

LTDP-TF Recommendations & Assertions

EOSC Association. Long Term Data Preservation Task Force
February 2024

EOSC Association AISBL

Rue du Luxembourg 3, BE-1000 Brussels, Belgium
+32 2 537 73 18 | info@eosc.eu | www.eosc.eu
Reg. number: 0755 723 931 | VAT number: BE0755 723 931

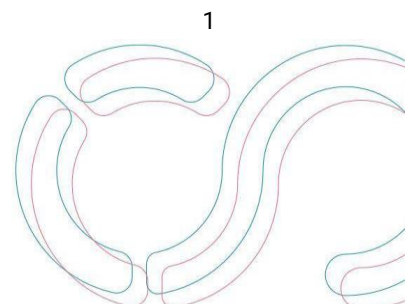
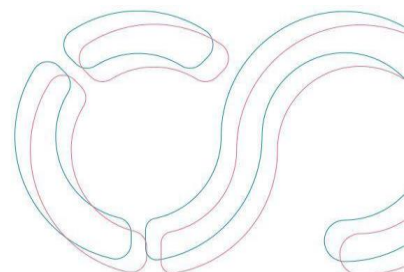


Table of contents

INTRODUCTION	3
STRUCTURE AND PRESENTATION OF THIS PAPER.....	4
A VISION FOR OPTIMAL PRESERVATION OF FAIR DIGITAL OBJECTS WITHIN EOSC	5
ASSERTIONS AND RECOMMENDATIONS FOR OPTIMAL PRESERVATION OUTCOMES	5
REPOSITORIES AND THE DIGITAL OBJECT LIFECYCLE	5
CURATION & PRESERVATION LEVELS	7
FAIR + TIME	8
PRESERVATION SYSTEMS, ACTIONS & OUTCOMES.....	8
TRUSTWORTHY DIGITAL REPOSITORIES (TDR) AND OTHER DATA AND METADATA SERVICES	9
STANDARDS, ASSESSMENT, CERTIFICATION & ALIGNMENT	11
OUTCOMES, JUDGEMENT AND GATEKEEPING	12
RETENTION, APPRAISAL & RE-APPRAISAL.....	13
TRANSPARENCY OF SERVICES, ARTEFACTS AND LEVELS OF CARE FOR OBJECTS.....	13
GENERIC VERSUS DOMAIN, DISCIPLINE AND OBJECT TYPE SPECIFIC ISSUES.....	14
ROLES AND RESPONSIBILITIES	14
FINANCE & FUNDERS.....	16
NETWORK OF TDRS	18
ADDRESSING PRESERVATION AT EUROPEAN, NATIONAL AND INSTITUTIONAL LEVEL	19
RECOMMENDATIONS AT EUROPEAN LEVEL	19
RECOMMENDATIONS AT NATIONAL LEVEL	20
RECOMMENDATIONS AT INSTITUTIONAL LEVEL	20
VERSION HISTORY	21



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 and input to how a European network of FAIR⁴-enabling Trustworthy Data Repositories can align to the vision of EOSC⁵. A paper on the frame of reference used to guide the task force approach was presented⁶ at the PV2023 conference⁷. This paper updates the initial assertions and recommendations of the task force based on feedback received. The decision was taken to minimise prose and maximise the presentation of brief assertions and recommendations. Numerous scoping assertions are included because the work of the task force revealed varying interpretations of key research infrastructure concepts, including preservation. This paper informs the final report from this iteration of the LTDP-TF. These assertions and recommendations remain valid for future reference by a range of actors.

The work of the task force benefits from numerous prior efforts in research data management, curation and preservation, particularly the FAIR Forever report⁸, 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 overview discussion paper⁹ made it clear that the LTDP-TF perspective acknowledges the importance of FAIR data and trustworthy repositories as central to the EOSC vision¹⁰, but also notes that a full digital object 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, when necessary, take preservation *actions* on digital objects (data 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.

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 Wilkinson, M., Dumontier, M., Aalbersberg, I. et al. The FAIR Guiding Principles for scientific data management and stewardship. *Sci Data* 3, 160018 (2016). <https://doi.org/10.1038/sdata.2016.18>

⁵ EOSC Preservation: Overview Discussion Paper. Zenodo. <https://doi.org/10.5281/zenodo.7516259>

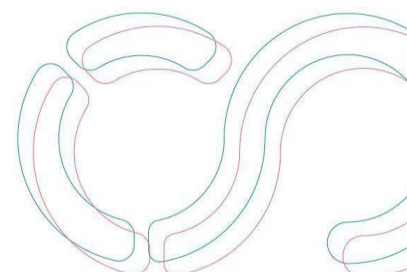
⁶ L'Hours, H., & Wyns, R. (2023, December 1). Preservation in the context of EOSC. Sustainable repositories curating digital objects from a long term FAIR enabling perspective. <https://doi.org/10.5281/zenodo.10020725>

⁷ Preservation in the context of EOSC. Sustainable repositories curating digital objects from a long-term FAIR enabling perspective <https://indi.to/dpnjt>

⁸ FAIR Forever? Long Term Data Preservation Roles and Responsibilities, Final Report <https://doi.org/10.5281/zenodo.4574234>

⁹ EOSC Preservation: Overview Discussion Paper <https://doi.org/10.5281/zenodo.7516259>

¹⁰ [https://www.eosc.eu/sites/default/files/2023-05/EOSC-A_GA%236_EOSC%20slide%20deck%20\(May%202023\).pdf](https://www.eosc.eu/sites/default/files/2023-05/EOSC-A_GA%236_EOSC%20slide%20deck%20(May%202023).pdf)



While the task force notes the importance of agreeing on standards and developing metrics and tests for assessment, it is important to recognise that repositories 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 repositories 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 Roadmap¹¹ that have already resulted in funding 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 digital objects 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, storage¹² 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 value¹³ we require the resources and skills to deliver preservation systems, actions and outcomes. These depend on sustainable repositories with a long term remit to monitor drivers for change including community needs¹⁴ and technical evolution¹⁵. 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. Digital objects 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 Presentation of this Paper

- The assertions and recommendations are presented under a sequence of headings.

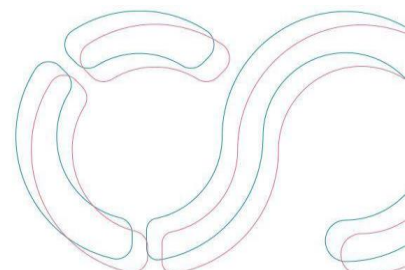
¹¹ <https://eosc.eu/sria-mar>

¹² i.e. common criteria on minimum acceptable integrity and multi-copy, multi-site redundancy.

¹³ See assertion 10, below.

¹⁴ including new research interests, methodologies, and semantic artefacts such as ontologies.

¹⁵ Including file format risks, new software, linked open data and emulation



- In the body of this document text that is taken directly from the FAIR Forever report¹⁶ is presented in quotes, though numerous other items reflect that report's recommendations.
- Recommendation bullets are preceded by an @ symbol¹⁷.
- Recommendations that were marked as priorities by those providing feedback during the consultation are presented in bold and marked with an asterisk*.
- 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 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

The LTDP-TF 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 active 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 their funders, depositors and users.

Achieving this vision is at the core of the task force's recommendations.

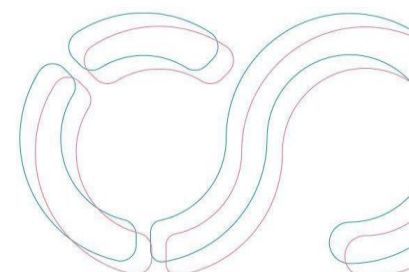
Assertions and Recommendations for Optimal Preservation Outcomes

Repositories and the Digital Object Lifecycle

Desirable characteristics of digital objects, including their FAIRness, at the point of re-use depend on a range of storage, curation, active 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.

¹⁶ FAIR Forever? Long Term Data Preservation Roles and Responsibilities, Final Report <https://doi.org/10.5281/zenodo.4574234>

¹⁷ '@' is selected for accessibility reasons as this symbol is commonly expressed by screen readers.

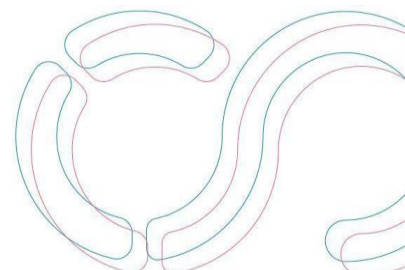


1. Digital objects are constructs of data and metadata including, but not limited to, research outputs such as datasets, and software.
2. Numerous types of digital objects exist. Some characteristics are applicable to many types of objects, others are more specialised.
3. In addition to digital objects 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 force¹⁸.
5. Projects¹⁹ collect or create digital objects of relevance to research.
6. Digital objects must be appropriately cared for during projects.
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 the operational services required for preservation and maintaining FAIRness for the longer term.
9. Digital objects can retain their value beyond the lifetime of the project that collected 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 repository may support the care of digital objects' data and metadata.
12. Different data and metadata services²⁰ are required to support research across the full digital object lifecycle.

¹⁸ 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 objects.

¹⁹ The term 'project' is used broadly to apply to activities (whether academic, commercial or public) in the pre-repository phase of the lifecycle that conceive, collect and/or create digital objects of interest to researchers of all kinds.

²⁰ References to data and metadata 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.



13. Any digital object service holding data or metadata at any level of care (retention, curation or preservation), and at any stage of the lifecycle could define itself as a repository.
14. @Additional work is needed to define different types of digital object systems that offer data and metadata services, including the activities and functions 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 should be defined as a foundation for all levels of retention, curation and preservation services.
17. Repositories benefit from being transparent on their current level of storage, curation and preservation practice, as this increases trust by the user and funders.

Curation & Preservation Levels

18. This paper defines the different levels of care that a repository may apply to digital objects as follows²¹:

Z. Level Zero. Content distributed as deposited. Unattended deposit-storage-access. Data content and supporting metadata are stored for a given time period, or indefinitely. This may include multiple copies and monitoring of bitstreams for integrity. Data content and supporting metadata are distributed to users exactly as they are provided by depositors. Beyond these measures, there is no appraisal, 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, metadata elements, and compliance with legal and ethical norms.²²

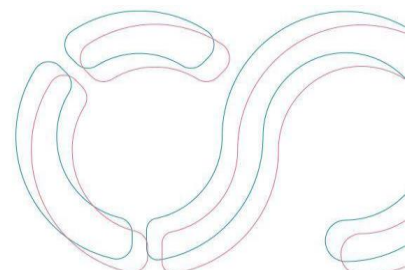
C. Initial Curation

In addition to Level D above, if these criteria are not met the digital objects are curated by the repository to meet the defined criteria. This initial curation for access and use may include, e.g., the correction or enhancement of metadata and/or data content, or the creation of dissemination formats.

B. Logical-Technical Curation

²¹ CoreTrustSeal Standards and Certification Board. (2023). Curation & Preservation Levels: CoreTrustSeal Discussion Paper (v02.00). Zenodo. <https://doi.org/10.5281/zenodo.8083359>

²² The actions that follow these checks are determined by the repository. For example, a repository may choose to return digital objects that do not meet the deposit criteria to the depositor, or to ingest the digital object and document non-compliance, or to undertake initial curation to ensure compliance.



In addition to D and/or C above the repository takes long-term responsibility for ensuring that the data and metadata can be rendered as required by the designated community.

This entails the responsibility for updating hard- and software environments, archival and dissemination formats of digital objects, and metadata in response to the threat of technological obsolescence and/or to accommodate changing needs of the Designated Community.

A. Conceptual preservation for understanding and reuse

In addition to B above, the repository takes long-term responsibility that the data content and metadata can be independently understood by the designated community. This entails the responsibility for updating the content of metadata elements and other semantic artefacts such as controlled vocabularies and ontologies if necessary. It may include responsibility for editing the structure and content of deposited data, for example in response to changes in legal regulations.

- 19. @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*.**

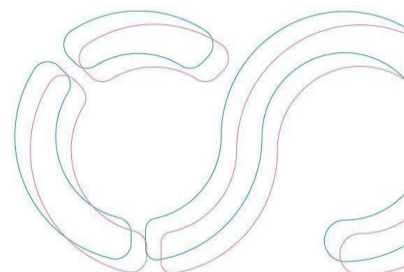
FAIR + Time

As time passes, technologies and the needs of digital object (re)users evolve. This may result in digital objects becoming less FAIR over time. Transparent FAIR-enabling repositories 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 repositories should be defined.
- 24. @FAIR-enabling practices undertaken by repositories should be made transparent to users and funders to increase trust in services*.**

Preservation Systems, Actions & Outcomes

To achieve preservation *outcomes* (long-term access to and use of digital objects that maintain desirable 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).

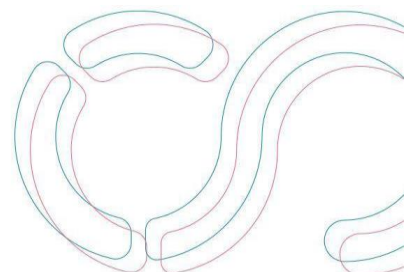
25. Preservation systems (E.g. organisations, partnerships, archives, repositories, libraries, galleries, museums) provide a sustainable organisational infrastructure to monitor the evolution of technical infrastructures and the needs of user communities and to undertake preservation actions as necessary.
26. Preservation actions include any action on a digital objects' data (e.g. format updates), metadata (e.g. schema updates) or on the environment provided to interact with the digital object (e.g. emulation) to ensure the object retains desirable characteristics (e.g. FAIRness).
27. Preservation outcomes are demonstrated by an object maintaining desirable characteristics (e.g. FAIRness) over time.
28. Many repository activities and functions are necessary for preservation, but preservation also requires additional activities with associated roles and costs.
29. @ Unique preservation activities and functions should be defined alongside activities and functions that apply to all repositories.

Trustworthy Digital Repositories (TDR) and other Data and Metadata Services

Trustworthy Digital Repositories (TDR) comply with a set of organisational infrastructures, digital object management, technology and security requirements. TDR standards, such as the CoreTrustSeal²³ currently require that repositories provide active preservation (Levels B and A above) to enable 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 should 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).

²³ <https://www.coretrustseal.org/about/>



31. The CoreTrustSeal is a not-for-profit foundation, developed in response to an RDA²⁴ 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 papers²⁵ and groups²⁶ that the CoreTrustSeal provides an appropriate mechanism to define core expectations of TDRs and an exemplar for offering assessment and certification services*.**
33. With the exception of the requirements for offering preservation services (CoreTrustSeal Requirement 09) and considering reuse (CoreTrustSeal Requirement 13) the CoreTrustSeal²⁷ is relevant to all repositories²⁸.
34. @Other ongoing work to define repository and data service characteristics and expectations exist and should be encouraged and supported²⁹.
- 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' organisational infrastructure, digital object management, technology and security, as well as the roadmap towards improvement of TRUST³⁰ 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.

²⁴ <https://rd-alliance.org/>

²⁵ 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>

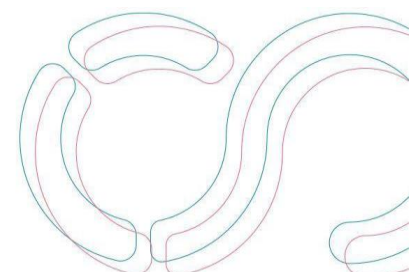
²⁶ European Commission, Directorate-General for Research and Innovation, Jones, S., Aronsen, J., Beyan, O. et al., Recommendations on certifying services required to enable FAIR within EOSC – , Genova, F.(editor), Publications Office, 2021, <https://data.europa.eu/doi/10.2777/127253>

²⁷ CoreTrustSeal Requirements 2023-2025 <https://doi.org/10.5281/zenodo.7051012>

²⁸ Review of CoreTrustSeal Applicability to non-Preservation (Trustworthy) Data Services <https://doi.org/10.5281/zenodo.7646134>

²⁹ Repository & (meta)data Services Functions & Activities: Crosswalk <https://doi.org/10.5281/zenodo.7690658>

³⁰ 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>



- 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 on them where necessary*.
- 39. @Additional work is needed to define different types of digital object systems, the activities and functions 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*.
- 41. @Technical repository service providers' (storage providers, ARCHIVER³¹ etc) portfolio of service offerings should be clear and comparable for client end-users^{32*}.

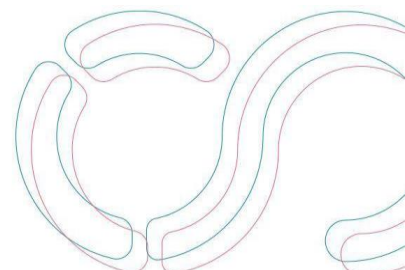
Standards, Assessment, Certification & Alignment

Defining standards for objects and repositories 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.

³¹ <https://archiver-project.eu/>

³² E.g. via efforts such as <https://www.rd-alliance.org/rda-tiger>

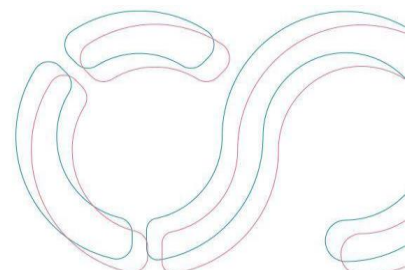


48. @ Standards and guidance should be developed, coordinated and maintained to provide full lifecycle information on preservation to researchers and preservation practitioners.

Outcomes, Judgement and Gatekeeping

Digital objects and repositories at all levels of maturity are on a journey towards FAIR and trustworthiness. Negative judgements about repositories or the objects they contain, and gatekeeping (e.g. exclusion from the EOSC) as a result of the outcomes of assessments should be avoided in favour of investment and assistance. Transparency of practice and assessments should be used as a tool for developing roadmaps for improvement.

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 compliance with 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 assistance and 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 for improvement*.**
55. @No repository or digital object should be unnecessarily excluded from any part of EOSC.
- 56. @FAIRness and trustworthiness must continue to be a supported journey for all parties*.**



Retention, Appraisal & Re-Appraisal

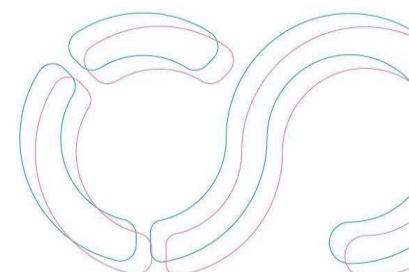
Any activities that retain, identify and evaluate a greater proportion of relevant digital objects are desirable. Some of these objects may be appraised as having sufficient value to be retained, curated or actively preserved. Reappraisal over time may change the level of care received by a digital object, or lead to a decision to delete the object. As for 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 digital objects vary widely, and so there are different perspectives and insights into the kinds of preservation action that may be required.
60. Knowing which characteristics matter, and what metadata they might require is a subject specialist skill that may be best, and most cost-efficiently 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. Transparency can be designed to mitigate any risks to information security or competitive advantage.
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 repository should be transparent about the levels of care provided*.**



- 66. **@Each object should have a clear level of care associated with the repositories 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 metadata about levels of care 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

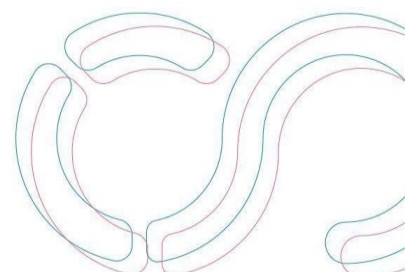
All of the assertions and recommendations above, and the roles and actors described below are specified at a high level due to the scope and timescales of the task force. Further exploration of the specific needs implied by discipline, domain and object type is required.

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

Roles and Responsibilities

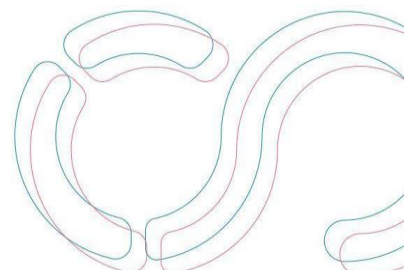
Roles and responsibilities across partners and the lifecycle are not always well defined and many of them (e.g. those involving storage or initial curation) are important but not sufficient for preservation. Clarity on roles informs the need for specialist skills (e.g. disciplinary or by digital object type). Clear 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 also depend on a conceptual understanding of the digital objects 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 for digital objects' 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. @From the point of conception of a digital object defined roles and responsibilities should be in place to ensure that preservation actions are considered throughout the lifecycle.
82. Roles and responsibilities include the early definition of preservation actions in a data management plan³³. These include multi-copy/multi-location redundancy, integrity, data protection, digital object design, provenance information, appraisal criteria, desirable retention and preservation periods, and target 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 (during the conception/collection/creation phase) versus the broader expertise and opportunities for economies of scale during the repository phase.
84. @Different roles should use a living data management plan as a key artefact for periodic audit, review and revision.

³³ FAIR Forever Use Case. A mechanism to ensure accountability and implementation of preservation in DMPs.

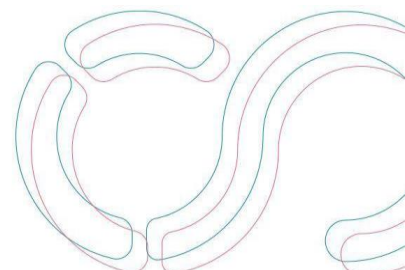


85. **@Policy makers³⁴ must make the development and implementation of active preservation explicit in policies applicable to all stakeholders and across the lifecycle. They are accountable for periodic review and revision of policy*.**
86. @Executives must adopt a preservation policy into their operational 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 technology 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. **@Digital objects' 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 early in the lifecycle enable preservation*.**
91. **@All of the preservation-specific and supporting research data management roles across the data lifecycle require sustained training*.**
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 repository activities and functions are necessary inputs into financial calculations related to salaries and funding streams.

Finance & Funders

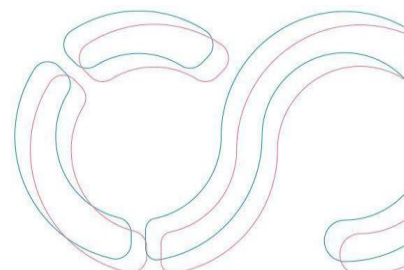
94. "Researchers, disciplinary and cross-disciplinary research communities, Member States and the EOSC Community are all expecting commitments from each other but lack the support to make commitments of their own".

³⁴ 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".



95. @Further research and analysis are necessary to support business planning based on qualitative and quantitative risks and benefits³⁵.
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 between partners and throughout the lifecycle.
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 sustainable 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.
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".

³⁵ FF Use Case A business case factory or service for preservation cost modelling.



Network of TDRs

107. Active engagement between EOSC, international networks of trustworthy digital repositories, research performing organisations, and public and private sector data service providers would generate mutual benefits³⁶.
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:**
- a. Networking and knowledge exchange³⁷, improving FAIR-enabling capabilities and trustworthiness.
 - b. Stakeholder advocacy and engagement³⁸, acting as a "unified voice" of the repository community.
 - c. Aligning with and providing input to the EOSC ecosystem, including repository landscape monitoring and defining the requirements of repositories and their users.
 - d. Coordination and development of frameworks for research data repository policies and routines³⁹, such as a strategic framework to achieve baseline certification, to audit data management plans⁴⁰, and to identify preservation pathways for data⁴¹.
 - e. Evaluation of FAIR metrics and tools and provision of feedback on efforts to align certification requirements with FAIR⁴².
 - f. Identification of costs of action versus inaction with respect to high value, critically endangered content⁴³.

³⁶ (FF16) "Establish an ongoing basis for partnership in the digital preservation community, including beyond the research data community".

³⁷ Towards a European network of FAIR-enabling Trustworthy Digital Repositories (TDRs) - A Working Paper <https://doi.org/10.5281/zenodo.7034315>

³⁸ <https://doi.org/10.5281/zenodo.7034315>

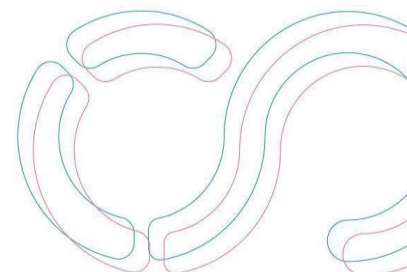
³⁹ <https://doi.org/10.5281/zenodo.7034315>

⁴⁰ (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"

⁴¹ (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"

⁴² How a European network of FAIR-enabling Trustworthy Data Repositories can align to the vision of EOSC <https://doi.org/10.5281/zenodo.7568400>

⁴³ (FF14) "Identify costs of action versus inaction with respect to high value, critically endangered content"



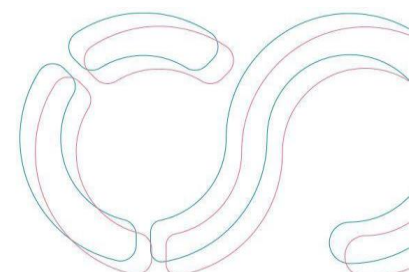
Addressing Preservation at European, National and Institutional Level

109. **@Greater clarity is needed in the EOSC vision 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. EOSC cannot achieve its goals in the long-term unless preservation is addressed. Preservation is critical to EOSC digital objects, reputation and sustainability.
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. A high-level digital preservation policy across EOSC with defined connection points to national and institutional preservation policies.
 - c. The monitoring of policy implementation across EOSC partners.
 - d. A definition of objectives, challenges, and implications for the preservation of digital objects.
 - e. 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 and metadata services, including repositories, to be FAIR-enabling at whatever level of care they provide*.**
116. **@Establish pathways for continuous quality improvement that reflect*:**
- a. The wide range of digital objects and repositories 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 repositories, including preservation services⁴⁴.
118. @Support the development of digital rights management standards and mechanisms that enable the transition of digital objects and repositories' licences towards 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.

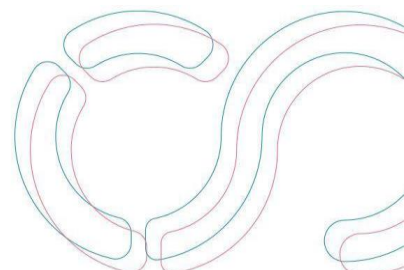
Recommendations at National Level

119. @Actions taken at European level should be adopted, adapted for national needs, and support provided for implementation.
120. @When actions 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 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 infrastructure and practice across the lifecycle.
124. @Funders should commission repositories to conduct audits, and repositories should undertake these audits.

⁴⁴ Cf: FAIR Forever Use Case 3: A Business Case Factory or Service for Preservation Cost Modelling within EOSC.



Version History

02.00: This revised version is based on community feedback after over 700 downloads of version 1.0. No specific objections were received to any assertion or recommendation. Some feedback focussed either on areas that were beyond the scope of the task force, or on points of clarification. The revised text focuses less on the wider (and less well defined) data and metadata services, in favour of referencing repositories. These changes are to increase clarity and do not reduce the need to define and support a wider range of service types. References to 'data and metadata' were perceived by some respondents to reduce the applicability of the guidelines to additional data object types, such as software. To reduce this perception a number of references to 'data and metadata' have been changed to 'digital objects'.

01.00: The basis of this initial draft document is to integrate⁴⁵ the recommendations of the prior report, FAIR Forever⁴⁶, 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 Overview⁴⁷ 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 beyond⁴⁸ the task force. Consultation and iteration will inform feedback related to the MAR and SRIA⁴⁹ with a major iteration of the recommendations expected to be released later in 2023⁵⁰.

⁴⁵ The authors do not and cannot incorporate all elements of this prior work and suggest that it should be read as a reference.

⁴⁶ FAIR Forever? Long Term Data Preservation Roles and Responsibilities, Final Report <https://doi.org/10.5281/zenodo.4574234>

⁴⁷ EOSC Preservation: Overview Discussion Paper <https://doi.org/10.5281/zenodo.7516259>

⁴⁸ 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.

⁴⁹ <https://eosc.eu/sria-mar>

⁵⁰ The exact timeline depends on ongoing discussion within the EOSC Association about the role and timeframes for Task Forces.

