDATIONS AT NATIONAL LEVEL	19
RECOMMENDATIONS AT INSTITUTIONAL LEVEL	19
VERSION HISTORY	20



Introduction

The EOSC Association's Long Term Data Preservation Task Force (LTDP-TF)¹, which functions under the Advisory Group (AG) *Technical challenges EOSC*² has previously shared an overview³ document for public feedback. A paper on the frame of reference used to guide the task force approach was presented⁴ at the PV2023 conference.

In this initial consultation draft of the task force outcomes the decision has been taken to minimise prose and maximise the presentation of brief assertions and recommendations for open feedback. Numerous scoping assertions are included because the work of the task force to date has revealed varying interpretations of some key research infrastructure concepts that need to be contextualised⁵. This approach supports granular feedback on each assertion and recommendation that will inform future interactions of this document. We would appreciate responses from a wide range of stakeholders across the digital object management lifecycle. Respondents are asked to propose amendments to text or provide explanatory comments if they object to the statements as presented. The feedback template can be found [here](#) which can be sent to preservation-tf@eosc.eu.

The results of this consultation period will be used to revise the assertions and recommendations for future reference by a range of actors. These will inform a final report for this iteration of the LTDP-TF alongside proposals to the EOSC Association on how further progress can be guided, monitored and supported through their activities, including, but not limited to future task forces.

The work of the task force benefits from numerous prior efforts in research data management, curation and preservation, particularly the FAIR Forever report⁶, which notes that “Digital preservation involves the series of managed activities necessary to ensure the continued access to research data for as long as necessary, which encompass actions and interventions throughout the lifecycle—not just at the creation of FAIR data or the transfer and ingest to a certified archival repository.” The task force’s overview discussion paper⁷ made it clear that the TF perspective acknowledges the importance of FAIR data and trustworthy repositories as central to the EOSC vision, but also notes that a full (meta)data lifecycle perspective is critical. The overview further highlights that to achieve preservation *outcomes* we depend on sustained and sustainable preservation *systems* that take responsibility for monitoring the technical and user environments and, where necessary, take preservation *actions* on digital objects (data

¹ <https://www.eosc.eu/advisory-groups/long-term-data-preservation>

² <https://eosc.eu/eosc-task-forces>

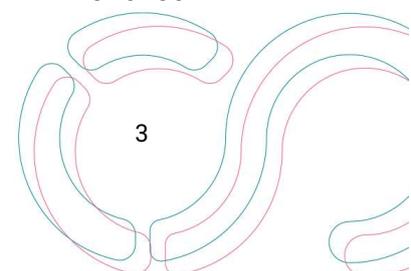
³ EOSC Preservation: Overview Discussion Paper <https://doi.org/10.5281/zenodo.7516259>

⁴ Preservation in the context of EOSC. Sustainable repositories curating digital objects from a long-term FAIR enabling perspective <https://indi.to/dpnjt>

⁵ Where the TF has used an internal working definition this is flagged with WG in superscript^{WG}

⁶ FAIR Forever? Long Term Data Preservation Roles and Responsibilities, Final Report <https://doi.org/10.5281/zenodo.4574234>

⁷ EOSC Preservation: Overview Discussion Paper <https://doi.org/10.5281/zenodo.7516259>



and/or metadata). These preservation systems (including, but not limited to entities that self-describe as libraries, archives or repositories) exist within a wider network of data and metadata services. Trust across these services, through transparent practice, is vital to the success of any federated research infrastructure, including EOSC.

While the task force notes the importance of agreeing on standards and developing metrics and tests for assessment, it is important to recognise that data services and digital objects are in the relatively early stages of a journey towards trustworthiness and FAIRness. Successes should be acknowledged and credited but less mature data services and digital objects must be invested in for improvement. The Task Force has worked to provide targeted guidance through The Strategic Research & Innovation Agenda and its Multi-Annual Roadmap⁸ that have already resulted in calls that will invest in preservation. However, the journey towards agreed community criteria and broad adoption must be seen as a marathon and not a sprint. The scale of data and metadata collected by, created by, or of interest to researchers of all kinds, whether academic, commercial or public is vast. The baseline task of identifying and providing basic deposit, storage and access for these digital objects is already a challenge. Initial curation to ensure digital objects meet desirable criteria, such as FAIRness, requires an initial investment. For digital objects with a long term value we require the resources and skills to deliver preservation systems, actions and outcomes. Meanwhile, appraisal (review, evaluation, selection and decisions on the level of care to be applied) and reappraisal of digital objects for their value as digital assets is complex but necessary. The cost of inaction is the potential loss of resources necessary for reproducibility, replication and reuse; the benefit of action is the continued availability of resources for new and novel explorations within and across disciplines.

References to 'digital objects' within this paper are not limited to data points created or collected during the course of research and include a wide range of data and metadata constructs, including software, that are relevant to EOSC. It may not be possible, practical or even desirable to retain, curate or preserve every digital object, but any related appraisal decisions should be transparent and the resultant levels of 'care' (retention, curation or preservation) being offered by repositories and received by digital objects should be clear. Addressing individual object types, disciplinary issues and aligning with the full contextual legal and interoperability framework of the EOSC are prerequisites, but beyond the scope of this task force and paper.

Structure and reading of the document

- The assertions and recommendations are presented under a sequence of headings.
- Working definitions developed within the Task Force are presented with a superscriptWD.

⁸ <https://eosc.eu/sria-mar>

- In the body of this document quoted text is taken directly from the FAIR Forever report⁹, though numerous other items reflect the report’s recommendations¹⁰.
- Recommendation bullets are preceded by an @ symbol¹¹.
- If a recommendation suggests that something needs to be ‘defined’ then the implied follow up actions are ‘adopted’ and ‘made transparent’.
- Most of the recommendations made in this paper apply to the wider data repository landscape and associated stakeholders on a European, national, institutional, and thematic (domain, discipline) level. Those recommendations that are specific to a certain context are headed accordingly.

A Vision for Optimal Preservation of FAIR Digital Objects within EOSC

In the first output of the task force shared with the community for consultation, the task force defined the following vision for optimal preservation of FAIR Digital Objects within the context of EOSC:

Digital objects that act as inputs to, or outputs from, research are identified, findable and accessible in environments that support good storage practice. These objects are subject to appraisal, and reappraisal over time, to assess their value, their impact and the associated costs, risks and benefits. Ongoing appraisal informs the level of investment in the retention, curation and long-term preservation of digital objects. The levels of care, and changes to levels of care, provided by repositories and assigned to digital objects are transparent to (meta)data funders, depositors and users.

This vision is at the core of the task force’s recommendations that are listed in this document.

Full overview of Assertions and Recommendations for Optimal Preservation Outcomes

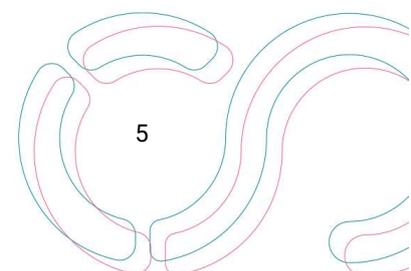
Digital Objects, Research Projects, Repository and (meta)data Services

Desirable characteristics of digital objects, including FAIR, at the point of re-use depend on a range of storage, curation, preservation and appraisal activities undertaken by different data services, including repositories, throughout the lifecycle. These outcomes depend on clear criteria, transparency and a long-term perspective.

⁹ FAIR Forever? Long Term Data Preservation Roles and Responsibilities, Final Report
<https://doi.org/10.5281/zenodo.4574234>

¹⁰ In a separate process the LTDP TF is highlighting FAIR Forever items that are relevant to other EOSC Task Forces.

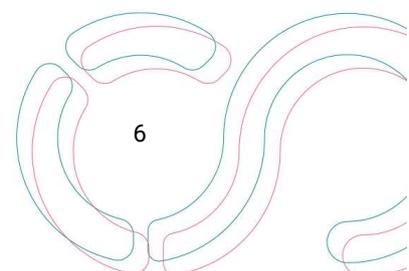
¹¹ Selected for accessibility reasons as this symbol is commonly expressed by screen readers.



1. Digital objects are constructs of data and metadata including, but not limited to, research outputs such as datasets, and software.
2. Numerous types of digital objects exist. Some characteristics are applicable to many types of objects, others are more specialised.
3. In addition to disciplinary-specific requirements and digital objects stored as files in a single location there are numerous complex or new and novel digital objects including distributed artefacts such as ontologies.
4. Coverage of all objects by type is important, but beyond the scope of this task force¹².
5. Projects collect or create digital objects of relevance to research.
6. Digital objects must be appropriately cared for during research.
7. The Principles that digital objects should be Findable, Accessible, Interoperable and Reusable (FAIR) define desirable outcomes of digital objects' care.
8. Projects are often time limited, reducing their ability to offer operational services required for preservation for the longer term.
9. Digital objects can retain their value beyond the life time of the project that curated or created them.
10. The value of digital objects is derived in part from different reuse cases including for replication, reproducibility, supporting assertions made in published papers or integration into new and novel research.
11. During or after a research project other digital object services may support the care of data and metadata.
12. Different data and metadata services¹³ are required to support research across the full digital object lifecycle.
13. Any digital object service holding data or metadata at any level of care (retention, curation or preservation) can be very broadly defined as a repository.

¹² The FAIR Forever report provides a use case covering a A Legacy Code Software Preservation Service within EOSC. This is just one of many valid scenarios for the development of preservation services for specific categories of digital object

¹³ References to 'Data Services' in the context of this paper include the full range of entities (organisations, partnerships etc) that hold or curate or preserve or provide other ancillary services around all types of digital objects.



14. @Additional work is needed to define different types of digital object systems that offer (meta)data services, including the functions and activities they require and the levels of care they provide for data and metadata.
15. Effective storage, including multi-copy redundancy and integrity measures, is necessary but not sufficient for preservation.
16. @Minimum criteria for acceptable storage practices in different scenarios should be defined as a foundation for all levels of retention, curation and preservation services.
17. Different types of data services benefit from being transparent on their current level of storage, curation and preservation practice, as this increases trust by the user and funders alike.

Curation & Preservation Levels

18. This paper defines the different service levels that a repository may apply as follows¹⁴:
 - Z. Level Zero. Content distributed as deposited.** Unattended deposit-storage-access. No curation or long-term preservation.
 - D. Deposit Compliance.** Data content and supporting metadata deposited are checked at the point of deposit for compliance with defined criteria e.g. data formats and metadata elements. If these criteria are not met the digital objects are either returned to the depositor for change, or the repository undertakes the necessary curation steps to ensure they comply (reaching level C, below).
 - C. Initial Curation.** Data content and supporting metadata deposited are curated to meet defined criteria e.g. for data formats and metadata elements. Curation for initial access and use, but no long-term preservation.
 - B. Logical-Technical Preservation.** In addition to D and/or C above the repository takes long-term responsibility for ensuring that the data and metadata are updated over time to newer standards and formats in response to technical risks (e.g. file format or software obsolescence), the changing needs of the designated community (e.g. newer alternate formats become necessary for reuse).
 - A. Conceptual preservation for understanding and reuse.** In addition to B above, the repository monitors changes to the definition and demands of their designated community, including their knowledge base, and takes responsibility for the preservation actions that ensure digital objects can be understood and re-used. Often this will involve updates to the content of metadata elements and other semantic artefacts such as controlled

¹⁴ CoreTrustSeal Standards and Certification Board. (2023). Curation & Preservation Levels: CoreTrustSeal Discussion Paper (v02.00). Zenodo. <https://doi.org/10.5281/zenodo.8083359>

vocabularies and ontologies. For some repositories it may include responsibility for editing the structure and content of deposited data.

19. @Data services, including repositories should specify all the levels of care they apply to objects within their collection, including through repository and digital object registry metadata.
20. @Digital objects should include metadata that specify their level of care and the timeframes or criteria for reappraisal of the level of care.

FAIR + Time

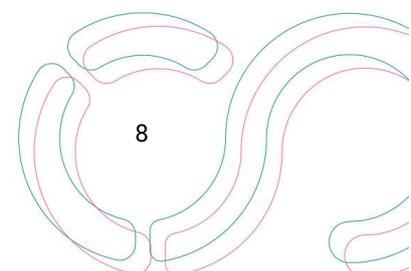
As time passes, technologies and the needs of (meta)data (re)users evolve. This may result in digital objects becoming less FAIR over time. Transparent FAIR enabling data services are required to monitor the situation and take additional FAIR enabling preservation actions over time as necessary.

21. Objects may be made FAIR before deposit in a repository (Level D) or made FAIR by initial curation within the repository (Level C).
22. As technical infrastructure and the needs of user communities evolve, digital objects that were initially made FAIR may require additional preservation actions over time to ensure they remain FAIR.
23. @FAIR-enabling practices to be undertaken by all data services should be defined.
24. @FAIR-enabling practices undertaken by data services should be made transparent to users and funders to increase trust in services.

Preservation Systems, Actions & Outcomes

To achieve preservation outcomes (long-term access and use of data and metadata maintaining key characteristics, include FAIRness) we depend on sustainable preservation systems (organisations, partnerships or other entities) that take responsibility for monitoring the technical and user environments and, where necessary, take preservation actions on digital objects (data and/or metadata).

25. Preservation systems^{WD} (organisations, partnerships, archives, repositories, libraries, galleries, museums) provide a sustainable organisational infrastructure to monitor the evolution of a technical infrastructure and the needs of user communities and to undertake preservation actions as necessary.
26. Preservation actions^{WD} include any action on a digital objects' data, metadata (e.g. format of schema updates) or an environment provided to interact with the digital object (e.g. emulation) that ensure the object retains desirable characteristics (e.g. FAIRness).



27. Preservation outcomes^{WD} are defined as an object maintaining desirable characteristics (e.g. FAIRness) over time.
28. Many (meta)data services, functions and activities are necessary for preservation, but preservation also requires additional actions with associated roles and costs.
29. @ Unique preservation functions and activities should be defined alongside functions and activities that apply to all (meta)data services.

Trustworthy Digital Repositories (TDR) and other (meta)data Services

Trustworthy Digital Repositories (TDR) comply with a set of organisational infrastructure, digital object management, technology and security requirements. TDR standards, such as the CoreTrustSeal currently require that repositories provide active preservation (Levels B and A above) to ensure ongoing re-use of digital objects. But many of the TDR requirements, including CoreTrustSeal, are applicable to a broader range of repositories and data services. Certification in general, and CoreTrustSeal in particular, are not always the appropriate solution, but the CoreTrustSeal Requirements provide a common reference and a transparent consultation approach that other criteria for repositories and data services would usefully align with.

30. Repositories that undertake to offer preservation services and can meet defined criteria for their organisational infrastructure, digital object management, technical infrastructure and security provision are candidates to be a Trustworthy Digital Repository (TDR).
31. The CoreTrustSeal¹⁵ is a not-for-profit foundation, developed in response to an RDA mission and maintained through the RDA that provides 16 Requirements and an associated peer review and certification process for TDRs.
32. The LTDP TF reasserts the conclusion of previous EOSC-relevant papers¹⁶ and groups¹⁷ that the CoreTrustSeal provides an appropriate mechanism to define core expectations of TDR and an exemplar for offering assessment and certification services.

¹⁵ <https://www.coretrustseal.org/about/>

¹⁶ European Commission, Directorate-General for Research and Innovation, Turning FAIR into reality – Final report and action plan from the European Commission expert group on FAIR data, Publications Office, 2018, <https://data.europa.eu/doi/10.2777/1524>

¹⁷ European Commission, Directorate-General for Research and Innovation, Jones, S., Aronsen, J., Beyan, O. et al., Recommendations on certifying services required to enable FAIR within EOSC – , Genova, F.(editor), Publications Office, 2021, <https://data.europa.eu/doi/10.2777/127253>

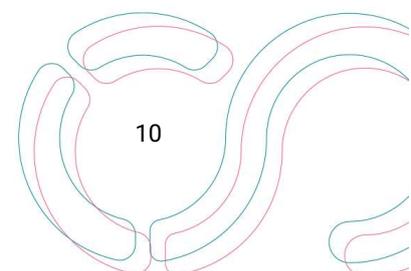
33. With the exception of the requirements for offering preservation services (CoreTrustSeal Requirement 09) and considering reuse (CoreTrustSeal Requirement 13) the CoreTrustSeal¹⁸ is relevant to all (meta)data services¹⁹.
34. @Other ongoing work to define repository and data service characteristics and expectations exist and should be encouraged and supported²⁰.
35. @To maintain clarity and alignment these other efforts should map and crosswalk their own criteria to the CoreTrustSeal. Any reductions, additions or variations versus the CoreTrustSeal should be documented and explained to support interoperability of standards and approaches.
36. Transparency on the current status of repositories for organisational infrastructure, digital object management actions, technology and security, as well as the roadmap towards improvement of TRUST²¹ and FAIR, will increase the knowledge base of funders, repositories and other data services, and increase the trust of digital object depositors and reusers.
37. The CoreTrustSeal applies to both generic and specialist repositories. It is domain aware (applicants must state any disciplinary communities that are supported) but domain agnostic.
38. @Efforts to define more specific domain/disciplinary criteria, or criteria that define expectations for specific types of digital objects should adopt the CoreTrustSeal requirements where possible, and elaborate around them where necessary.
39. @Additional work is needed to define different types of digital object systems, the functions and activities they undertake and the levels of care they provide for data and metadata.
40. @Roles and responsibilities including for complex partnerships, third party relationships and outsourcing should be understood and transparent.

¹⁸ CoreTrustSeal Requirements 2023-2025 <https://doi.org/10.5281/zenodo.7051012>

¹⁹ Review of CoreTrustSeal Applicability to non-Preservation (Trustworthy) Data Services <https://doi.org/10.5281/zenodo.7646134>

²⁰ Repository & (meta)data Services Functions & Activities: Crosswalk <https://doi.org/10.5281/zenodo.7690658>

²¹ Lin, D., Crabtree, J., Dillo, I. et al. The TRUST Principles for digital repositories. Sci Data 7, 144 (2020). <https://doi.org/10.1038/s41597-020-0486-7>



41. @Technical repository service providers' (storage providers, ARCHIVER²² etc) portfolio of service offerings should be clear and comparable for client end-users²³.

Standards, Assessment, Certification & Alignment

Defining standards for objects, repositories and other data services supports a common understanding of characteristics and goals. Assessment permits benchmarking of current status and planning for future improvements. Certification is one possible outcome of assessment to acknowledge good practice. Alignment of standards and assessment processes is an effective mechanism for agreeing desirable outcomes and providing guidance and support.

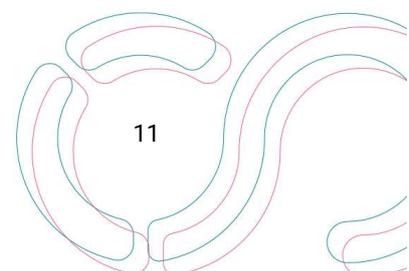
42. The development and application of standards provides mechanisms for agreeing and implementing a range of consistent practices.
43. Indicators, metrics and tests can be designed to assess compliance with standards.
44. Standards may be used to support, self-, peer- or third-party assessment approaches.
45. Assessments can be used to acknowledge quality (e.g. certification of repositories) and help service providers understand their current state and plan actions for future priorities and goals.
46. When an assessed entity has changed, the assessment may no longer be valid and may need to be repeated.
47. Repeated assessment may support an understanding of trends and progress over time.
48. @ standards and guidance should be developed, coordinated and maintained to provide full lifecycle information on preservation to researchers and preservation practitioners.

Outcomes, Judgement and Gatekeeping

Digital objects, repositories and data services at differing levels of maturity are on a journey towards FAIR and trustworthiness. Transparency of practice and assessments should be used as a tool for developing roadmaps on a journey of improvement. Exclusionary judgements and gatekeeping (e.g. inclusion in the EOSC) as a result of the outcomes of assessments must be avoided.

²² <https://archiver-project.eu/>

²³ E.g. via efforts such as <https://www.rd-alliance.org/rda-tiger>



49. Outcomes of assessments may include certification of repositories, assessing objects for their degree of FAIRness or compliance with a wide range of other criteria.
50. Perceptions of negative judgement due to outcomes of assessments could negatively impact services and reduce transparency.
51. Unnecessary gatekeeping of participation in or access to services (including EOSC) risks excluding potentially high-quality services and high-value digital objects from research infrastructures.
52. Any cases where an assessment may result in binary inclusion/exclusion outcome must be clearly justified and documented, e.g. a strict assessment outcome may be justified to protect sensitive data through information security measures, or to protect critical services through technical interoperability criteria.
53. @Digital Objects and the services that enable their FAIRness, deposit, storage, curation, access and preservation should be supported in transparent efforts to use assessment as a route to improvement.
54. @Efforts by repositories and other data services to share transparent information about their functions, activities and objects should be rewarded by targeted investment towards improvement.
55. @No data service or digital object should be unnecessarily excluded from any part of EOSC.
56. @FAIRness and TRUST must continue to be a supported journey for all parties.

Retention, Appraisal & Re-Appraisal

Any progress to retain, identify and evaluate a greater proportion of relevant digital objects is desirable. Some of these objects may be appraised as having sufficient value to be retained, curated or actively preserved. Reappraisal over time may change the level of care received by a digital object, or lead to a decision to delete the object. Like curation and preservation, appraisal and reappraisal depend on appropriate expertise. Levels of care, and changes to levels of care should be justified and transparent.

57. A long-tail of data exists that has not been brought into a managed storage, curation, or preservation system.
58. Any decision taken to evaluate a digital object with a view to delete it, retain it or assign or change a level of care is an appraisal decision.
59. The role and nature of data vary widely, and so there are different perspectives and insights into the kinds of preservation action that may be required.

60. Knowing which elements matter, and what metadata they might require is a subject specialist skill that in many cases can only be captured at the point of creation.
61. @Retention and reappraisal decisions and timescales, including guaranteed preservation timescales, should be transparent.

Transparency of Services, Artefacts and Levels of Care for Objects

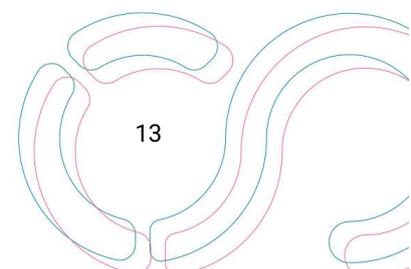
Exposing the different levels of service offered, and the levels of care received by objects, along with the evidence artefacts that support these assertions provides a level of transparency that informs understanding, cooperation and interoperability between the many actors and objects across the research data lifecycle.

62. Transparency of metadata, supporting artefacts (policies, preservation plans, data management plans) and digital objects fosters understanding, interoperability and continuous improvement between peer services.
63. Information shared through transparency can be designed to mitigate any risks to information security or competitive advantages.
64. Transparency supports the establishment of feedback mechanisms to engage expert communities of practice in the evaluation and improvement of services and objects.
65. @Each service should be transparent about the levels of service provided.
66. @Each object should have a clear level of care associated with the service(s) that take responsibility for them.
67. @Living and machine-actionable data management plans should form the basis of continuity through the research data lifecycle.
68. @Registries of repositories and other data services should align service-level metadata and supporting information.
69. @Registries of digital objects should align metadata relating to retention periods, appraisal periods and levels of curation and preservation.

Generic versus Domain, Discipline and Object Type Specific Issues

All of the assertions and recommendations above, and the roles and actors described below are specified at a high level due to the scope and timescales of the task force. All of these imply a need for more exploration of specific needs as disciplines, domain and object type level.

70. @ Addressing the challenges of metadata and interoperability in and across scientific domains and disciplines must be supported in further investment to identify granular needs for specific types of digital objects and disciplines.



Roles and Responsibilities

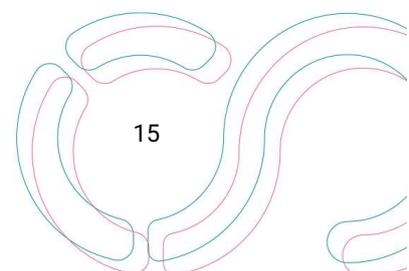
Roles and responsibilities across partners and lifecycles are not always well defined and many of them (e.g. those involving storage or initial curation) are important for, but not sufficient for, preservation. Clarity on roles informs the need for specialist skills (e.g. disciplinary or by digital object type). Clarity on responsibility is of particular importance when defining workflows within and between organisations. Organisational cooperation and interoperability depend on rights management agreements that are ultimately built on the roles and responsibilities around digital objects.

71. Preservation roles, responsibilities and accountability remain unclear, this includes for data stewards as individuals and for organisations such as libraries, archives and institutional repositories.
72. Preservation roles and skills are not limited to technical manipulation of digital objects: optimal preservation outcomes depend on a conceptual understanding of the data and metadata being cared for.
73. Assigning clear responsibilities within groups, partnerships and across lifecycles makes outcomes, including preservation, more effective and accountable.
74. Assigning roles to individuals with clear responsibility makes outcomes, including preservation, more effective and accountable.
75. @Where responsibility is distributed, accountability should remain clear, including accountability for (meta)data loss or destruction.
76. Lack of clarity and transparency of roles and responsibilities presents a risk to digital objects and their preservation.
77. @Digital object management outcomes, including preservation, should be integrated into a roles and responsibilities framework that integrates all actors and actions.
78. @The roles and responsibilities framework should be aligned with clear process models that meet the needs of different stakeholder communities.
79. Preservation roles must include monitoring the changing needs of communities at the point of reuse. This community watch must be aware of the knowledge base, methodologies and technologies of the user communities.
80. Preservation roles must include monitoring the changing nature of available technologies for the deposit, storage, curation, discovery, access and reuse of data and metadata. This technology watch must continue to meet the needs identified through community watch and be proactive as well as reactive.

81. @Defined roles should be in place to take responsibility from the point of conception to ensure that preservation actions are considered throughout the life cycles.
82. Roles and responsibilities include the early definition of preservation actions in a data management plan²⁴ such as multi-copy/multi-location redundancy, integrity, data protection, digital object design, provenance information, appraisal criteria, desirable retention and preservation periods and future repositories.
83. @Risk analysis approaches should be used to identify when in the lifecycle it is appropriate to take preservation actions. This includes the availability of individual researcher expertise about the digital objects (conception/collection/creation phase) versus the broader expertise and opportunities for economies of scale at the repository phase.
84. @Different roles should use a living data management plan as a key artefact for periodic audit, review and revision.
85. @Policy makers²⁵ must make the development and implementation of digital preservation explicit in policy applicable to all stakeholders and across the lifecycle. They are accountable for periodic review and revision of policy.
86. @Executives must adopt preservation policy into their operational and strategic planning.
87. @Managers must integrate preservation planning into operational management including staffing, funding, service development and procurement.
88. @Practitioners must provide guidance, community and technical monitoring and, where necessary, take preservation actions to ensure optimal preservation outcomes.
89. Technicians develop and maintain the hardware and software infrastructure that supports preservation systems including integrity, data protection, automation and audit.
90. @Data and metadata creators, collectors and reusers, including researchers, should develop the knowledge and skills at a general level and within their own disciplinary and domain area of expertise, so that their actions are preservation-aware.

²⁴ FAIR Forever Use Case. A mechanism to ensure accountability and implementation of preservation in DMPs

²⁵ FAIR Forever Use Case 5, A mechanism for digital preservation policy across institutions within EOSC, would offer valuable resources to help research policy makers with these responsibilities.”

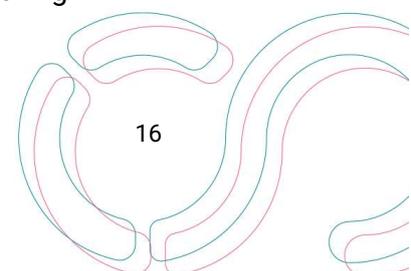


91. @All of the preservation-specific and supporting research data management roles across the data lifecycle require sustained training based on a rich knowledge base of preservation information.
92. @Clear responsibilities must be in place for developing standards and guidance, for communication and for training.
93. Transparency about and analysis of current roles and responsibilities associated with data services functions and activities are necessary inputs into financial calculations related to salaries and funding streams.

Finance & Funders

94. “Researchers, disciplinary and cross-disciplinary research communities, Member States and the EOSC Community are all expecting commitments from each other but lack the support to make commitments of their own”.
95. @Further research and analysis are necessary to support business planning based on qualitative and quantitative risks and benefits ²⁶.
96. @Funders should clarify to all grant holders that the FAIR Principles and the potential need for preservation are full lifecycle issues.
97. @Funders should integrate into their calls the costs required to meet the needs for compliance with the FAIR Principles and any long-term preservation
98. Data management plans are the critical reference point across partners and lifecycles.
99. @Audit pathways are essential for all research outputs.
100. @Data management plans should define obligations and support accountability.
101. @Once identified, critically endangered content that retains value requires investment.
102. @Full lifecycle costs should be assessed, including the unique costs of preservation using methodologies that include the potential costs of inaction.
103. @Identify and support staffing costs for preservation specific roles and responsibilities.
104. Commitment to funding compliance with emerging policies is critical at the European, national and institutional level.

²⁶ FF Use Case A business case factory or service for preservation cost modelling



105. Any decision to store, retain and then undertake initial curation will imply costs (Level C above). Any *additional* costs needed to deliver active preservation (Levels B and A above) are uniquely preservation costs.
106. “Investigations into preservation roles and responsibilities and their associated costs (e.g. case studies, labour market data analysis) are necessary to establish accurate costing parameters for preservation programmes and services provided in EOSC.”

Network of TDRs

107. EOSC and the global digital preservation community would both benefit from active engagement in international networks of trustworthy digital repositories, research performing organisations, and public and private sector data service providers²⁷.
108. Current and future networks of preservation practitioners, including TDRs, can support the development, evaluation, implementation and monitoring of all the recommendations made in this paper and provide a platform for:
- Networking and knowledge exchange ²⁸, improving FAIR-enabling capabilities and trustworthiness.
 - Stakeholder advocacy and engagement ²⁹, acting as a "unified voice" of the repository community.
 - Align with and input to the EOSC ecosystem, including repository landscape monitoring and defining the requirements of repositories and their data and metadata users.
 - Coordinate and develop frameworks for research data repository policies and routines³⁰, such as a strategic framework to achieve baseline certification, to audit data management plans³¹, and identifying preservation pathways for data³².
 - Evaluate FAIR metrics and tools and provide feedback on efforts to align certification requirements with FAIR³³.

²⁷ (FF16) “Establish an ongoing basis for partnership in the digital preservation community, including beyond the research data community”

²⁸ <https://doi.org/10.5281/zenodo.7034315>

²⁹ <https://doi.org/10.5281/zenodo.7034315>

³⁰ <https://doi.org/10.5281/zenodo.7034315>

³¹ (FF10) “Provide strategic framework for audit of data management plans. Adapt workplans to include quality improvement mechanisms where these do not already exist, including DPC Rapid Assessment Model, establishing thereby a strategic framework to achieve baseline certification for primary preservation services, or identifying preservation pathways for data”

³² (FF04) “Adapt workplans to include quality improvement mechanisms where these do not already exist, including DPC Rapid Assessment Model, establishing thereby a strategic framework to achieve baseline certification for primary preservation services, or identifying preservation pathways for data”

³³ <https://doi.org/10.5281/zenodo.7568400>

- f. Identify costs of action versus inaction with respect to high value, critically endangered content³⁴.

Addressing Preservation at European, National and Institutional Level

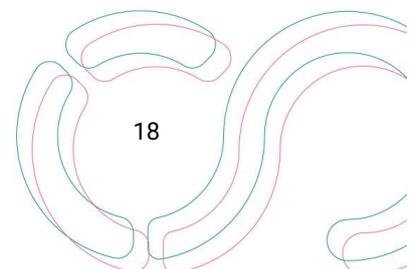
109. @Greater clarity is needed in the EOSC vision and for the application of preservation activity across the lifecycle and at EU, National and institutional level.

Recommendations at European Level

European level influence may act directly on institutions or via the intermediary of national level policies and practices. In all cases actions should be aligned at the relevant European, National or Institutional Level.

110. Addressing preservation is critical to EOSC data, reputation and sustainability. EOSC cannot achieve its goals in the long-term unless they are addressed.
111. @Digital preservation risks and opportunities must be made explicit in the EOSC vision and monitored and addressed as the EOSC evolves.
112. @“EOSC cannot rely on a single generic canon of 'digital preservation practice'. Instead, workflows should leverage large scale infrastructures while remaining faithful to discipline-specific requirements.”
113. @Roles, responsibilities, and accountabilities for preservation in EOSC should be clarified.
114. @EOSC should establish a workplan for policy development and implementation within EOSC services and partners including delivery or support of:
 - a. A dedicated preservation group with an EOSC Board member providing communication and liaison.
 - b. To establish a high-level digital preservation policy across EOSC with defined connection points to national and institutional preservation policies.
 - c. To monitor policy implementation across EOSC partners.
 - d. Define objectives, challenges, and implications for the preservation of research data.
 - e. To act as a point of contact between EOSC and other digital preservation communities within and beyond the research data community.
115. @Support the alignment between the interpretations of FAIR for digital objects and the criteria for data services, including repositories to be FAIR enabling at whatever level of care they provide.
116. @Establish pathways for continuous quality improvement that reflects:

³⁴ (FF14) “Identify costs of action versus inaction with respect to high value, critically endangered content”



- a. The wide range of digital objects and data services striving to engage with EOSC.
 - b. The scenarios where standards compliance and assessment are necessary achievements rather than desirable targets e.g. for the protection of sensitive data.
117. @Support the development of verified business models for data services, including preservation services³⁵.
118. @Support the development of digital rights management standards and mechanisms that support the transition of digital objects and data services' licences to explicit, machine-actionable, scalable interoperability including:
- a. Granular and dynamic digital objects.
 - b. Complex organisational partnerships and outsourcing.
 - c. Full lifecycle research data management, storage, curation and preservation.

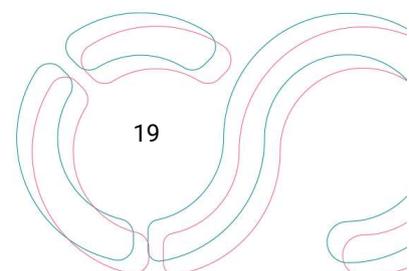
Recommendations at National Level

119. @Actions taken at European level should be adopted, adapted for national needs and support provided for implementation.
120. @When action cannot be taken at European level they should be developed and implemented at national level.

Recommendations at Institutional Level

121. @Support researchers in digital object design, creation, storage, curation and preservation whether through locally provided repositories and data services or via third parties.
122. @Take appropriate responsibility for relevant phases of the research data lifecycle including assurances that digital objects are made FAIR and that services are FAIR enabling and trustworthy.
123. @Support ongoing review and audit of practice across the lifecycle.
124. @Funders should commission repositories to conduct audits, and repositories should undertake these audits.

³⁵ Cf: FAIR Forever Use Case 3: A Business Case Factory or Service for Preservation Cost Modelling within EOSC



Version History

01.00: The basis of this initial draft document is to integrate³⁶ the recommendations of the prior report, FAIR Forever³⁷, and place them in the context of the developments that have taken place since it was published. The developments include the launch of the EOSC-Association and the associated advisory groups and task forces. The LTDP-TF Overview³⁸ provides the Task Force context and feedback to that document will be used to re-scope recommendations as necessary. Together this information and context will be used to develop a first set of recommendations for discussion within and beyond³⁹ the task force. Consultation and iteration will inform feedback related to the MAR and SRIA⁴⁰ with a major iteration of the recommendations expected to be released later in 2023⁴¹.

³⁶ The authors do not and cannot incorporate all elements of this prior work and suggest that it should be read as a reference.

³⁷ FAIR Forever? Long Term Data Preservation Roles and Responsibilities, Final Report
<https://doi.org/10.5281/zenodo.4574234>

³⁸ EOSC Preservation: Overview Discussion Paper <https://doi.org/10.5281/zenodo.7516259>

³⁹ Wider engagement will be guided by the EOSC Association, the work of Task Force members and through coordination efforts such as the FAIR Synchronisation Force from the FAIR IMPACT project.

⁴⁰ <https://eosc.eu/sria-mar>

⁴¹ The exact timeline depends on ongoing discussion within the EOSC Association about the role and timeframes for Task Forces.