Appendix 5: Analyses of Frameworks

This appendix contains the input for the “Roadmap for Preservation and Curation in the SSH” deliverable (D4.1). Before this roadmap could be created an extensive analysis of different frameworks was done. In this analysis the frameworks were compared with each other by identifying how specific categories were implemented in these frameworks. To define the categories the OAIS model was used.

Content

1. Introduction

This report will define the baseline for data curation and long term data preservation by analysing the current state of data management by making use of existing investigation results and by studying relevant reports. The goal is to establish an appropriate state-of-the-art description and guidelines and a common roadmap for Social Sciences and Humanities (SSH) institutions and researchers with respect to proper rules for data management and to determine the required services and measures.

Since the SSH represent a broad variety of different research fields and practices, it is impossible to create one roadmap which would cover all specific situations. This roadmap should therefore be seen as a general set of recommendations. Not all recommendations will suite all, and not all recommendations thinkable for a specific situation are included.

For this roadmap we first conducted an inventory of the commonly accepted sets of guidelines and framework models. The focus was restricted to reports which were purposed to be generally applicable for different research fields. Practical reports of a single organisation, such as best practices, were not included in this analysis. The sources which were included are:

OAIS

> OAIS is a technical recommendation of practice for providing long term preservation of digital information. An organization can establish an OAIS archive by adhering to the recommendations and standards stated by OAIS. The term 'open' in OAIS indicates, that these recommendations and standards are not final and are subject to discussion in open forums, it does not imply that the access to the particular archive is unrestricted.

TRAC-TDR

> The Audit and Certification of Trustworthy Digital Repositories, published in 2011, is the CCDS (The Consultative Committee for Space Data Systems (CCSDS) recommended practice. It is based on the Trustworthy Repositories Audit & Certification: Criteria and Checklist (TRAC) and builds on the RLG-OCLC Report “Trusted Digital Repositories: Attributes and Responsibilities” (TDR).
>
> > “OAIS included a Roadmap for follow-on standards which included ‘standard(s) for accreditation of archives’. It was agreed that RLG and National Archives and Records Administration (NARA) would take this topic forward and the later published the TRAC document which combined ideas from OAIS and Trusted Digital Repositories: Attributes and Responsibilities (TDR—reference). The current document follows on from TRAC in order to produce an ISO standard.”
>
> The document follows on from TRAC in order to produce an ISO standard. The document is meant to be used as a tool for (self-) audit and certification processes for digital repositories which seek to assess and demonstrate their trustworthiness. It provides a detailed specification of criteria by which digital repositories shall be audited. It presents high-level criteria to assess the organizational infrastructure, digital object management, and infrastructures and security risk management in digital repositories.

DRAMBORA

Developed jointly by the Digital Curation Centre (DCC) and Digital Preservation Europe (DPE), DRAMBORA (Digital Repository Audit Method Based on Risk Assessment) represents the main intellectual outcome of a period of pilot repository audits undertaken by the DCC throughout 2006 and 2007. It presents a methodology for self-assessment, encouraging organisations to establish a comprehensive self-awareness of their objectives, activities and assets before identifying, assessing and managing the risks implicit within their organisation. As opposed to other data curation approaches, DRAMBORA focuses on self-assessment and risk-management instead of setting rules for data ingestion.

DCC Cycle

The DCC Curation Lifecycle Model provides a graphical, high-level overview of the stages required for successful curation and preservation of data from initial conceptualisation or receipt through the iterative curation cycle.

The model van be used to plan activities within an organisation or consortium to ensure that all of the necessary steps in the curation lifecycle are covered.

DPC

The Handbook was first compiled by Neil Beagrie and Maggie Jones and is now maintained and updated by the Digital Preservation Coalition (DPC) in collaboration with the National Library of Australia and the Preserving Access to Digital Information (PADI) Gateway. The handbook is intended to provide a bridge between broad, high level overviews and explicit, detailed guidelines applicable to the needs of a specific institution. The strategic overviews are intended to link to operational activities in order to reinforce the need to develop practical procedures grounded firmly in the business mission of the institution. Some of the issues and segments are written from a UK perspective and legislation in this area will vary from country to country. However, many of the broader suggestions have wider implications.

DPE Platter

The PLATTER is a framework, which provides a basis for a digital repository to plan the development of its goals, objectives and performance targets over the course of its lifetime in a manner which will contribute to the repository establishing trusted status amongst its stakeholders. PLATTER cannot be perceived as an audit or certification tool itself, but rather a framework that will allow new repositories to incorporate the goal of achieving trust into their planning from an early stage. The PLATTER process is centred around a group of Strategic Objective Plans (SOPs) through which a repository specifies its current objectives, targets, or key

performance indicators in those areas which have been identified as central to the process of establishing trust.

Nestor

The work of the Nestor working group "Trusted repositories" has the main focus on the challenge of earning trust and communicating this, be it to fulfill a legal mandate or simply to survive in the market. It identifies criteria which permit the trustworthiness of a digital repository to be evaluated, both at the organizational and technical levels.

DSA

The Data Seal of Approval (DSA) is a set of 16 guidelines for data repositories or data archives. It regards three stakeholders of data: the data producer, the data consumer and the data archive. The main responsibility belongs to the data archive. The guidelines are specified in the form of questions, to which applicants should response. All information an applicant provides during the assessment should be available online. The DSA can be seen as the basis level certification for a data repository.

Digiplanet

The Blue Ribbon Task Force on Sustainable Digital Preservation and Access was created in late 2007 to address, among other, the issues of availability, long-term preservation and access and economic sustainability of digital information. In 2010, the Taskforce published its Final Report, called "Sustainable Economics for a Digital Planet: Ensuring Long-Term Access to Digital Information." The report provides general principles and actions to support long-term economic sustainability; context-specific recommendations tailored to specific scenarios analyzed in the report. It is based on an investigation into preservation in the United States and the United Kingdom.

The recommendations of the abovementioned sources are analysed in chapter three based on a topic related framework. The most important recommendations, preferably shared by several sources, are presented in an overview.

2. Terminology

In this section some of the terms used in this document will be explained.

Archival Information Package (AIP)

AIP, as defined in OAIS reference model (CCSDS 650.0-B-1 Reference Model for an Open Archival Information System (OAIS)), is an information package that is used to transmit archival objects into a digital archival system, store the objects within the system, and transmit objects from the system. An AIP contains both metadata that describes the structure and content of an archived essence and the actual essence itself. It consists of multiple data files that hold either a logically or physically packaged

entity. The implementation of AIP can vary from archive to another archive; it specifies, however, a container that contains all the necessary information to allow long term preservation and access to archival holdings. The metadata model of OAIS is based on METS specifications.

Source: http://www.iasa-web.org/tc04/archival-information-package-aip

Best practice

a practice which is most appropriate under the circumstances, esp. as considered acceptable or regulated in business; a technique or methodology that, through experience and research, has reliably led to a desired or optimum result

Source: http://dictionary.reference.com

Data Consumer

The person or organisation that extracts data from a data archive.

(Data) Curation

Digital curation is the selection, preservation, maintenance, collection and archiving of digital assets. Digital curation is generally referred to the process of establishing and developing long term repositories of digital assets for current and future reference by researchers, scientists, historians, and scholars. Enterprises are starting to utilize digital curation to improve the quality of information and data within their operational and strategic processes.

Source: http://en.wikipedia.org/wiki/Digital_curation

Data Producer

The person or organisation that creates data to supply it to an archive.

Data Archive

The person or organisation that stores data created by data producers to make it available for data consumers.

Framework

A conceptual framework is used in research to outline possible courses of action or to present a preferred approach to an idea or thought. Conceptual frameworks (theoretical frameworks) are a type of intermediate theory that attempt to connect to all aspects of inquiry (e.g., problem definition, purpose, literature review, methodology, data collection and analysis). Conceptual frameworks can act like maps that give coherence to empirical inquiry. Because conceptual frameworks are potentially so close to empirical inquiry, they take different forms depending upon the research question or problem.

Source: http://en.wikipedia.org/wiki/Conceptual_framework

Guideline

A principle put forward to set standards or determine a course of action

Source: http://dictionary.reference.com

Metadata

A set of data that describes and gives information about other data.

Source: http://oxforddictionaries.com/definition/english/metadata

Policy

a course or principle of action adopted or proposed by an organization or individual

Source: http://oxforddictionaries.com/definition/english/policy

Preservation

Digital preservation can be understood as the series of managed activities necessary to ensure continued access to digital materials for as long as necessary. It combines policies, strategies and actions to ensure access to reformatted and born digital content regardless of the challenges of media failure and technological change. The goal of digital preservation is the accurate rendering of authenticated content over time.

Source: http://en.wikipedia.org/wiki/Digital_preservation

Reference model

A reference model in systems, enterprise, and software engineering is an abstract framework or domain-specific ontology consisting of an interlinked set of clearly defined concepts produced by an expert or body of experts in order to encourage clear communication. A reference model can represent the component parts of any consistent idea, from business functions to system components, as long as it represents a complete set. This frame of reference can then be used to communicate ideas clearly among members of the same community.

Reference models are often illustrated as a set of concepts with some indication of the relationships between the concepts.

Source: http://en.wikipedia.org/wiki/Reference_model

Submission Information Package (SIP)

The SIP is an Information Package that is delivered to the repository and digital storage system for ingest.

Source: http://www.iasa-web.org/tc04/submission-information-package-sip

Stakeholder

A person with an interest or concern in something, especially a business.

Source: http://oxforddictionaries.com/definition/english/stakeholder

3. Summary of comparisons

The complete comparison of the different frameworks is included in the appendix of this document. In this chapter we try to summarize the different analyses.

For almost all topics the sources emphasize the importance of following generally acknowledged standards and following well-established, documented procedures. One should define the roles, responsibilities of and expectations for each party involved with the different phases of data-archiving. For agreements between different parties, such as Data Producer and Data-Archive, or Data-Archive and Data Consumer, the Data-Archive should use explicit contracts. When archiving and disseminating the data, one should consider the different needs of data, e.g. privacy delicate data should be approached differently than other data. Acting transparently and communicating the procedures to the designated community of the Data-Archive are also emphasized.

Table of contents for the different sources:

B DRAMBORA

C DCC Cycle

D DPC

E DPE Platter

N Nestor

O OAIS

P Digiplanet

S DSA

T TRAC TDR

3.1 Organizational Framework

3.1.1 Purpose and Requirements

Identify the mandate and responsibilities of the data-archive and communicate these centrally.

Recommendation	*Found in*
The repository shall have identified and established the duties that it needs to perform and shall have appointed staff with adequate skills and experience to fulfil these duties.	*T*
Determine the general functional type of the repository: *What is the source of the repository's mandate? Is the Repository for profit or non-profit? Does the Repository receive a significant proportion of its material from a legally mandated source? What is the operational status of the repository (not yet running, running but still under development, mature)*	*E*
DRAMBORA enables archives and repositories to better fulfil their responsibilities and achieve their strategic goals by: identifying the strengths and weaknesses of their digital repository; and assisting them to plan effectively to mitigate these risks.	*B*

3.1.1.1 Scope and objectives

A data repository should have an explicit mission statement and strategy on digital archiving, long-term preservation and data access. It should define its designated community and the type of material in the repository. The related policies and procedures should be documented.

Recommendation	*Found in*
Make the 'ingest' process as straight-forward as possible and provide support and guidance wherever you can; automate processes if you can.	*C*
Don't be afraid to be critical when reviewing 'best practice' and recommended approaches. They might work for the specific scenario for which they were created but not for you. Do you know the criteria used to rate things like 'preferred' formats?	*C*
The repository shall have a mission statement that reflects a commitment to the preservation of, long term retention of, management of, and access to digital information. *An explicit mission in digital archiving which is documented in a mission statement*	*S T*
The repository shall have defined its Designated Community and	*T N*

associated knowledge base(s) and shall have these definitions appropriately accessible. *The digital repository has defined its designated community/communities*	
A document which outlines the way in which the mission statement is implemented	*S*
A continuation plan for digital assets	*S*
The digital repository has defined its goals	*N*
The digital repository has developed criteria for the selection of its digital objects	*N*
Define the overall scale of the repository, whether expressed in human, technical, or financial terms: *What is the amount of digital material you expect to archive? How many distinct digital objects do you expect to archive? How many fulltime-equivalent staff does the Repository expect to employ? How many distinct end-users are expected to access material in the Repository over the course of a calendar year?*	*E*
Is the majority of the material in the Repository simple, moderately complex, or highly complex?	*E*
How specialized is the data in the Repository (low, medium or high)?	*E*
How sensitive is the most sensitive material in the Repository (low, medium, high)?	*E*
In which of the three access classes (open, restricted, closed) does the Repository have significant holdings?	*E*
The repository shall have documented policies and procedures for preserving its AIPs, they should never be deleted unless allowed as part of an approved policy, there should be no ad-hoc deletions.	*O*
The producers and consumers communities should be provided with submission and dissemination standards, policies and procedures to support the preservation objectives of the OAIS.	*O*
define a coherent global strategy from the outset, ensuring that everyone concerned is involved	*D*
a clear focus should be on the end purpose of the archiving process which is to serve the consumers or "designated communities" of current and future users	*D*
Key stakeholders and decision-makers should be motivated to contribute to the medium to long term preservation of digital materials. These key stakeholders include the producer, the rights holder, the repository and the consumer, who each may or may not be the same entity depending on the organisation.	*D*
The repository shall negotiate and accept appropriate information from information producers, obtain sufficient control of the information, determine which communities are the Designated Communities, define the information provided accordingly, hence define a knowledge base, ensure that the information to be preserved is independently understandable to the designated community without the need of experts, follow documented policies and procedures to prevent data loss by deletion or	*O*

any other means, disseminate information as copies to the designated communities.	
[...] auditors must identify the repository's mandate, which, it is anticipated, will be described in an organisational mission statement or enacting documentation. The subsequent task requires auditors to identify, within that mandate, each organisational goal and objective relevant to the repository.	*B*

3.1.1.2 Collection policy

Maintain a data collection policy. Specify data acceptation criteria such as the type of data that are preserved. Require information on data collection methods, e.g. a methodology report. Provide information on Institutional Review Boards (IRBs) or other official approval for data collection if available. Use deposit agreements including the scope of material, delivery form, file formats, accompanying metadata. Actively follow a monitoring system, including quality control checks, to control on this.

Recommendation	*Found in*
Return data which fails validation procedures for further appraisal and re-selection.	*C*
The repository shall have a Collection Policy or other document that specifies the type of information it will preserve, retain, manage, and provide access to.	*T*
Provide information on Institutional Review Boards (IRBs) or other official approval for data collection	*S*
Information on data collection method: e.g. a methodology report	*S*
Solid reputation of data producer	*S*
Which of the three acquisition strategies (push, pull, self-creation) account for a significant portion of the total material in the Repository?	*E*
Acquire relevant material: specify quantitative targets for the material to be acquired by the repository	*E*
Negotiate deposit agreements: Scope of material, Delivery form, File formats, Accompanying metadata, etc.	*E*
Obtain Physical Control of Materials	*E*
It is essential that a repository have in place a monitoring system to determine that the required material is actually made available by the producer or depositors.	*E*
The repository should institute procedures to monitor the relevance of any deposit agreement, taking into account the same issues considered in the initial development of the agreement	*E*
The repository shell have a policy that the information to be preserved is understandable to the communities without the need of an expert.	*O*
a clear set of guidance documents such as the organisation's mission and	*D*

collection (selection) policies and guidelines will reduce long term cost defining the aim and direction of collections and services for more efficient decision making	
Quality control checks to ensure that the data producer adheres to the request for metadata	*S*
The policies DRAMBORA defines for archives in order to be fit for data preservation are: *1. Mandate & Commitment to Digital Object Maintenance* *2. Organisational Fitness* *3. Legal & Regulatory Legitimacy* *4. Efficient & Effective Policies* *5. Adequate Technical Infrastructure* *6. Acquisition & Ingest* *7. Preservation of Digital Object Integrity, Authenticity & Usability* *8. Metadata Management & Audit Trails* *9. Dissemination* *10.Preservation Planning & Action*	B

3.1.1.3 Criteria for evaluating data

[This chapter is closely related to the items under the previous chapter 'Collection policy' and these two could possibly be integrated.]

Always check the identities of depositors and the authenticity and integrity of the digital objects.

Recommendation	*Found in*
Check the identities of depositors	*S*
The digital repository ensures the authenticity of the digital objects during all stages of processing. These stages are Ingest, Archival storage, Access	*N*
Archives may express an evaluation of authenticity of its holdings, based on community practise and recommendations.	*O*
The digital repository ensures the integrity of the digital objects during all processing stages. These stages are Ingest, Archival storage, Access	*N*
Risks are assessed according to the following three characteristics: *1. impact: the potential impact that the risk would have should occur* *2. impact expression: the way in which the negative effects of the risk's occurrence manifest themselves* *3. probability: the likelihood of or frequency with which the risk will occur.*	*B*

3.1.2 Legal and Regulatory Framework

Identify roles and conditions of use concerning data and maintain contracts or agreements with data consumers and producers. Define an intellectual property rights, copyright laws and determine who owns the digital materials. Comply with national and international laws and, if applicable, specific regulations within the research community. The archive itself should be a legal entity. Define special procedures for data with disclosure risk. Use a legal advisor if possible to develop the appropriate contracts.

Recommendation	*Found in*
Identify any data protection requirements that you need to address in the course of your research and ensure that these are communicated to all staff.	*C*
The repository shall commit to transparency and accountability in all actions supporting the operation and management of the repository that affect the preservation of digital content over time.	*T*
The repository shall have and maintain appropriate contracts or deposit agreements for digital materials that it manages, preserves, and/or to which it provides access. [These shall] specify and transfer all necessary preservation rights, and those rights transferred shall be documented. *Model contracts with data producers and consumers* *Legal contracts exist between producers and the digital repository.*	*S T N*
The repository shall have specified all appropriate aspects of acquisition, maintenance, access, and withdrawal in written agreements with depositors and other relevant parties […] to ensure that the respective roles of repository, producers, and contributors in the depositing of digital content and transfer of responsibility for preservation are understood and accepted by all parties.	*T*
The repository shall have policies in place to address liability and challenges to ownership/rights […] in order to minimize potential liability and challenges to the rights of the repository. *[These might include] a definition of rights, licenses, and permissions to be obtained from producers and contributors of digital content; citations to relevant laws and regulations; policy on responding to challenges […].*	*T*
Repository needs to be a legal entity	*S*
Published uses of conditions	*S*
Comply with national and international laws	*S*
Defined measures if conditions are not complied with	*S*
Special procedures to manage data with disclosure risk: reviewing disclosure risk, limited/secure access, conditions of distribution, data anonymization	*S*
Staff training to manage data with disclosure risk	*S*

In carrying out its archiving tasks, the digital repository acts on the basis of legal arrangements.	*N*
With regards to use, the digital repository acts on the basis of legal arrangements	*N*
Archives will honour all applicable restrictions of information. They shall understand the intellectual property rights concepts, such as copyrights and any other applicable laws prior to accepting copyrighted materials into the OAIS.	*O*
Use a legal advisor to guide your rights management policy and develop documents.	*D*
Develop model letters for rights clearance, model deposit agreements, model licences and clauses for preservation activities	*D*
If you are licensing material from third parties ensure they have addressed future access to subscribed material in the licence and have robust procedures to support this	*D*
Prepare reasoned arguments and explanations for your preservation activities; rights holders will need to be convinced of the need and persuaded that their interests will be safeguarded.	*D*
Keep detailed records of rights negotiations	*D*
Treat licences and rights correspondence as key institutional records to be retained in fireproof and secure environments	*D*
Make a schedule clearly identifying a list of materials deposited and covered by the licence	*D*
IPR: A clause should be drafted to cover the following: • *Permissions needed for content.* • *Permissions needed for associated software.* • *Permissions needed for copying for the purposes of preservation.* • *Permissions needed for future migration of content to new formats for the purposes of preservation.* • *Permissions needed for emulation for the purposes of preservation.* • *Permissions in respect of copyright protection mechanisms.*	*D*
Statutory and Contractual Issues: A clause should be drafted to cover the following: • *Statutory permissions and legal deposit obligations in respect of electronic materials.* • *Grant and contractual obligations in respect of electronic materials.* • *Conditions, rights and appropriate interests of authors, publishers and other funders.* • *Confidential information and protection of the confidentiality of individuals and institutions.* • *Protecting the integrity and reputation of data creators or other stakeholders.*	*D*

Investment by the preservation agency: A clause should be drafted to cover the following: • *IPR in any value added by the preservation agency.* • *Withdrawal clauses (and associated fees).*	*D*
A broad definition of the repository's regulatory framework is assumed, incorporating acts or provisions with both external and internal origins. Relevant extrinsic commitments and influences include statutory legislation and statutory instruments, global or business-related regulations, de facto or established standards and codes of practice. Internally arising commitments may be traceable to contracts, policies, strategic planning, or accepted business norms.	*B*
One area of priority for near-term action in achieving sustainability concerns public policy action (including modifying copyright laws and to enable digital preservation; creating incentives and requirements for private entities to preserve on behalf of the public; sponsoring public-private partnerships; clarifying rights issues associated with Web-based materials; empowering stewardship organisations to protect digital orphans form unacceptable loss).	*P*
Issues surrounding copyright laws for digital data need careful consideration when choosing preservations strategies.	*P*
"The US Copyright office and the Library of Congress have made recommendations to change the law [of copyrighting digital materials] in light of digital technology realities." The DigiPlanet team "urge Congress to take up this matter expeditiously".	*P*

3.1.3 Funding and Resource Planning

Create and maintain short- and long-term business and financial plans and a funding model. Monitor and analyze financial risks, benefits, investments and expenditures. Carry out marketing and promotional activities suitable for the repository's needs. Maintain a residual fund, if possible, to cover for situations of financial shortfalls.

Concerning human resources, maintain a good level of expertise by actively supporting training and development and internal and external as well as multidisciplinary collaborations.

Recommendation	*Found in*
The repository shall have the appropriate number of staff to support all functions and services. *Sufficient numbers of appropriately qualified staff are available.*	*N T*
The repository shall have in place an active professional development program that provides staff with skills and expertise development	*T*

opportunities.	
The repository shall have short- and long-term business planning processes in place to sustain the repository over time.	*T*
The repository shall have financial practices and procedures which are transparent, compliant with relevant accounting standards and practices, and audited by third parties in accordance with territorial legal requirements.	*T*
The repository shall have an ongoing commitment to analyze and report on financial risk, benefit, investment, and expenditure (including assets, licenses, and liabilities).	*T*
The repository shall have procedures, commitment and funding to replace hardware [and software] when evaluation indicates the need to do so.	*T*
Carry out promotional activities	*S*
Adequate financing of the digital repository is secured.	*N*
In order for the business plan to remain current, regular review is needed. As sources of income cannot normally be guaranteed for more than a few years, monitoring is needed to identify such shortages before they develop into a budget deficiency.	*E*
The repository must achieve an income sufficient for routine functioning.	*E*
Define and maintain marketing and outreach plans suitable for the repository's needs	*E*
Management is often the primary source of funding for an OAIS and may provide for resource utilization (personnel, equipment, facilities).	*O*
Offering bit-level preservation, where the only undertaking is to guarantee storage and delivery of the sequence of bits, will have lower costs than a repository managing full migration paths or emulation solutions.	*D*
Consider engagement in greater collaboration within and between organizations in order effectively to confront and overcome the challenges of digital preservation. Collaboration forms that should be considered include a) Internal collaborations; b) External collaborations; and c) Informal arrangements.	*D*
Preservation strategies enacted early in the life cycle are likely to be more cost effective than salvage attempts left until technology has already moved on significantly	*D*
Develop an Information Strategy which integrates IT training with the overall mission of the institution	*D*
Identify, in consultation with key staff, a skills audit, to determine what specific competencies are required to meet organizational objectives	*D*
Establishing a balance between recruiting specific skills and effectively developing existing talent	*D*
Ensure staff have access to appropriate equipment	*D*
Ensure access to practical "hands on" training and practice	*D*
Encourage networking between colleagues in other institutions	*D*
Consider strategies such as short-term secondment to an institution which may have more experience in a specific area	*D*
Involving staff in designing training and development programmes	*D*
Facilitate effective multi-disciplinary communication	*D*

Take a broad view of what constitutes training and development (i.e. combination of formal courses, both generic and tailor-made, informal training within the organisation, skills transfer within the organisation, networking etc.).	*D*
Work together - researchers and information managers need to communicate regularly. Neither can do their job in isolation.	*C*
The repositories must have strategies to deal with budget shortfalls. Common strategies include willing creditors to boost the repository for a period, or having savings. The repository should also have a prioritized list of the services it provides, beyond what is absolutely required of it, to ease in selecting which could be terminated, if income shortages arrive.	*E*
In order to avoid risks of loss of data and guarantee reliable data preservation, archives and repositories should adhere to the following strategies: *♦ Develop self-sustainability with charged-for services* *♦ Seek assurances of level of budget* *In the event of risk's execution:* *♦ Solicit additional funding to enable achievement of organisational objectives* *♦ Revise objectives if funding stream is insufficiently flexible* *♦ Maintain residual fund where possible to meet shortfalls*	*B*
Define roles and responsibilities among stakeholders to ensure an ongoing and efficient flow of resources to preservation throughout the digital lifecycle.	*P*
Define roles and responsibilities among stakeholders to ensure an ongoing and efficient flow of resources to preservation throughout the digital lifecycle.	*P*
Funding models should be tailored to the norms and expectations of anticipated users.	*P*
One area of priority for near-term action in achieving sustainability concerns organisational action (including developing public-private partnerships; ensuring that organisations have access to skilled personnel, from domain experts to legal and business specialists; creating and sustaining secure chains of stewardship between organisations over time; achieving economies of scale and scope; addressing the free-rider problem).	*P*
The selection of an appropriate funding model requires an in-depth knowledge of the circumstances surrounding the effort, preservation goals, the stakeholder community etc.	*P*
Stimulating growth of capacity and funding to meet the demand [of preservation] is crucial.	*P*
Preservation works best when the interests and actions of users, owners, and archives can be aligned in an economic strategy and operationalized in a business model.	*P*
An important choice available to decision makers in designing suitable preservation strategies is to determine who preserves the digital materials.	*P*
An important choice available to decision makers in designing suitable	*P*

preservation strategies is to determine who pays for the preservation of the digital materials.	
Advice on what to do when roles and responsibilities among stakeholders and actors are unclear.	*P*
There is no single "best" funding model for digital preservation. Common funding models for digital preservation outlined as well as strategies to reduce costs.	*P*
Condition of sustainability: appropriate organisation and governance of preservation activities. Principle of action: Roles and responsibilities among stakeholders must be clear, transparent, and well integrated; and handoffs between responsible parties must be ensured at key moments of risk in the digital lifecycle. Actions: Create effective governance mechanisms to aggregate and rationalize collective preservation interests and costs. Create mechanisms to address free-rider problems in the provision of preservation.	*P*
Condition of sustainability: ongoing and efficient allocation of resources to preservation. Principle of action: Funding models must reflect community norms, be flexible to adjust to disruptions over time, and leverage economies of scale and scope as appropriate. Actions: Choose funding models according to norms and expectations of anticipated users; leverage economies of scale and scope; lower costs of preservation overall.	*P*
Action agenda for National and International Archives: 1.Create mechanisms for public-private partnerships to align or reconcile benefits that accrue to commercial and cultural entities. These agencies can play a critical role in convening stakeholders, sponsoring cooperation and collaboration, and ensuring representation of all stakeholders. 2. Convene expert communities to address the selection and preservation needs of materials of particular interest to the public for which there is no stewardship (Web materials, digital orphans).3. Act expeditiously to reform national and international copyright legislation to address digital preservation needs.4. Create financial incentives to encourage private entities to preserve digital materials on the public behalf.	*P*
Action agenda for Funders and Sponsors of Data creation:1. Create preservation mandates when possible, ensuring that they adhere to community selection criteria, and specifying roles and responsibilities of individuals and organizations. 2. Invest in building capacity throughout the system. The Library of Congress, the National Archives and Records Administration, the National Science Foundation, and JISC have set important precedents for supporting capacity building within specific communities of practice. Seeding stewardship capacity and developing sustainable funding models should, however, be a high priority for all funders. 3. Provide leadership in training and education for 21st century preservation, including domain expertise and core competencies in STEM. Such organizations as the National Archives, Library of Congress, National Library of Medicine, National Agricultural Library, National Science Foundation, Smithsonian Institution, Institute of Museum and Library Services, National Endowment for the Arts, and National Endowment for	*P*

the Humanities in the United States; and the British Library, National Archives, JISC, Digital Curation Centre, and Digital Preservation Coalition in the United Kingdom each have a remit for promoting digital preservation skills. 4. Fund the modeling and testing of domain-specific preservation strategies. This would entail developing domain-specific requirements for lifecycle management to create a timeline of predictable risks, strategies to meet them, and triggering mechanisms to address them.	
Action agenda for Stakeholder organisations: 1. Secure preservation of high-value institutional materials by making explicit roles and responsibilities across organizational boundaries. 2. Develop preservation strategies that assign responsibilities for achieving outcomes. Service-level agreements and MOUs with third-party archives should include contingency plans for handoffs and clauses for putting internal monitoring systems in place. 3. Leverage resources; create economies of scope and economies of scale by partnering with related organizations and industry professional associations. 4. Work with domain and preservation experts to ensure that personnel are fully equipped with the technical skills needed for selecting, curating, and preserving materials. 5. Fund internal preservation and access activities as core infrastructure.	*P*
Action agenda for individuals: 1. Provide nonexclusive rights to preserve content they create and to distribute this content through publicly accessible venues. 2. Partner with preservation experts early in the lifecycle of one's own digital data, to ensure that data are ready to hand off to an archive in forms that will be useful over the long term. 3. Actively participate in professional societies and relevant organizations in developing stewardship best practices and selection priorities.	*P*
Flexibility in preservation decision making is necessary as investment in digital preservation is not necessarily once-and-for-all, or all-or-nothing. Making clear that this flexibility exists when an investment is first considered may make the scope of the perceived economic commitment less daunting, and correspondingly increase decision makers' willingness to make this commitment.	*P*

3.1.4 Long-Term Preservation Policy

Have a preservation strategy plan to act organisational or data related changes such as:

1. Repository ceases to exist.

Maintain a succession plan and contingency plans for financial cutbacks or emergencies. Have an escrow arrangement in case the archive ceases to exist, for example an agreement with another repository about housing vital data.

2. Obsolescence of file formats.

Take measures to ensure long-term data usability. [see more in chapter 2 on this issue]

Define a long-term preservation policy and consider the broader economic setting in which preservation decisions are made. Base the decisions which data to preserve on the perceived value associated with the digital materials over time. One should keep investing in the maintenance to ensure the value creating capacity over time.

Recommendation	*Found in*
Be aware of, and undertake management and administrative actions planned to promote curation and preservation throughout the curation lifecycle.	*C*
Work with researchers and information managers to develop policies and to identify realistic and implementable workflows.	*C*
Appraise for the here and now but with an eye to the future.	*C*
The repository shall have a Preservation Strategic Plan that defines the approach the repository will take in the long-term support of its mission […] to help the repository make administrative decisions, shape policies, and allocate resources in order to successfully preserve its holdings. *Provide a plan for long-term preservation for digital assets*	*S T*
The repository shall have Preservation Policies in place to ensure its Preservation Strategic Plan will be met.	*T*
The repository shall have mechanisms for review, update, and ongoing development of its Preservation Policies as the repository grows and as technology and community practice evolve.	*T*
The repository shall identify the Content Information and the Information Properties that the repository will preserve […] in order to make it clear to funders, depositors, and users what responsibilities the repository is taking on and what aspects are excluded. It is also a necessary step in defining the information which is needed from the information producers or depositors.	*T*
The repository shall have a procedure(s) for identifying those Information Properties that it will preserve [in order] to establish a clear understanding with depositors, funders, and the repository's Designated Communities how the repository determines and checks what the characteristics and properties of preserved items will be over the long term. These procedures will be necessary to confirm authenticity or to identify erroneous claims of authenticity of the preserved digital record.	*T*
The repository shall have a record of the Content Information and the Information Properties that it will preserve […]. *This is necessary in order to identify in writing the Content Information of*	*T*

the records for which it has taken preservation responsibility and the Information Properties it has committed to preserve for those records based on their Content Information.	
Take future obsolescence of file formats into account	*S*
Take measures to ensure long-term data usability	*S*
The digital repository assumes responsibility for long-term preservation of the information represented by the digital objects	*N*
Appropriate organisational structures exist for the digital repository.	*N*
The digital repository engages in long-term planning.	*N*
The digital repository reacts to substantial changes	*N*
The digital repository implements strategies for the long-term preservation of the archival information packages (AIPs).	*N*
The preservation tasks are ensured even beyond the existence of the digital repository by an agreement with other repositories to take over the collections. The very purpose of the succession plan(s) is to detail agreement(s) about who will inherit the digital data if the repository ceases to function.	*E*
Being explicit about preservation commitments is a crucial element of preservation policy and procedures.	*D*
Policies to identify which digital resources should be stored online	*D*
Retention policies to determine at what stage (if ever) online storage of digital resources will be re-assessed	*D*
The archive should have a formal succession plan, contingency plans, and/or escrow arrangements in place, in case the archive ceases to operate or the governing or funding institution substantially changes its scope. *The repository shall have an appropriate succession plan, contingency plans, and/or escrow arrangements in place in case the repository ceases to operate or the governing or funding institution substantially changes its scope.*	*O* *T*
Ensure contingency plans for financial cutbacks or emergencies: have an agreement with another repository about housing vital data; prioritise services to be retained in the case of financial problems	*E*
For each risk identified, the final stage of the self-assessment process is to define appropriate risk management measures and targets[...]. For each risk details of treatment, avoidance measures, and anticipated outcomes should be recorded and monitored over time. A timeframe for regular reassessment should also be recorded at this stage to ensure that new risks are identified and mitigated as the repository matures.	*B*
Consider the broader economic setting in which preservation decisions are made as these are fundamental in shaping the prospects for achieving long-term sustainability.	*P*
Articulate a compelling value proposition– e.g. when making the case for preservation, make the case for use.	*P*
Condition of sustainability: recognition of the benefits of preservation by decision makers. Principle: To make a case for preservation, make the case	*P*

for use. Actions: Articulate the value proposition in terms of use cases; identify at-risk materials whose damage or loss is unacceptable; use proxy organizations to aggregate diffuse demand; use option strategies in cases of uncertain value.	
Provide clear incentives to preserve in the public interest. Condition of sustainability: incentives for decision makers to act in the public interest. Principle of action: Incentives must be strengthened when they are weak; aligned when they diverge among different stakeholder communities; and created where there are none. Actions: When there are insufficient incentives, use preservation mandates as appropriate. Provide financial incentives for private owners to preserve on behalf of the public. Bring copyright law and mandatory deposit requirements up to date for digital preservation. Remove barriers to creating efficient decentralized stewardship mechanisms by use of nonexclusive licenses granting preservation rights to third parties.	*P*
One area of priority for near-term action in achieving sustainability concerns Education and Public Outreach action (including: promoting education and training for 21st century digital preservation and raising awareness of the urgency to take timely preservation actions).	*P*
Five conditions required for economic sustainability of digital preservation are: -recognition of the benefits of preservation by decision makers; -a process for selecting digital materials with long-term value; -incentives for decision makers to preserve in the public interest; -appropriate organization and governance of digital preservation activities; and -mechanisms to secure an ongoing, efficient allocation of resources to digital preservation activities.	*P*
A sustainable preservation strategy must be flexible enough to span generations of data formats, access platforms, owners and users.	*P*
Demand for preservation must be articulated well enough to ensure there is sufficient supply.	*P*
It is important to frame the demand for preservation services as a demand for the product of those services (the digital materials) for future use.	*P*
The decision to preserve will ultimately be based on the perceived value associated with the digital materials over time. We must make ongoing investments in their maintenance to sustain their value-creating capacity over time.	*P*
Preserving digital orphans is a priority for public institutions. ('Digital orphans' is the name given to materials on the Web whose ownership or provenance are uncertain).	*P*
Condition of sustainability: selection of materials with long-term value	*P*
Principle of action: Priorities must be made for selecting materials that have the greatest promise of returning value to users over time. Action: Make decisions about priorities among collections competing for scarce resources based on projected future use.	*P*
Policy interventions for digital preservation can be necessary when the level of investment in preservation by private decision-makers to meet their own needs falls short of what is optimal for society.	*P*

3.1.5 Access Policy

Require the data consumer to comply with an access agreement or user contract, in which the conditions of use are defined. In this document, refer to and include requirements to comply with generally accepted regulations and codes of conduct for data usage. Make sure these conditions and the access policy are well communicated to the designated community. If applicable, use special contracts for accessing confidential data. Define measures which are taken if conditions are not complied with.

Recommendation	*Found in*
Ensure that any restrictions on access and use are communicated and respected.	*C*
The repository shall comply with Access Policies […] to ensure [it] has fully addressed all aspects of usage which might affect the trustworthiness of the repository, particularly with reference to support of the user community.	*T*
Require the data consumer to comply with access regulations and generally accepted codes of conduct (e.g. End User Licences)	*S*
Define measures which are taken if conditions are not complied with.	*S*
Special contracts to grant access to confidential data	*S*
The digital repository ensures its designated community/ communities can access the digital objects.	*N*
The digital repository ensures that the designated ommunity/communities can interpret the digital objects.	*N*
The digital repository permits usage of the digital objects based on defined criteria.	*N*
The digital repository acquires adequate metadata to record the corresponding usage rights and conditions.	*N*
Create, Maintain and Review a Mission Statement which reflects the Repository's Mandate	*E*
The definition of the designated community should be available on the repository website and the community should be very well aware of available delivery and access options. It is essential to monitor and reflect all changes inside the community over the time.	*E*
Supporting tools, such as access and use mechanisms, i.e. Content Data Object (CDO) are determined to remain adequate for an extended period of time. Once this software will cease to function it shall be re-implemented, or alternatively emulated by new software.	*O*
A clause should be drafted to cover permissions and conditions in respect of access to the material	*D*
To guarantee reliable long-term data access archives and repositories should adhere to the following strategies: *· Define policies describing available information delivery services and communicate these to the user community* *· Implement appropriate systems to meet delivery policy*	*B*

requirements *· Establish sufficiently robust technical infrastructure to satisfy demands of proposed delivery services*	
One way to strengthen preservation incentives is for stakeholders to aggregate and leverage their demand-side power to negotiate preservation requirements as part of access arrangements. An example is the NESLi2 Model License for Journals in the United Kingdom, used by the Joint Information Systems Committee (JISC) in negotiating e-journal licenses on behalf of UK higher education institutions (See http://www.nesli2.ac.uk)	*P*

3.2 Technological Environment

3.2.1 IT Architecture

The IT infrastructure must be able to guarantee the integrity, security and availability of the data. Identify and manage the risks to the repository's preservation operations regarding the system infrastructure. Structure the metadata elements/digital objects based on established standards. Use adequate unique persistent identifiers to fit the repository's current but also future requirements taking account e.g. the numbers of objects.

Recommendation	Found in
Definition of data: *Data, any information in binary digital form, is at the centre of the Curation Lifecycle.* *This includes:* *Digital Objects: simple digital objects (discrete digital items such as text files, image files or sound files, along with their related identifiers and metadata) or complex digital objects (discrete digital objects made by combining a number of other digital objects, such as websites).* *Databases: structured collections of records or data stored in a computer system.*	*C*
The system of [unique persistent] identifiers shall be adequate to fit the repository's current and foreseeable future requirements such as numbers of objects.	*T*
The repository shall identify and manage the risks to its preservation operations and goals associated with system infrastructure. *[This can, for example, be achieved by means of] infrastructure inventory of system components; periodic technology assessments; estimates of system component lifetime; export of authentic records to an independent system; use of strongly community supported software e.g., Apache, iRODS, Fedora); re-creation of archives from backups.*	*T*

The repository shall have hardware [and software] technologies appropriate to the services it provides to its designated communities.	*T*
Derive metadata elements from established metadata standards, registries or conventions	*S*
Archival storage of the digital objects is undertaken to defined specifications.	*N*
The data management system is capable of providing the necessary digital repository functions.	*N*
The digital repository uniquely and persistently identifies its objects and their relationships.	*N*
The IT infrastructure is appropriate.	*N*
The IT infrastructure implements the object management requirements.	*N*
The IT infrastructure implements the security requirements of the IT security system.	*N*
The infrastructure protects the digital repository and its digital objects.	*N*
IT Infrastructure must be capable of coping with the scale of data storage, processing and transport appropriate for the repository.	*E*
IT infrastructure must be able to guarantee the integrity and security of the stored data.	*E*
The IT infrastructure must be able to guarantee that certain services remain available to the users. Identify which services of the repository are no longer required and can be closed and which are in great demand.	*E*
OAIS associations can be categorized technically by both external and internal factors. External factors include characteristics of the Producer and Consumer communities. Internal factors could include common implementations of the information models presented in 3.2, or multi-Archive sharing of one or more of the functional areas presented in 3.1.	*O*
DRAMBORA itself does not describe any architectural constructs. The nestor criteria catalogue, and ISO 27001:2005 is mentioned briefly.	*B*
One area of priority for near-term action in achieving sustainability concerns technical action (including building capacity to support stewardship in all areas; lowering the cost of preservation overall; determining the optimal level of technical curation to operationalize an option strategy for all types of digital material.	*P*

3.2.2 Standards and Formats

Identify the needs for data formats for data depositing, data archiving and data dissemination. Use widely adopted formats acknowledged by the community.

Check that the data are deposited in the required formats.

Store data in a format that is immune to hardware or software obsolescence and/or be ready to migrate the data if necessary.

Disseminate data in open, non-proprietary and well-documented file formats when possible.

Recommendation	Found in
Store the data in a secure manner adhering to relevant standards.	*C*
Create new data from the original, for example: by migration into a different format, or by creating a subset, by selection or query, to create newly derived results, perhaps for publication	*C*
Migrate data to a different format. This may be done to accord with the storage environment or to ensure the data's immunity from hardware or software obsolescence.	*C*
The repository shall have adequate specifications enabling recognition and parsing of the SIPs. *[This can, for example, take the form of] Packaging Information for the SIPs; Representation Information for the SIP Content Data, including documented file format specifications; published data standards; documentation of valid object construction.*	*T*
The repository shall have a description of how AIPs are constructed from SIPs.	*T*
The repository shall have and use a convention that generates persistent, unique identifiers for all AIPs [so as to] uniquely identify each AIP [and its components] within the repository.	*T*
Provide a plan for infrastructural development	*S*
Technical infrastructure should be based on internationally accepted archival standards like OAIS	*S*
Provide data in formats used by the research community	*S*
Publish a list of preferred formats	*S*
Check that data are deposited in preferred formats	*S*
Require detailed information about file formats from depositors	*S*
The digital repository records adequate metadata for formal and content based description and identification of the digital objects.	*N*
The digital repository records adequate metadata for structural description of the digital objects.	*N*
The digital repository records adequate metadata to document all the changes made by the digital repository to the digital objects.	*N*
The digital repository acquires adequate metadata for technical description of the digital objects.	*N*
The package structure is preserved at all times.	*N*
What are the main sources of bibliographic and descriptive metadata in the repository?	*E*
What interoperability standards are implemented in the Repository?	*E*
A data provider encodes his data in a package format (Submission Information Package) acceptable to the repository. The repository receives these SIPs, and repackages them for storage (Archival Information Package). A repository user requests the data, and the repository repackages it in a format appropriate for the user (Dissemination	*E*

Information Package).	
For Submission Information Package: Specify the digital object formats and accompanying meta-data the repository will accept. Formats (potentially) containing encryption, or embedded objects and files, as well as commercially protected formats can sometimes prove impossible to convert.	*E*
For Archival Information Package: Specify the data format and metadata content for archiving digital objects. Where the SIPs defined in the previous goal should strike some balance between commonly used formats, and formats useful for archiving, the AIPs do not need to.	*E*
For Dissemination Information Package: Specify the data formats used for disseminating digital objects.	*E*
To preserve the meaning of an Information Object, its Representation Information must also be preserved. This is most easily accomplished when the Representation Information objects are expressed in forms that are easily understandable, such as text descriptions that use widely supported standards such as ASCII characters for electronic versions. One problem with the use of only text descriptions is that such descriptions can be ambiguous. This is addressed by the use of standardized, formal description languages containing well-defined constructs with which to describe data structures. These languages may need to be augmented with text descriptions to convey fully the semantics of the Representation Information.	*O*
Selection of cataloguing and documentation standard	*D*
Knowledge of all relevant standards for all categories of digital resources acquired by the institution.	*D*
Written guidelines on preferred and acceptable standards	*D*
Institutional strategies for outreach, collaboration, standards and best practice.	*D*
Technology watch on standards activities	*D*
Use "open" non-proprietary, well-documented file formats wherever possible.	*D*
Alternatively utilise file formats which are well-developed, have been widely adopted and are de facto standards in the marketplace.	*D*
Identify formats acceptable for the purposes of transfer, storage and distribution to users (these may be distinct).	*D*
Minimise the number of file formats to be managed as far as is feasible/desirable	*D*
Do not use encryption or compression for archival files if possible	*D*
Standards that are recommended by DRAMBORA: *♦ ISO 9000:2000 Quality Management Systems Series* *♦ ISO 27001:2005 Information technology — Security techniques — Information security management systems — Requirements*	*B*

3.2.3 Security and Risk Management / Media Monitoring and Refreshing Strategy

The data-archive should define and maintain a policy for system and physical security.

The documentation should include an analysis of security risk factors and a crisis management plan.

The repository should take the following measures regarding changes:

-Documented change management process that identifies changes to critical processes

-Periodic risk analysis and systematic reports

-Periodic long-term value analysis on digital assets

-Security updates and backups

-Defined processes for storage media or hardware change (e.g. refreshing, migration)

-Implement technology watches to flag changes which might lead to hardware or software retention.

Repeat the risk management cycle regularly since some risks may disappear over time while other new risks emerge.

Further, measures should be taken to unable unauthorized system or physical access to data.

Recommendation	Found in
Dispose of data, which has not been selected for long-term curation and preservation in accordance with documented policies, guidance or legal requirements. *Typically data may be transferred to another archive, repository, data centre or other custodian. In some instances data is destroyed. The data's nature may, for legal reasons, necessitate secure destruction.*	*C*
The repository shall have adequate hardware and software support for backup functionality sufficient for preserving the repository content and tracking repository functions. *[This can include, for example,] documentation of what is being backed up and how often; audit log/inventory of backups; validation of completed backups; disaster recovery plan, policy and documentation; fire drills; testing of backups; support contracts for hardware and software for backup mechanisms; demonstrated preservation of system metadata such as access controls, location of replicas, audit trails, checksum values.*	*T*

The repository shall have a process to record and react to the availability of new security updates based on a risk-benefit assessment.	*T*
The repository shall have defined processes for storage media and/or hardware change (e.g., refreshing, migration).	*T*
The repository shall have identified and documented critical processes that affect its ability to comply with its mandatory responsibilities […] to ensure that the critical processes can be monitored to ensure that they continue to meet the mandatory responsibilities and to ensure that any changes to those processes are examined and tested.	*T*
The repository shall have a documented change management process that identifies changes to critical processes that potentially affect the repository's ability to comply with its mandatory responsibilities.	*T*
The repository shall have a process for testing and evaluating the effect of changes to the repository's critical processes.	*T*
The repository shall maintain a systematic analysis of security risk factors associated with data, systems, personnel, and physical plant.	*T*
The repository shall have implemented controls to adequately address each of the defined security risks.	*T*
The repository staff shall have delineated roles, responsibilities, and authorizations related to implementing changes within the system.	*T*
The repository shall have suitable written disaster preparedness and recovery plan(s), including at least one off-site backup of all preserved information together with an offsite copy of the recovery plan(s).	*T*
Create a crisis management plan	*S*
Political upheaval can come to seemingly stable societies. Repositories caught in such circumstances could face opposing demands to both open up their collections, and to close them down.	*E*
Many repositories are a part of a larger organization, which due to some changes could have problems seeing the purpose of the repository. The repository must actively seek to demonstrate its value to those with executive power over it.	*E*
The repository should, so far as possible, detail procedures for dealing with environmental disasters in a timely manner.	*E*
Loss of educated key staff: • *the internal workings and trade secrets of the repository could be revealed to outsiders – the repository should prepare press statements and other communications for when the secrets break* • *the services of the repository could cease to function - the best course of action would be to shut down unmaintainable services until new staff have been hired or trained.*	*E*
The physical security of the repository can also be threatened. The repository should identify likely threats, and address them, in addition to having plans in the event of these threats materializing.	*E*
Media Decay: Digital media, over time, become increasingly unreliable as secure preservers of bits. Even those that are used with some level of error correction eventually need to be replaced. The net result of media decay is that AIP information must be moved to newer media.	*O*

Management will generally conduct some regular review process to evaluate the OAIS performance and progress toward Long Term goals, and assess the risks to which the OAIS and its holdings are exposed.	*O*
The Establish Standards and Policies function is responsible for establishing and maintaining the Archive system standards and policies. It receives budget information and policies such as the OAIS charter, scope, resource utilization guidelines, and pricing policies It provides Management with periodic reports. It receives from Management. Recommendations for Archive system enhancement, proposals for new Archive data standards, and periodic risk analysis reports from Preservation Planning. It will have to face risks from unforeseen events (unplanned down time due to network outage, software bugs, hardware failure, human error, disk crash, etc.) and make the appropriate decisions to minimize the risk of not fulfilling the Archive's commitments. It also receives performance information and Archive holding inventories from Manage System Configuration.	*O*
System and physical security policies and procedures should be in place to ensure the care and integrity of items during accessioning. These should be developed from and reflect the institutional policies and procedures on security.	*D*
Establish disaster recovery plan	*D*
Control access to storage facilities and processing areas. Store in separate, preferably lockable area	*D*
Ensure no unauthorised access	*D*
Design audit features into mass storage systems and computerised physical access controls. Undertake regular random checks if automated audits are not feasible	*D*
Establish procedures to ensure no deliberate or inadvertent changes can take place	*D*
Ensure all legal requirements are met	*D*
Establish procedures for ensuring authenticity	*D*
Use passwords and user ids, and other network security procedures	*D*
Define system and area access privileges for staff	*D*
Assign specific staff responsibilities for data security and storage facilities	*D*
Undertake a retrospective survey of digital holdings, a risk assessment and action plan	*D*
Implement a process of technology watch and/or implement procedures for standardisation and changes in technology in your IS strategy.	*D*
Maintain a list of hardware/software available within the institution and use this to flag implications for technology change and hardware/ software replacement/ retention.	*D*
Ensure you have good preservation metadata in a computerised catalogue which can form the basis for technology watch and monitoring.	*D*
Consider "digital archaeology" to retrieve access to data in obsolete formats.	*D*
Risks are assessed according to the following three characteristics:	*B*

1. impact: the potential impact that the risk would have should occur *2. impact expression: the way in which the negative effects of the risk's occurrence manifest themselves* *3. probability: the likelihood of or frequency with which the risk will occur.*	
♦ highlight what digital resources are at risk within their organisation; *♦ highlight the risks to which these digital resources were exposed;* *♦ highlight the risks to organisations posed by threats to digital resources (e.g. reputation, cessation of business);* *♦ categorise and prioritise risks in order to facilitate their management;* *♦ facilitate communication within the organisation about areas of risk;* *♦ stimulate risk management strategy development.*	*B*
Advice on what to do when the benefits from assets no longer outweigh the costs of maintaining them e.g. the best practice for all institutions is to have a plan for accessing long-term value periodically, at which point they decide whether to renew, handoff, or abandon a preservation commitment.	*P*

3.3 Data Curation

3.3.1 Pre-Ingest Function

Be aware of the expectations, roles, processes and intellectual property rights.

Recommendation	Found in
Conceive and plan the creation of data, including capture method and storage options.	*C*
Get into the habit of equating data curation with good research.	*C*
Know what your funding body expects you to do with your data and for how long. Assess your ability to be able to meet these expectations (i.e., do you need additional funding or staff?).	*C*
Determine intellectual property rights from the outset and ensure they are documented.	*C*
Identify any anticipated publication requirements (embargoes, restrictions on publishing over multiple sites).	*C*
Identify and document specific roles and responsibilities as early as possible.	*C*
The repository shall have contemporaneous records of actions and administration processes that are relevant to content acquisition.	*T*
The first contact between the OAIS and the Producer is a request that the OAIS preserve the data products created by the Producer. This contact may be initiated by the OAIS, the Producer or Management. The Producer	*O*

establishes a Submission Agreement with the OAIS, which identifies the SIPs to be submitted and may span any length of time for this submission. Some Submission Agreements will reflect a mandatory requirement to provide information to the OAIS, while others will reflect a voluntary offering of information and others may reflect any payments which may be involved. Even in the case where no formal Submission Agreement exists such as in the archiving of many World Wide Web sites, a virtual Submission Agreement may exist specifying the file formats and the general subject matter the OAIS will accept.	
Within DRAMBORA, digital curation is characterised as a risk-management activity; the job of digital curator is to rationalise the uncertainties and threats that inhibit efforts to maintain digital object authenticity and understandability, transforming them into manageable risks.	*B*

3.3.1.1 Information and guidance given to data producer

The repository shall have written policies that indicate when it accepts preservation responsibility for contents of each set of submitted data objects, such as submission agreements, deposit agreements, and deeds of gift; confirmation receipt sent back to producer/depositor.

Define the standards to use for content, syntax and structure and communicate these. Provide user friendly ways to enter metadata, e.g. clearly defined deposit forms. Specify the SIPs and criteria for accepting digital objects, provide the producer with responses at agreed points during the ingest process.

Recommendation	Found in
Agree from an early stage any standards you will be making use of for content, syntax, and structure. Once these have been agreed, make sure they are communicated - both to other researchers on the project and to the data/information managers you will be working with. Provide training if necessary.	*C*
Identify data quality metrics as soon as possible and ensure that these are communicated and monitored.	*C*
The repository shall have written policies that indicate when it accepts preservation responsibility for contents of each set of submitted data objects […] to avoid misunderstandings between the repository and producer/depositor as to when and how the transfer of responsibility for the digital content occurs. *[This can be demonstrated, for example, by] properly executed submission agreements, deposit agreements, and deeds of gift; confirmation receipt sent back to producer/depositor.*	*T*
The repository shall provide the producer/depositor with appropriate	*T*

responses at agreed points during the ingest processes.	
Provide information on the full package information that should be deposited	*S*
Provide: • *User friendly ways to enter metadata* • *Deposit forms with resource discovery metadata* • *Tools to create metadata at file level*	*S*
The digital repository accepts digital objects from the producers based on defined criteria.	*N*
The digital repository specifies its submission information packages (SIPs).	*N*
The digital repository identifies which characteristics of the digital objects are significant for information preservation.	*N*
The repository should have some description for content providers about how to package data and metadata or representation information. The repository should have a policy regarding the completeness and correctness of new data and what action to take regarding invalid or incomplete data.	*E*
Specify sources and formats for bibliographic and descriptive metadata in the received data. This is the basic information required to enable the object to be discovered (e.g. by a search engine) and interpreted.	*E*
Specify technical metadata in the SIP	*E*
The Receive Submission function provides the appropriate storage capability or devices to receive a SIP from the Producer (or from Administration). Digital SIPs may be delivered via electronic transfer (e.g., FTP), loaded from media submitted to the Archive, or simply mounted (e.g., CD-ROM) on the Archive file system for access. Non-digital SIPs would likely be delivered by conventional shipping procedures. The Receive Submission function may represent a legal transfer of custody for the Content Information in the SIP, and may require that special access controls be placed on the contents. This function provides a confirmation of receipt of a SIP to the Producer, which may include a request to resubmit a SIP in the case of errors resulting from the SIP submission.	*O*
Provide documentation to guide and support transfer of digital resources from suppliers.	*D*

3.3.2 Ingest Function

Ingest functions include receiving SIPs, performing quality assurance on SIPs, generating an Archival Information Package (AIP) which complies with the Archive's data formatting and documentation standards, extracting Descriptive Information from the AIPs for inclusion in the Archive database, and coordinating updates to Archival Storage and Data Management.

Think about who you are creating your data for. Follow documented procedures and define reasons why not to include a certain SIP. Related to tour data, create administrative, descriptive, structural, technical and preservation metadata.

Recommendation	Found in
Create data including administrative, descriptive, structural and technical metadata. Preservation metadata may also be added at the time of creation.	*C*
Know who you are creating your data for and what you want them to be able to do (and not do) with it. Communicate this with others on the project.	*C*
Be realistic – strike a balance between what is sufficient and what is ideal based on your practical realities.	*C*
Transfer data to an archive, repository, data centre or other custodian. Adhere to documented guidance, policies or legal requirements.	*C*
Remember - ingest does not necessarily need to mean deposit in a data centre or repository but rather moving to a 'curated' environment – could be as simple as a specific folder on a shared drive.	*C*
The repository shall document the final disposition of all SIPs. In particular […,] the repository shall follow documented procedures if a SIP is not incorporated into an AIP or discarded and shall indicate why the SIP was not incorporated or discarded.	*T*
The Ingest Functional Entity (labeled 'Ingest' in the figures in this section) provides the services and functions to accept Submission Information Packages (SIPs) from Producers (or from internal elements under Administration control) and prepare the contents for storage and management within the Archive. Ingest functions include receiving SIPs, performing quality assurance on SIPs, generating an Archival Information Package (AIP) which complies with the Archive's data formatting and documentation standards, extracting Descriptive Information from the AIPs for inclusion in the Archive database, and coordinating updates to Archival Storage and Data Management.	*O*
DRAMBORA does not give any specific guidelines regarding ingest functions. It focuses on beforehand risk assessment and the audit process.	*B*

3.3.2.1 Information and documentation from data producer

Specify the information that is needed from the data producer. Use Submission Agreements, which define information such as the content, format, and scheduled arrival times of the Submission Information Package (SIP). The SIP is an Information Package that is provided to the OAIS by the Producer.

Give a formal receipt for closure to the Data Producer. Good practice is also to provide references to publications, used methods and research techniques/data collection method. Provide sufficient information to others to assess the scientific quality of data.

Recommendation	Found in
Receive data, in accordance with documented collecting policies, from data creators, other archives, repositories or data centres, and if required assign appropriate metadata.	*C*
Make sure you know about any repository policies that might affect your deposit for long-term storage (i.e., what will they accept, are there preferred formats or normalisation processes).	*C*
Get a formal receipt (if possible) or an informal acknowledgement for closure and transfer of stewardship	*C*
The repository shall clearly specify the information that needs to be associated with specific Content Information at the time of its deposit [to achieve] a clear understanding of what needs to be acquired from the Producer.	*T*
Sufficient information to others to assess the scientific and scholarly quality of data	*S*
References to publications	*S*
Used methods and research techniques, including data collection	*S*
Information on how research data obtained shape	*S*
The data within the data Producer entity are private and may be in any format the Producer desires. However, when the decision is made to store the data in an OAIS, the Producer who is responsible for the data meets with archivists to negotiate a Submission Agreement as [..] This agreement defines information such as the content, format, and scheduled arrival times of the Submission Information Package (SIP). The SIP is an Information Package that is provided to the OAIS by the Producer. The SIP consists of the Content Information plus the data that is necessary to assure that those data can be maintained by the OAIS and that the data can be interpreted and used by Consumers who withdraw them from the OAIS in the future.	*O*
A minimum standard of information required for cataloguing	*D*
Data producers are to be aware of the quality assurance process, mentioned later in the report.	*B*

3.3.2.2 Quality assurance and data checking

Level of data quality should be defined by the fitness for purpose. One should define all processes and responsibilities, for example which function within the organisation is responsible for data quality assurance at the point of deposit. One should verify the identity of the producer of all materials and the

completeness and correctness of each SIP. To support this, one should use data submission agreements. To be able to verify correct procedure it is good to keep system logs from procedures, authentications, files received, etc. Through quality controls, and quality assurance procedures for each object ingested, one can check that the data are deposited in preferred formats and all metadata are included. For digital submissions, the mechanisms might include cyclic redundancy checks and checksums.

Recommendation	Found in
Ensure that your data meets minimum quality assurance metrics (based on intended use).	*C*
Decide on who is responsible for final aspects of data quality assurance at the point of deposit (researcher, archive, information manager, etc…). Ensure that this final point of QA is communicated to all stakeholders.	*C*
Data quality is not absolute. Level of data quality and cleaning must be assessed by fitness for purpose. So, 'high quality' data for one user group may be completely unsuitable for another user group.	*C*
The repository shall have mechanisms to appropriately verify the identity of the Producer of all materials […] to avoid providing erroneous provenance to the information which is preserved. This can be achieved, for example, with the help of] legally binding submission agreements/deposit agreements/deeds of gift, evidence of appropriate technological measures; logs from procedures and authentications.	*T*
The repository shall have an ingest process which verifies each SIP for completeness and correctness. [This can be achieved, for example, with the help of] system log files from system(s) performing ingest procedure(s); logs or registers of files received during the transfer and ingest process; documentation of standard operating procedures, detailed procedures, and/or workflows; format registries; definitions of completeness and correctness.	*T*
Quality controls to ensure data is deposited in preferred formats and that requested metadata are included	*S*
All processes and responsibilities have been defined.	*N*
The digital repository documents all its elements based on a defined process.	*N*
The Quality Assurance function validates (QA results) the successful transfer of the SIP to the temporary storage area. For digital submissions, these mechanisms might include Cyclic Redundancy Checks (CRCs) or checksums associated with each data file, or the use of system log files to record and identify any file transfer or media read/write errors.	*O*
Unique numbering of each item accessioned.	*D*
Marking and labeling procedures	*D*
Handling guidelines for different media	*D*
Validation procedures to check media, content, and structure	*D*

Repository requires staff to undertake quality assurance procedures for each object ingested, which takes on average 10 minutes, although an additional policy states that ingest should be completed in 10 minutes.	*B*
Before assessing their repository's aims or shortcomings, it is vital that repository administrators define the purpose and scope of what they are aiming to achieve with the self-assessment process. To assist with this, DRAMBORA Interactive permits users to register details about their own repository to help determine their self-assessment profile. DRAMBORA Interactive also helps users to identify, register, and provide access to additional institutional staff members who will contribute to the various steps of the self-assessment process. DRAMBORA Interactive features robust security features to ensure that sensitive corporate information is protected against non-authorised access.	*B*
Self-assessment is a key point in DRAMBORA. The self-assessment is divided into several phases, the first of which focuses on attaining a comprehensive overview of the objectives [...] and activities of the organisation supported by adequate documentation. In this phase, questions about the institution's mandate, its organisational constraints (legal, political, and/or technical limitations or obligations), overarching objectives, and business activities are explored.	*B*

3.3.2.3 Data documentation and enhancement

Making use of archival standards for hierarchical data description and metadata elements can be useful. The data repository should have a definition for each AIP that is adequate for long-term preservation. Use documented processed for acquiring Preservation Description Documentation and a method for verifying the completeness and correctness of newly generated AIPs. To enhance data, provide references to publications based on the data.

Recommendation	Found in
Making use of archival standards like ISAD-G can be useful for hierarchical data description. So, talk to information managers at your institution for advice.	*C*
The repository shall have for each AIP or class of AIPs preserved by the repository an associated definition [including appropriate Packaging Information] that is adequate for parsing the AIP and fit for long-term preservation needs.	*T*
The repository shall be able to identify which definition applies to which AIP. [This can be achieved, for example, by generating and employing] documentation clearly linking each AIP, or class of AIPs, to its definition.	*T*
The repository shall have a definition of each AIP that is adequate for long-term preservation, enabling the identification and parsing of all the required components within that AIP. *[The definitions shall be used] to extract Content Information and PDI*	*T*

(Provenance, Access Rights, Context, Reference, and Fixity Information) from AIPs.	
The repository shall have access to necessary tools and resources to provide authoritative Representation Information for all of the digital objects it contains. *[This includes] tools or methods to identify the file type of all submitted Data Objects [and] to determine what Representation Information is necessary to make each Data Object understandable to the Designated Community.*	*T*
The repository shall have access to the requisite Representation Information.	*T*
The repository shall have tools or methods to ensure that the requisite Representation Information is persistently associated with the relevant Data Objects.	*T*
The repository shall have documented processes for acquiring Preservation Description Information (PDI) for its associated Content Information and acquire PDI in accordance with the documented processes.	*T*
The repository shall ensure that the PDI is persistently associated with the relevant Content Information.	*T*
The repository shall capture or create minimum descriptive information and ensure that it is associated with the AIP.	*T*
Provide references to publications based on the research data	*S*
Pay attention to relevance of metadata items to data consumers	*S*
Use metadata elements which are derived from established metadata standards, registries or conventions.	*S*
The repository should have a method for verifying the completeness and correctness for newly generated AIPs, as with SIPs.	*E*
The metadata from the SIP should be conserved, and elaborated, rather than being changed. A file format, that strikes a balance between being human readable and being machine parseable, such as XML could be used to encode the metadata.	*E*
The Generate AIP function transforms one or more SIPs into one or more AIPs that conform to the Archive's data formatting standards and documentation standards. This may involve file format conversions, gathering adequate Representation Information, data representation conversions or reorganization of the Content Information in the SIPs. The Generate AIP function may issue report requests to Data Management to obtain reports of information needed by the Generate AIP function to produce the Descriptive Information that completes the AIP. This function sends SIPs or AIPs for audit to the Audit Submission function in Administration, and receives back an audit report. As a result of the audit report for example, it may be necessary to gather further Representation Information to ensure that the Content Information is understandable and usable by the Designated Community.	*O*
Guidelines for retrospective documentation or catalogue enhancement	*D*

3.3.3 Archival Storage and Preservation

Undertake actions to ensure long-term preservation and retention of the authoritative nature of data. These actions include data cleaning, validation, assigning preservation metadata, assigning representation information and ensuring acceptable data structures or file formats.

Be able to provide a complete list of all [unique persistent] identifiers and do spot checks for duplications.

Define AIPs, verify each AIP for completeness and correctness at the point it is created and ensure that they van al be located and retrieved.

Document preservation actions so that people know what has been done to the data over time.

Recommendation	Found in
Evaluate data and select for long-term curation and preservation. Adhere to documented guidance, policies or legal requirements.	*C*
Undertake actions to ensure long-term preservation and retention of the authoritative nature of data. Preservation actions should ensure that data remains authentic, reliable and usable while maintaining its integrity. Actions include data cleaning, validation, assigning preservation metadata, assigning representation information and ensuring acceptable data structures or file formats.	*C*
Document preservation actions so that people know what has been done to the data over time.	*C*
The repository shall obtain sufficient [legal and physical] control over the Digital Objects to preserve them.	*T*
The repository shall be able to provide a complete list of all [unique persistent] identifiers and do spot checks for duplications.	*T*
The repository shall verify each AIP for completeness and correctness at the point it is created.	*T*
The repository shall have contemporaneous records of actions and administration processes that are relevant to AIP creation.	*T*
The repository shall preserve the Content Information of AIPs.	*T*
The repository shall have contemporaneous records of actions and administration processes that are relevant to storage and preservation of the AIPs.	*T*
The repository shall maintain bi-directional linkage between each AIP and its descriptive information [over time] to ensure that all AIPs can be located and retrieved.	*T*
The digital repository defines its archival information packages (AIPs).	*N*
The digital repository takes care of transforming the submission information packages (SIPs) into archival information packages (AIPs).	*N*
The digital repository guarantees the storage and readability of the	*N*

archival information packages (AIPs).	
It is expected that the Ingest Functional Entity will coordinate the updates between Data Management and Archival Storage and provide appropriate coordination and error recovery. The AIP should first be stored in Archival Storage. The confirmation of that operation will include a unique identification to retrieve that AIP from Storage. This identifier should be merged into the Package Description prior to the addition of the Collection Description to Data Management.	*O*
The path-dependent nature of preservation decision making means that decisions made at any time shape future conditions and determine the range of future choices. Diagram showing traditional preservation path v. Digital Preservation Lifestyle can be found in 2.2.1 (p.29).	*P*
Strategies for overcoming problems with supply of preserved digital materials including lowering barriers to preserve; giving incentives to private parties to preserve in the public interest and imposing mandates to preserve.	*P*
There is an urgent need in all sectors of digital creation—public and private, cultural and scientific—for support in the near-term to model and test robust preservation strategies. All stakeholder communities must provide leadership and accept responsibility for the development of a common digital preservation infrastructure that is sustainable for generations to come.	*P*

3.3.3.1 Physical data preservation and storage

Data repository should implement the following:

- Preservation policy / Policy for maintaining documents on central file servers
- Strategy for backup / multiple copies
- Recovery provisions
- Risk management techniques used to inform the strategy
- Checks on the consistency of the archive (checksums)
- Policy to handle and monitor deterioration of storage (actively monitor the integrity of AIPs, record and report all incidents of data corruption or loss and take steps to repair or replace corrupt or lost data)
- Linking services which enable finding uniquely identified objects regardless their physical location
- Defined security levels
- Agreement with a security company to provide assured physical security services.

Recommendation	Found in
The repository shall define, collect, track, and appropriately provide its information integrity measurements.	*T*
The repository shall have a system of reliable linking/resolution services in order to find the uniquely identified object, regardless of its physical location [e.g. by means of naming conventions].	*T*

The repository shall provide an independent mechanism for verifying the integrity of the repository collection/content [in order] to enable the audit of the integrity of the collection as a whole.	*T*
The repository shall have specifications for how the AIPs are stored down to the bit level.	*T*
The repository shall actively monitor the integrity of AIPs [e.g. with the help of checksums].	*T*
The repository shall have effective mechanisms to detect bit corruption or loss.	*T*
The repository shall record and report to its administration all incidents of data corruption or loss, and steps shall be taken to repair/replace corrupt or lost data.	*T*
The repository shall manage the number and location of copies of all digital objects […] in order to assert that the repository is providing an authentic copy of a particular digital object.	*T*
Implement and monitor checksums	*S*
Data repository should implement the following: • *Preservation policy* • *Strategy for backup / multiple copies* • *Recovery provisions* • *Risk management techniques used to inform the strategy* • *Checks on the consistency of the archive* • *Defined security levels* • *Policy to handle and monitor deterioration of storage*	*S*
The digital repository has technical control of the digital objects in order to carry out long-term preservation measures.	*N*
What strategy is used for storage? (in-house, external, in-house under external support)?	*E*
The Receive Data function receives a storage request and an AIP from Ingest and moves the AIP to permanent storage within the Archive. The transfer request may need to indicate the anticipated frequency of utilization of the Data Objects making up the AIP in order to allow the appropriate storage devices or media to be selected for storing the AIP. This function will select the media type, prepare the devices or volumes, and perform the physical transfer to the Archival Storage volumes. Upon completion of the transfer, this function sends a storage confirmation message to Ingest, including the storage identification of the AIPs.	*O*
Generating multiple copies of an item as part of an institution's storage and preservation policy	*D*
Policies for maintaining documents on central file server	*D*
Strategies for migrating to larger file server before full capacity is reached	*D*
Transfer strategy: ♦ Establish service level agreement with third-party security company to provide assured physical security services	*B*
Transfer strategy:	*B*

♦ Acquire insurance against failure of hardware systems	
One of the choices available to decision makers in designing suitable preservation strategies is to determine who selects what to preserve.	*P*

3.3.3.2 Preservation strategy

Take actions to ensure that the information stored in the data-archive remains accessible to, and understandable by, the Designated Community over the long term, even if the original computing environment becomes obsolete.

Plan for preservation throughout the curation lifecycle of digital material. This would include plans for management and administration of all curation lifecycle actions. Use appropriate standards for metadata to ensure adequate description and control over long term. Monitor developments in standards, formats, hardware, software and storage technologies, user communities and reservation requirements. Stay aware of who you are keeping the data for and what your target audience's needs are.

Recommendation	Found in
Assign administrative, descriptive, technical, structural and preservation metadata, using appropriate standards, to ensure adequate description and control over the long-term.	*C*
Plan for preservation throughout the curation lifecycle of digital material. This would include plans for management and administration of all curation lifecycle actions.	*C*
Make a start on selection and appraisal from as early a point as possible (e.g., apply the new NERC criteria for identifying valuable data sets at the project plan stage).	*C*
Know who you are keeping it the data for and what you want them to be able do with it. This may affect the way you keep it and what you keep.	*C*
Conversely, know what you need to dispose of. Destruction is often vital to ensure compliance with legal requirements.	*C*
Re-appraisal can take place before ingest so review what you have and what you need to keep before depositing it to long-term storage.	*C*
Pin down the significant properties of your data and communicate them – make sure that the people carrying out preservation actions know what they are. This might be through metadata or other means.	*C*
The repository shall have documented preservation strategies relevant to its holdings.	*T*
The repository shall have mechanisms to change its preservation plans as a result of its monitoring activities.	*T*
Data repository should have a preservation policy (see also 2.2.3.3.1)	*S*
Provide procedural documentation for archiving data, including:	*S*

• *Workflows* • *Decision making process* • *Skills of employees* • *Types of data* • *Selection process* • *Measures regarding data that fall outside the mission* • *Guarding privacy of subjects* • *Information to data producers about handling the data*	
The digital repository has a strategic plan for its technical preservation measures (preservation planning).	*N*
What strategy is used for software management? • *support by the software supplier* • *support by a third party* • *self-support (i.e. in-house)* • *support by a user and developer community*	*E*
Repository must maintain understanding of all structural (e.g. file encoding) standards and formats	*E*
Repository must maintain understanding of contemporary and emerging hardware, software and storage technologies	*E*
Repository must maintain understanding of identified user communities and their associated competences and knowledge base	*E*
Repository must maintain understanding of preservation requirements for each stored information asset or class of information	*E*
Repository must maintain, exercise and evaluate preservation strategies capable of meeting specific preservation targets	*E*
Repository must maintain and exercise appropriate appraisal policies to determine which information must continue to be preserved	*E*
Fundamentally, approaches to information preservation in the face of changing technologies and Designated Community requirements require digital migration [...] Digital migration of an AIP can include: – *Copying Content Data Object or Representation Information bits to new media* – *Altering or adding to Content Data Object or Representation Information bits* – *Altering or adding to PDI bits* – *Altering or adding to operational software whose role is essential to Content Information preservation (i.e., it is part of Representation Information)* - *Altering or adding to the bits that make up the AIP's Packaging Information*	*O*
Procedures to update collection management databases	*D*
Targets for accessioning tasks and timescales for their completion	*D*
Written policies and guidelines, including selection policy for materials to be migrated	*D*
Quality control procedures	*D*
Rigorous documentation of migration procedure	*D*

Preservation metadata and documentation	*D*
Migrate data whenever there is a software upgrade or a new software application is installed	*D*
Ensure the migration results in little or no loss in content or context	*D*
Employ strict quality control procedures that may include testing the migration programme with a sample of records or bit/byte or checksum comparisons of migrated and original data.	*D*
Retain copies of the digital resource in its original format whenever some information or presentation of the resource may be lost or modified in migration	*D*
The Preservation Planning Functional Entity […] provides the services and functions for monitoring the environment of the OAIS, providing recommendations and preservation plans to ensure that the information stored in the OAIS remains accessible to, and understandable by, the Designated Community over the Long Term, even if the original computing environment becomes obsolete. Preservation Planning functions include evaluating the contents of the Archive and periodically recommending archival information updates, recommending the migration of current Archive holdings, developing recommendations for Archive standards and policies, providing periodic risk analysis reports, and monitoring changes in the technology environment and in the Designated Community's service requirements and Knowledge Base. Preservation Planning also designs Information Package templates and provides design assistance and review to specialize these templates into SIPs and AIPs for specific submissions. Preservation Planning also develops detailed Migration plans, software prototypes and test plans to enable implementation of Administration migration goals.	*O*
Tolerance strategy: *♦ Implement policy to define the parameters of acceptable loss resulting from preservation activities*	*B*
Tolerance strategy: *♦ Implement redundant backup storage*	*B*
Actions recommended for sustainability – Scholarly discourse - There are particular needs to align preservation incentives among commercial and non-profit providers; ensure handoffs between commercial publishers and stewardship organizations in the interest of long-term preservation of the scholarly record; and address the free-rider problem. Clarification of the long-term value of emerging genres of digital scholarship, such as academic blogs and grey literature, is a high priority. Research and education institutions, professional societies, publishers, libraries, and scholars all have leading roles to play in creating sustainable preservation strategies for the materials that are valuable to them.	*P*
Actions recommended for sustainability – Research data - there are few robust systems for making decisions about what to preserve; and there is often a lack of coordination of roles, responsibilities, and funding sources among those best positioned to preserve data (researchers) and the preservation infrastructure (curation and archiving services) that should	*P*

support them. Research and education institutions, professional societies, archives, researchers, and the funding agencies that support data creation all have leading roles to play in creating sustainable preservation strategies.	
Actions recommended for sustainability – commercially owned cultural content - There are well-established preservation and access strategies undergoing fundamental changes as a result of new information technologies. This includes the creation, distribution, and consumption of cultural content, most evident in the emergence of interactive genres such as games and the creation of a long tail of use and reuse. As a result, there may be two forms of benefits— commercial and cultural, or private and public—that compete with one another. When that occurs, proxy organizations must step in to represent the public interest. Leading players in preserving this content include private creators, owners, and trade associations, stewardship organizations, regulatory authorities, and leading national and international institutions that can sponsor public-private partnerships to ensure the long-term access to our digital cultural heritage.	*P*
Actions recommended for sustainability – collectively created Web content - The Web environment is marked by great dynamism, uncertainty about long-term value of digital content, and obscure ownership and rights issues for many collectively produced Web assets. The priority here is for stewardship organizations, content creators, hosting sites, platform providers, and users to model and test preservation strategies, and to provide clarification about long-term value and selection criteria.	*P*
Engaging context experts to identify context-specific attributes is crucial for developing a sound preservation strategy.* **features specific to a data type or user community that constrain choices for preservation.*	*P*
Scholarly discourse – recommendation 1- libraries, scholars, and professional societies should develop selection criteria for emerging digital genres in scholarly discourse, and prototype preservation and access strategies to support them.	*P*
Scholarly discourse – recommendation 2-publishers reserving the right to preserve should partner with third-party archives libraries to ensure long-term preservation.	*P*
Scholarly discourse – recommendation 3-Scholars should consider granting nonexclusive rights to publish and preserve, to enable decentralised and distributed preservation of emerging scholarly discourse.	*P*
Scholarly discourse – recommendation 4 – Libraries should create a mechanism to organize and clarify their governance issues and responsibilities to preserve monographs and emerging scholarly discourse along lines similar to those for e-journals.	*P*
Scholarly discourse – recommendation 5 – All open-access strategies that assume the persistence of information over time must consider provisions for the funding of preservation.	*P*
Research data – recommendation 1- Each domain, through professional	*P*

societies or other consensus-making bodies, should set priorities for data selection, level of curation, and length of retention.	
Research data - recommendation 2- Funders should impose preservation mandates, when appropriate. When mandates are imposed, funders should also specify selection criteria, funds to be used, and responsible organizations to provide archiving.	*P*
Research data - recommendation 3 - Funding agencies should explicitly recognize "data under stewardship" as a core indicator of scientific effort and include this information in standard reporting mechanisms.	*P*
Research data - recommendation 4 - Preservation services should reduce curation and archiving costs by leveraging economies of scale when possible.	*P*
Research data - recommendation 5 - Agreements with third-party archives should stipulate processes, outcomes, retention periods, and handoff triggers.	*P*
Commercially owned cultural content - recommendation 1 - Leading cultural organizations should convene expert communities to address the selection and preservation needs of commercially owned cultural content and digital orphans.	*P*
Commercially owned cultural content - recommendation 2 - Regulatory authorities should bring current requirements for mandatory copyright deposit into harmony with the demands of digital preservation and access.	*P*
Commercially owned cultural content - recommendation 3 - Regulatory authorities should provide financial and other incentives to preserve privately held cultural content in the public interest.	*P*
Commercially owned cultural content - recommendation 4 - Leading stewardship organizations should model and test mechanisms to ensure flexible long-term public-private partnerships that foster cooperative preservation of privately held materials in the public interest.	*P*
Collectively produced web content- recommendation 1 - Leading stewardship organizations should convene stakeholders and experts to address the selection and preservation needs of collectively produced Web content.	*P*
Collectively produced web content- recommendation 2 - Creators, contributors, and host sites could lower barriers to third-party archiving by using a default license to grant nonexclusive rights for archiving.	*P*
Collectively produced web content- recommendation 3 - Regulatory authorities should create incentives, such as preservation subsidies, for host sites to preserve their own content or seek third party archives as preservation partners.	*P*
Collectively produced web content- recommendation 4 - Regulatory authorities should take expeditious action to reform legislation to grant authority to stewardship institutions to preserve at-risk Web content.	*P*
Collectively produced web content- recommendation 5 - Leading stewardship organizations should develop partnerships with one or more major content providers to explore the technical, legal, and financial dimensions of long-term preservation.	*P*

Action agenda for Scholarly Discourse – 1. Libraries, scholars, and professional societies should develop selection criteria for emerging genres in scholarly discourse, and prototype preservation and access strategies to support them. 2. Publishers reserving the right to preserve should partner with third-party archives or libraries to ensure long-term preservation. 3. Scholars should consider granting nonexclusive rights to publish and preserve, to enable decentralized and distributed preservation of emerging scholarly discourse. 4. Libraries should create a mechanism to organize and clarify their governance issues and responsibilities to preserve monographs and emerging scholarly discourse along lines similar to those for e-journals. 5. All open-access strategies that assume the persistence of information over time must consider provisions for the funding of preservation.	*P*
Action agenda for Research Data – 1. Each domain, through professional societies or other consensus-making bodies, should set priorities for data selection, level of curation, and length of retention. 2. Funders should impose preservation mandates, when appropriate. When mandates are imposed, funders should also specify selection criteria, funds to be used, and responsible organizations to provide archiving. 3. Funding agencies should explicitly recognize "data under stewardship" as a core indicator of scientific effort and include this information in standard reporting mechanisms. 4. Preservation services should reduce curation and archiving costs by leveraging economies of scale when possible. 5. Agreements with third-party archives should stipulate processes, outcomes, retention periods, and handoff triggers.	*P*
Action agenda for Commercially owned cultural content – 1. Leading cultural organizations should convene expert communities to address the selection and preservation needs of commercially owned cultural content and digital orphans. 2. Regulatory authorities should bring current requirements for mandatory copyright deposit into harmony with the demands of digital preservation and access. 3. Regulatory authorities should provide financial and other incentives to preserve privately held cultural content in the public interest. 4. Leading stewardship organizations should model and test mechanisms to ensure flexible long-term public-private partnerships that foster cooperative preservation of privately held materials in the public interest.	*P*
Action agenda for collectively produced web content – 1. Leading stewardship organizations should convene stakeholders and experts to address the selection and preservation needs of collectively produced Web content. 2. Creators, contributors, and host sites could lower barriers to third-party archiving by using a default license to grant nonexclusive rights for archiving. 3. Regulatory authorities should create incentives, such as preservation subsidies, for host sites to preserve their own content or seek third-party archives as preservation partners. 4. Regulatory authorities should take expeditious action to reform legislation to grant authority to stewardship institutions to preserve at-risk Web content. 5. Leading stewardship organizations should develop partnerships with one or more major content providers to explore the technical, legal, and financial dimensions of long-term preservation.	*P*

3.3.3.3 Version control/change procedures

Maintain a documented history of the changes and provide an audit trail to original data versions. Compare properties of different versions of a file. Define a versioning strategy and document processes used for changes. Define a strategy for data changes and make also the data producers and consumers aware of this strategy.

Recommendation	Found in
The repository shall have a documented history of the changes to its operations, procedures, software, and hardware […] in order to provide an 'audit trail' through which stakeholders can identify and trace decisions made by the repository.	*T*
Documentation shall describe any processes used for changes to [unique persistent] identifiers.	*T*
The repository shall have procedures for all actions taken on AIPs […] in order to ensure that any actions performed against an AIP do not alter the AIP information in a manner unacceptable to its Designated Communities.	*T*
The repository shall be able to demonstrate that any actions taken on AIPs were compliant with the specification of those actions.	*T*
The repository shall have mechanisms in place to ensure any/multiple copies of digital objects are synchronized.	*T*
Define a versioning strategy	*S*
Define a strategy for data changes and make also data producers aware of this strategy	*S*
Maintain provenance data and related audit trails	*S*
Compare essential properties of different versions of a file	*S*
Procedure to maintain links to metadata and other datasets	*S*
Digital Migrations that require some changes to the Content Information or PDI are referred to as Transformations. These changes will be to some of the bits in the Content Information or PDI with corresponding changes in the associated Representation Information. In all cases the intent is to provide maximum information preservation. The resulting AIP is intended to be a full replacement for the AIP that is undergoing Transformation. The new AIP qualifies as a new AIP Version of the previous AIP. The first version of the AIP is referred to as the original AIP and may be retained for verification of information preservation.	*O*
Procedures for updating, and managing versions or editions of an item.	*D*

3.3.4 Dissemination

Know what you want users to be able to do with your data. Specify minimum information requirements to enable the designated community to discover and identify data (enough context). Ensure transformation of AIPs into Dissemination Information Packages (DIPs) so that the DIPs remain traceable to the originals, with evidence supporting the authenticity of data.

Recommendation	Found in
Plan for what you think you'll need to keep to support your research findings. What is the minimum you'll need to support your findings over time?	*C*
Know what you want users to be able to do with your data and for how long.	*C*
Pin down and communicate the significant properties of your data.	*C*
Ensure that you provide enough context to ensure that your data can be located and used – either by the originally designated user community or new users over time.	*C*
The repository shall specify minimum information requirements to enable the Designated Community to discover and identify material of interest.	*T*
The repository shall follow policies and procedures that enable the dissemination of digital objects that are traceable to the originals, with evidence supporting their authenticity […]. This is necessary to establish an auditable chain of authenticity from the AIP to disseminated digital objects.	*T*
The digital repository defines its dissemination information packages (DIPs).	*N*
The digital repository ensures transformation of archival information packages (AIPs) into dissemination information packages (DIPs).	*N*
The OAIS may choose to provide this API as an implementation alternative to the production and delivery of a physical DIP for dissemination. This type of service allows the Consumer, as a client, to develop applications that appear to directly access the AIPs. This sort of access could be very useful for applications such as data mining where the creation and shipping of DIPs containing large AICs is impractical. This API could allow an application to virtually navigate through an AIC, deliver the bits of the Content Data Object of selected AIUs to the application and identify locations for obtaining associated Representation Information and PDI. However, as technology evolves, the OAIS moves to new hardware, new media, and new operating systems. If the OAIS wishes to maintain the same API for its Consumers, it will need to provide a 'wrapper' around part of its new infrastructure to match its services to the established API. The API will need to be adequately documented and tested to ensure it correctly delivers the AIU Content Information using this new Access Software. This approach should not result in any changes to software developed by the Consumer community. When the API is applicable across a wide range of AIUs in the OAIS or there are a significant number of Consumer applications based on the API, this wrapping approach is clearly	*O*

feasible and may result in a favorable cost/benefit ratio to the OAIS and its Designated Community. The 'Layered Model of Information' presented in annex E of this document further describes some potentially standard APIs.	
Each activity is linked to one or more key assets of the repository. Furthermore, each activity is supported by a number of technological systems and solutions that members of staff rely upon. Technology, software and various support systems are included in the assets category in this self-audit. For example a web server may be used to offer one or more key repository services, including the dissemination of digital content to users by the user services department.	*B*
Strategies for overcoming problems with demand for preserved digital materials including use or formulation of proxy organisations; benefits for future users & avoiding irreversible loss.	*P*

3.3.4.1 Visibility

Provide easy to use search facilities. Keep this in mind when you structure and define your metadata.

Recommendation	Found in
Provide search facilities, OAIS harvesting, deep searching	*S*
Carry out promotional activities	*S*
The Package Description is not required for the Long Term Preservation of the Content Information but is needed to provide visibility and access into the contents of an Archive. The contents of the Package Description are highly dependent on the structure of the Content Information and PDI it describes.	*O*
Tolerance strategy: *♦ Define policy to commit to the delivery of minimal service levels, incorporating breathing space for tolerable downtime or information non-availability*	*B*

3.3.4.2 Availability and accessibility

Ensure accessibility of data, use persistent identifiers, support data sharing and re-use. Use access agreements. [see also chapter 3.1.5 Access Policy]

Recommendation	Found in
Ensure that data is accessible to both designated users and reusers, on a day-to-day basis. This may be in the form of publicly available published information. Robust access controls and authentication procedures may be applicable.	*C*
Offer persistent identifiers	*S*

Promote data sharing and reuse	*S*
Create and Implement A Repository Access Policy taking into account: copyright law, law on dissemination of personal or commercially-sensitive data, National security, Libel, obscenity, hate-speech and blasphemy laws, Specific contractual restrictions imposed in deposit agreements	*E*
Specify and fulfill technical requirements for dissemination and access	*E*
The Coordinate Access Activities function provides one or more interfaces to the information holdings of the Archive. This interface will normally be via computer network or dial-up link to an on-line service, but might also be implemented in the form of a walk-in facility, printed catalog ordering service, or fax-back type service. Three categories of Consumer requests are distinguished: query requests, which are executed in Data Management and return immediate query responses for presentation to the user; report requests, which may require a number of queries and produce formatted reports for delivery to the Consumer; and orders, which may access either or both Data Management and Archival Storage to prepare a formal Dissemination Information Package (DIP) for on- or off-line delivery. The Generate DIP function accepts a dissemination request, retrieves the AIP from Archival Storage, and moves a copy of the data to a temporary storage area for further processing. This function also transmits a report request to Data Management to obtain Descriptive Information needed for the DIP. The Deliver Response function handles both on-line and off-line deliveries of responses (DIPs, query responses, reports and assistance) to Consumers. For on-line delivery, it accepts a response from Coordinate Access Activities and prepares it for on-line distribution in real time via communication links.	*O*
Several choices need to be made by decision makers when designing suitable preservation strategies including:-Who benefits from use of the preserved asset? -Who selects what to preserve?-Who owns the asset?-Who preserves the asset?-Who pays?	*P*
Because a significant portion of digital assets with long-term value are privately created and owned, strong and robust mechanisms for partnering between public and private sectors are necessary to ensure the long-term public interest in private assets.	*P*
Organizations can help dilute the risk of free riding. A commonly used mechanism for avoiding free-rider consumption is to make access to or use of the goods exclusive to a group, usually a fee-paying group. Tiered access is a standard remedy for levelling the costs of access to institutions with different financial means without prejudicing access by the community at large. Excluding access obviously does not work for those assets that communities wish to share freely.	*P*
Condition for sustainability: timely actions to ensure access	*P*

3.3.4.3 Tools and interfaces

Provide easy to use deposit forms. Watch and if possible, participate in the development of shared standards, tools and suitable software.

Recommendation	Found in
Maintain a watch on appropriate community activities, and participate in the development of shared standards, tools and suitable software.	*C*
Use of deposit forms (see also 2.2.3.1.1)	*S*
The OAIS Consumer interface in Access provides one or more Content Information IDs, with associated name spaces, to assist in identifying a particular Content Information object of interest. One or more of these Content Information IDs will be included in the PDI. *Reference Information associated with that Content Information object. The Descriptive Information in Data Management will map each of these IDs to the same AIP ID. The Access Function uses this information to obtain the AIP ID and gives it to Archival Storage to retrieve the associated AIP.*	*O*

3.3.4.4 Monitoring, review and feedback

The repository should periodically evaluate usage statistics and watch for services going out of favour with the users. Giving the users a way to request new features could also help the repository staying current. Monitor also the availability of data and checksums.

Monitor the environment and technological trends, evaluate which could be potentially harmful, and adapt.

Recommendation	Found in
The repository shall monitor its organizational environment to determine when to execute its succession plan, contingency plans, and/or escrow arrangements.	*T*
The repository shall commit to a regular schedule of self-assessment and external certification.	*T*
The repository shall track and manage intellectual property rights and restrictions on use of repository content as required by deposit agreement, contract, or license.	*T*
The repository shall ensure that the Content Information of the AIPs is understandable for their Designated Community at the time of creation of the AIP. In particular, the […] repository shall have [and execute] a documented process for testing understandability for their Designated Communities of the Content Information of the AIPs at their creation.	*T*
The repository shall bring the Content Information of the AIP up to the required level of understandability if it fails the understandability testing.	*T*

The repository shall have mechanisms in place for monitoring its preservation environment […] so that the repository can react to changes and thereby ensure that the preserved information remains understandable and usable by the Designated Community.	*T*
The repository shall have mechanisms in place for monitoring and notification when Representation Information is inadequate for the Designated Community to understand the data holdings.	*T*
The repository shall have mechanisms for creating, identifying or gathering any extra Representation Information required.	*T*
The repository shall provide evidence of the effectiveness of its preservation activities […] to assure the Designated Community that the repository will be able to make the information available and usable over the mid-to-long-term.	*T*
The repository shall log and review all access management failures and anomalies.	*T*
The repository shall record and act upon problem reports about errors in data or responses from users.	*T*
The repository shall employ technology watches or other technology monitoring notification systems.	*T*
The repository shall have procedures in place to monitor and receive notifications when hardware technology [or software] changes are needed.	*T*
The repository shall have procedures in place to evaluate when changes are needed to current hardware [or software].	*T*
Monitor availability of data	*S*
Monitor chekcsums	*S*
The repository should periodically evaluate usage statistics and watch for services going out of favor with the users. Giving the users a way to request new features could also help the repository staying current.	*E*
Technological changes can greatly affect the business model of a repository, and at times come about very quickly. Repositories need to watch the technological trends, evaluate which could be potentially harmful, and adapt.	*E*
Regular review of preservation priorities is important.	*P*
A principal outcome from the successful completion of this stage is a risk register with risk management features included. The risk management exercise cannot and should not stop with the creation of a risk register. Ongoing review and monitoring is essential to ensure that the risk management plan remains relevant. Factors affecting the likelihood and consequences of a risk may change, as may the factors that affect the suitability or cost of the risk mitigation measure. Also the repository's business, regulatory or social context will change over time and therefore some risks may disappear or become less important while other new risks may emerge. It is therefore necessary to repeat the risk management cycle regularly and review the target outcomes when their deadlines are reached.	*B*
The information on activities and assets of the repository should be made	*B*

available to the auditor in a complete and unabridged form. If this proves difficult, senior management should be contacted and the necessary authorisations acquired to gain access to the required information. Further usage restrictions can be agreed for the list of activities and assets that is created in this Stage in order to protect the sensitive information it may contain. Not having or not being able to access the information on what the repository is doing can be considered as a considerable risk in and of itself.	
Although documentation is crucial, it is not the only source of evidence that should be pursued. As a result of the DCC audits we have also identified experimental, testimonial and observational evidence. Assuming that a single individual is completing the self-audit process, it is unlikely that they will themselves have a comprehensive knowledge of every aspect of the repository's activities. With this in mind it is vital that the process should be an open one where a wide range of staff are capable of contributing their thoughts and noting risks for inclusion within the overall risk register. Individuals that should be consulted include the repository overall administrators; hardware and software administrators; and officers responsible for the core functions of ingest, archiving, preservation, documentation and access. Everyone from the head of the organisation to the janitor or cleaning staff can provide insights into the process of repository management. These additional contributions need not be too lengthy, but must be considered if the outcomes of the process are to be representative of the repository as a whole.	*B*

4. Appendix

4.a DRAMBORA

Digital Repository Audit Method Based on Risk Assessment

Developed jointly by the Digital Curation Centre (DCC) and Digital Preservation Europe (DPE), the Digital Repository Audit Method Based on Risk Assessment (DRAMBORA) represents the main intellectual outcome of a period of pilot repository audits undertaken by the DCC throughout 2006 and 2007. It presents a methodology for self-assessment, encouraging organisations to establish a comprehensive self-awareness of their objectives, activities and assets before identifying, assessing and managing the risks implicit within their organisation. As opposed to other data curation approaches, DRAMBORA focuses on self-assessment and risk-management instead of setting rules for data ingestion. Based on the analysis of the results of the self-assessment, implicit shortcomings are expressed in terms of organisational risk. These should be tempered by the introduction of appropriate risk management mechanisms.

Source: http://www.repositoryaudit.eu/

4.a.1 Organizational Framework

4.a.1.1 Purpose and Requirements

Recommendation	Found in
DRAMBORA enables archives and repositories to better fulfil their responsibilities and achieve their strategic goals by: identifying the strengths and weaknesses of their digital repository; and assisting them to plan effectively to mitigate these risks.	

4.a.1.1.1 Scope and objectives

Recommendation	Found in
[…] auditors must identify the repository's mandate, which, it is anticipated, will be described in an organisational mission statement or enacting documentation. The subsequent task requires auditors to identify, within that mandate, each organisational goal and objective relevant to the repository.	*5.7.2*

4.a.1.1.2 Collection policy

Recommendation	Found in
The policies DRAMBORA defines for archives in order to be fit for data	

preservation are: *1. Mandate & Commitment to Digital Object Maintenance* *2. Organisational Fitness* *3. Legal & Regulatory Legitimacy* *4. Efficient & Effective Policies* *5. Adequate Technical Infrastructure* *6. Acquisition & Ingest* *7. Preservation of Digital Object Integrity, Authenticity & Usability* *8. Metadata Management & Audit Trails* *9. Dissemination* *10.Preservation Planning & Action*	

4.a.1.1.3 Criteria for evaluating data

Recommendation	Found in
Risks are assessed according to the following three characteristics: *1. impact: the potential impact that the risk would have should occur* *2. impact expression: the way in which the negative effects of the risk's occurrence manifest themselves* *3. probability: the likelihood of or frequency with which the risk will occur.*	

4.a.1.2 Legal and Regulatory Framework

Recommendation	Found in
A broad definition of the repository's regulatory framework is assumed, incorporating acts or provisions with both external and internal origins. Relevant extrinsic commitments and influences include statutory legislation and statutory instruments, global or business-related regulations, de facto or established standards and codes of practice. Internally arising commitments may be traceable to contracts, policies, strategic planning, or accepted business norms.	*5.8.1*

4.a.1.3 Funding and Resource Planning

Recommendation	Found in
In order to avoid risks of loss of data and guarantee reliable data preservation, archives and repositories should adhere to the following strategies: *♦ Develop self-sustainability with charged-for services* *♦ Seek assurances of level of budget* *In the event of risk's execution:* *♦ Solicit additional funding to enable achievement of organisational*	*5.13.1.1*

objectives ♦ *Revise objectives if funding stream is insufficiently flexible* ♦ *Maintain residual fund where possible to meet shortfalls*	

4.a.1.4 Long-Term Preservation Policy

Recommendation	Found in
For each risk identified, the final stage of the self-assessment process is to define appropriate risk management measures and targets[...]. For each risk details of treatment, avoidance measures, and anticipated outcomes should be recorded and monitored over time. A timeframe for regular reassessment should also be recorded at this stage to ensure that new risks are identified and mitigated as the repository matures.	

4.a.1.5 Access Policy

Recommendation	Found in
To guarantee reliable long-term data access archives and repositories should adhere to the following strategies: *· Define policies describing available information delivery services and communicate these to the user community* *· Implement appropriate systems to meet delivery policy requirements* *· Establish sufficiently robust technical infrastructure to satisfy demands of proposed delivery services*	*7.3.8*

4.a.2 Technological Environment

4.a.2.1 IT Architecture

Recommendation	Found in
DRAMBORA itself does not describe any architectural constructs. The nestor criteria catalogue, and ISO 27001:2005 is mentioned briefly.	*5.9.9*

4.a.2.2 Standards and Formats

Recommendation	Found in
Standards that are recommended by DRAMBORA: ♦ *ISO 9000:2000 Quality Management Systems Series* ♦ *ISO 27001:2005 Information technology — Security techniques — Information security management systems — Requirements*	*5.8.9*

4.a.2.3 Security and Risk Management / Media Monitoring and Refreshing Strategy

Recommendation	Found in
Risks are assessed according to the following three characteristics: *1. impact: the potential impact that the risk would have should occur* *2. impact expression: the way in which the negative effects of the risk's occurrence manifest themselves* *3. probability: the likelihood of or frequency with which the risk will occur.*	
♦ *highlight what digital resources are at risk within their organisation;* ♦ *highlight the risks to which these digital resources were exposed;* ♦ *highlight the risks to organisations posed by threats to digital resources (e.g. reputation, cessation of business);* ♦ *categorise and prioritise risks in order to facilitate their management;* ♦ *facilitate communication within the organisation about areas of risk;* ♦ *stimulate risk management strategy development.*	4.3.1

4.a.3 Data Curation

4.a.3.1 Pre-Ingest Function

Recommendation	Found in
Within DRAMBORA, digital curation is characterised as a risk-management activity; the job of digital curator is to rationalise the uncertainties and threats that inhibit efforts to maintain digital object authenticity and understandability, transforming them into manageable risks.	

4.a.3.1.1 Information and guidance given to data producer

Recommendation	Found in
N.A.	

4.a.3.2 Ingest Function

Recommendation	Found in
DRAMBORA does not give any specific guidelines regarding ingest functions. It focuses on beforehand risk assessment and the audit process.	

4.a.3.2.1 Information and documentation from data producer

Recommendation	Found in
Data producers are to be aware of the quality assurance process, mentioned later in the report.	

4.a.3.2.2 Quality assurance and data checking

Recommendation	Found in
Repository requires staff to undertake quality assurance procedures for each object ingested, which takes on average 10 minutes, although an additional policy states that ingest should be completed in 10 minutes.	7.4
Before assessing their repository's aims or shortcomings, it is vital that repository administrators define the purpose and scope of what they are aiming to achieve with the self-assessment process. To assist with this, DRAMBORA Interactive permits users to register details about their own repository to help determine their self-assessment profile. DRAMBORA Interactive also helps users to identify, register, and provide access to additional institutional staff members who will contribute to the various steps of the self-assessment process. DRAMBORA Interactive features robust security features to ensure that sensitive corporate information is protected against non-authorised access.	
Self-assessment is a key point in DRAMBORA. The self-assessment is divided into several phases, the first of which focuses on attaining a comprehensive overview of the objectives [...] and activities of the organisation supported by adequate documentation. In this phase, questions about the institution's mandate, its organisational constraints (legal, political, and/or technical limitations or obligations), overarching objectives, and business activities are explored.	

4.a.3.2.3 Data documentation and enhancement

Recommendation	Found in
N.A.	

4.a.3.3 Archival Storage and Preservation

Recommendation	Found in
N.A.	

4.a.3.3.1 Physical data preservation and storage

Recommendation	Found in
Transfer strategy: *♦ Establish service level agreement with third-party security company to provide assured physical security services*	5.13.1.2
Transfer strategy: *♦ Acquire insurance against failure of hardware systems*	5.13.1.2

4.a.3.3.2 Preservation strategy

Recommendation	Found in

Tolerance strategy: *♦ Implement policy to define the parameters of acceptable loss resulting from preservation activities*	5.13.1.3
Tolerance strategy: *♦ Implement redundant backup storage*	5.13.1.3

4.a.3.3.3 Version control/change procedures

Recommendation	Found in
N.A.	

4.a.3.4 Dissemination

Recommendation	Found in
Each activity is linked to one or more key assets of the repository. Furthermore, each activity is supported by a number of technological systems and solutions that members of staff rely upon. Technology, software and various support systems are included in the assets category in this self-audit. For example a web server may be used to offer one or more key repository services, including the dissemination of digital content to users by the user services department.	5.9.1

4.a.3.4.1 Visibility

Recommendation	Found in
Tolerance strategy: *♦ Define policy to commit to the delivery of minimal service levels, incorporating breathing space for tolerable downtime or information non-availability*	*5.13.1.3*

4.a.3.4.2 Availability and accessibility

Recommendation	Found in
N.A.	

4.a.3.4.3 Tools and interfaces

Recommendation	Found in
N.A.	

4.a.3.4.4 Monitoring, review and feedback

Recommendation	Found in
A principal outcome from the successful completion of this stage is a risk register with risk management features included. The risk management exercise cannot and should not stop with the creation of a risk register. Ongoing review and monitoring is essential to ensure that the risk management plan remains relevant. Factors affecting the likelihood and consequences of a risk may change, as may the factors that affect the suitability or cost of the risk mitigation measure. Also the repository's business, regulatory or social context will change over time and therefore some risks may disappear or become less important while other new risks may emerge. It is therefore necessary to repeat the risk management cycle	5.12.3

regularly and review the target outcomes when their deadlines are reached.	
The information on activities and assets of the repository should be made available to the auditor in a complete and unabridged form. If this proves difficult, senior management should be contacted and the necessary authorisations acquired to gain access to the required information. Further usage restrictions can be agreed for the list of activities and assets that is created in this Stage in order to protect the sensitive information it may contain. Not having or not being able to access the information on what the repository is doing can be considered as a considerable risk in and of itself.	5.9.8
Although documentation is crucial, it is not the only source of evidence that should be pursued. As a result of the DCC audits we have also identified experimental, testimonial and observational evidence. Assuming that a single individual is completing the self-audit process, it is unlikely that they will themselves have a comprehensive knowledge of every aspect of the repository's activities. With this in mind it is vital that the process should be an open one where a wide range of staff are capable of contributing their thoughts and noting risks for inclusion within the overall risk register. Individuals that should be consulted include the repository overall administrators; hardware and software administrators; and officers responsible for the core functions of ingest, archiving, preservation, documentation and access. Everyone from the head of the organisation to the janitor or cleaning staff can provide insights into the process of repository management. These additional contributions need not be too lengthy, but must be considered if the outcomes of the process are to be representative of the repository as a whole.	5.2.4

4.b Appendix B DCC Cycle

The DCC Curation Lifecycle Model provides a graphical, high-level overview of the stages required for successful curation and preservation of data from initial conceptualisation or receipt through the iterative curation cycle.

You can use the model to plan activities within your organisation or consortium to ensure that all of the necessary steps in the curation lifecycle are covered.

It is important to note that the model is an ideal. In reality, users of the model may enter at any stage of the lifecycle depending on their current area of need. For instance, a digital repository manager may engage with the model for this first time when considering curation from the point of ingest. The repository manager may then work backwards to refine the support they offer during the conceptualisation and creation processes to improve data management and longer-term curation.

4.b.1 Organizational Framework

4.b.1.1 Purpose and Requirements

Recommendation	*Found in*
N.A.	

4.b.1.1.1 Scope and objectives

Recommendation	Found in
Make the 'ingest' process as straight-forward as possible and provide support and guidance wherever you can; automate processes if you can.	*DCC Ingest / DCC Store*
Don't be afraid to be critical when reviewing 'best practice' and recommended approaches. They might work for the specific scenario for which they were created but not for you. Do you know the criteria used to rate things like 'preferred' formats?	*DCC Preservation action*

4.b.1.1.2 Collection policy

Recommendation	Found in
Work together - researchers and information managers need to communicate regularly. Neither can do their job in isolation.	*DCC Create or Receive*
Return data which fails validation procedures for further appraisal and re-selection.	*DCC Reappraise*

5b.1.1.3 Criteria for evaluating data

Recommendation	*Found in*
N.A.	

4.b.1.2 Legal and Regulatory Framework

Recommendation	Found in
Identify any data protection requirements that you need to address in the course of your research and ensure that these are communicated to all staff.	*DCC Create or Receive*

4.b.1.3 Funding and Resource Planning

Recommendation	*Found in*
N.A.	

4.b.1.4 Long-Term Preservation Policy

Recommendation	Found in
Be aware of, and undertake management and administrative actions planned to promote curation and preservation throughout the curation lifecycle.	*DCC Curate and Preserve*
Work with researchers and information managers to develop policies and to identify realistic and implementable workflows.	*DCC Appraise and Select*
Appraise for the here and now but with an eye to the future.	*DCC Appraise and Select*

4.b.1.5 Access Policy

Recommendation	Found in
Ensure that any restrictions on access and use are communicated and respected.	*DCC Access, Use and Reuse*

4.b.2 Technological Environment

4.b.2.1 IT Architecture

Recommendation	Found in
Definition of data: *Data, any information in binary digital form, is at the centre of the Curation Lifecycle.* *This includes:* *Digital Objects: simple digital objects (discrete digital items such as text files, image files or sound files, along with their related identifiers and metadata) or complex digital objects (discrete digital objects made by combining a number of other digital objects, such as websites).* *Databases: structured collections of records or data stored in a computer system.*	*DCC Data*

4.b.2.2 Standards and Formats

Recommendation	Found in
Store the data in a secure manner adhering to relevant standards.	*DCC Store*
Create new data from the original, for example: by migration into a different format, or by creating a subset, by selection or query, to create newly derived results, perhaps for publication	*DCC Transform*
Migrate data to a different format. This may be done to accord with the storage environment or to ensure the data's immunity from hardware or software obsolescence.	*DCC Migrate*

4.b.2.3 Security and Risk Management / Media Monitoring and Refreshing Strategy

Recommendation	Found in
Dispose of data, which has not been selected for long-term curation and preservation in accordance with documented policies, guidance or legal requirements. *Typically data may be transferred to another archive, repository, data centre or other custodian. In some instances data is destroyed. The data's nature may, for legal reasons, necessitate secure destruction.*	*DCC Dispose*

4.b.3 Data Curation

4.b.3.1 Pre-Ingest Function

Recommendation	Found in
Conceive and plan the creation of data, including capture method and storage options.	*DCC Conceptualise*
Get into the habit of equating data curation with good research.	*DCC Conceptualise*
Know what your funding body expects you to do with your data and for how long. Assess your ability to be able to meet these expectations (i.e., do you need additional funding or staff?).	*DCC Conceptualise*
Determine intellectual property rights from the outset and ensure they are documented.	*DCC Conceptualise*
Identify any anticipated publication requirements (embargoes, restrictions on publishing over multiple sites).	*DCC Conceptualise*
Identify and document specific roles and responsibilities as early as possible.	*DCC Conceptualise*

4.b.3.1.1 Information and guidance given to data producer

Recommendation	Found in
Agree from an early stage any standards you will be making use of for content, syntax, and structure. Once these have been agreed, make sure they are communicated - both to other researchers on the project and to the data/information managers you will be working with. Provide training if necessary.	*DCC Create or Receive*
Identify data quality metrics as soon as possible and ensure that these are communicated and monitored.	*DCC Create or Receive*

4.b.3.2 Ingest Function

Recommendation	Found in
Create data including administrative, descriptive, structural and technical metadata. Preservation metadata may also be added at the time of creation.	*DCC Create or Receive*
Know who you are creating your data for and what you want them to be able to do (and not do) with it. Communicate this with others on the project.	*DCC Create or Receive*
Be realistic – strike a balance between what is sufficient and what is ideal based on your practical realities.	*DCC Create or Receive*
Transfer data to an archive, repository, data centre or other custodian. Adhere to documented guidance, policies or legal requirements.	*DCC Ingest*
Remember - ingest does not necessarily need to mean deposit in a data centre or repository but rather moving to a 'curated' environment – could be as simple as a specific folder on a shared drive.	*DCC ingest / DCC Store*

4.b.3.2.1 Information and documentation from data producer

Recommendation	Found in
Receive data, in accordance with documented collecting policies, from data creators, other archives, repositories or data centres, and if required assign appropriate metadata.	*DCC Create or Receive*
Make sure you know about any repository policies that might affect your deposit for long-term storage (i.e., what will they accept, are there preferred formats or normalisation processes).	*DCC Ingest / DCC Store*
Get a formal receipt (if possible) or an informal acknowledgement for closure and transfer of stewardship	*DCC Ingest / DCC Store*

4.b.3.2.2 Quality assurance and data checking

Recommendation	Found in
Ensure that your data meets minimum quality assurance metrics (based on intended use).	*DCC Appraise and Select*
Decide on who is responsible for final aspects of data quality assurance at the point of deposit (researcher, archive, information manager, etc…). Ensure that this final point of QA is communicated to all stakeholders.	*DCC Ingest / DCC Store*
Data quality is not absolute. Level of data quality and cleaning must be assessed by fitness for purpose. So, 'high quality' data for one user group	*DCC Ingest / DCC Store*

may be completely unsuitable for another user group.	

4.b.3.2.3 Data documentation and enhancement

Recommendation	Found in
Making use of archival standards like ISAD-G can be useful for hierarchical data description. So, talk to information managers at your institution for advice.	*DCC Ingest / DCC Store*

4.b.3.3 Archival Storage and Preservation

Recommendation	Found in
Evaluate data and select for long-term curation and preservation. Adhere to documented guidance, policies or legal requirements.	*DCC Appraise and Select*
Undertake actions to ensure long-term preservation and retention of the authoritative nature of data. Preservation actions should ensure that data remains authentic, reliable and usable while maintaining its integrity. Actions include data cleaning, validation, assigning preservation metadata, assigning representation information and ensuring acceptable data structures or file formats.	*DCC Preservation action*
Document preservation actions so that people know what has been done to the data over time.	*DCC Preservation action*

4.b.3.3.1 Physical data preservation and storage

Recommendation	*Found in*
N.A.	

4.b.3.3.2 Preservation strategy

Recommendation	Found in
Assign administrative, descriptive, technical, structural and preservation metadata, using appropriate standards, to ensure adequate description and control over the long-term.	*DCC Description and Representation Information*
Plan for preservation throughout the curation lifecycle of digital material. This would include plans for management and administration of all curation lifecycle actions.	*DCC Preservation Planning*
Make a start on selection and appraisal from as early a point as possible (e.g., apply the new NERC criteria for identifying valuable data sets at the project plan stage).	*DCC Appraise and Select*
Know who you are keeping it the data for and what you want them to be able do with it. This may affect the way you keep it and what you keep.	*DCC Appraise and Select/ DCC Preservation action*
Conversely, know what you need to dispose of. Destruction is often vital to ensure compliance with legal requirements.	*DCC Appraise and Select*
Re-appraisal can take place before ingest so review what you have and what you need to keep before depositing it to long-term storage.	*DCC Appraise and Select*
Pin down the significant properties of your data and communicate them – make sure that the people carrying out preservation actions know what they are. This might be through metadata or other means.	*DCC Preservation action*

4.b.3.3.3 Version control/change procedures

Recommendation	Found in
N.A.	

4.b.3.4 Dissemination

Recommendation	Found in
Plan for what you think you'll need to keep to support your research findings. What is the minimum you'll need to support your findings over time?	*DCC Appraise and Select*
Know what you want users to be able to do with your data and for how long.	*DCC Access, Use and Reuse*
Pin down and communicate the significant properties of your data.	*DCC Access, Use and Reuse*
Ensure that you provide enough context to ensure that your data can be located and used – either by the originally designated user community or new users over time.	*DCC Access, Use and Reuse*

4.b.3.4.1 Visibility

Recommendation	*Found in*
N.A.	

4.b.3.4.2 Availability and accessibility

Recommendation	Found in
Ensure that data is accessible to both designated users and reusers, on a day-to-day basis. This may be in the form of publicly available published information. Robust access controls and authentication procedures may be applicable.	*DCC Access, Use and Reuse*

4.b.3.4.3 Tools and interfaces

Recommendation	Found in
Maintain a watch on appropriate community activities, and participate in the development of shared standards, tools and suitable software.	*DCC Community Watch and Participation*

4.b.3.4.4 Monitoring, review and feedback

Recommendation	*Found in*
N.A.	

4.b.4 Sources:

http://www.dcc.ac.uk/resources/curation-lifecycle-model

4.c Appendix C DPC

The Handbook was first compiled by Neil Beagrie and Maggie Jones[1] and is now maintained and updated by the Digital Preservation Coalition (DPC) in collaboration with the National Library of Australia and the Preserving Access to Digital Information (PADI) Gateway[2].

The handbook is intended to provide a bridge between broad, high level overviews and explicit, detailed guidelines applicable to the needs of a specific institution. The strategic overviews are intended to link to operational activities in order to reinforce the need to develop practical procedures grounded firmly in the business mission of the institution.

Some of the issues and segments are written from a UK perspective and legislation in this area will vary from country to country. However, many of the broader suggestions have wider implications.

4.c.1 Organizational Framework

4.c.1.1 Purpose and requirements

4.c.1.1.1 Scope and objectives

Recommendation	Found in
Define a coherent global strategy from the outset, ensuring that everyone concerned is involved.	*4.1.2*
A clear focus should be on the end purpose of the archiving process which is to serve the consumers or "designated communities" of current and future users.	*3.7*
Key stakeholders and decision-makers should be motivated to contribute to the medium to long term preservation of digital materials. These key stakeholders include the producer, the rights holder, the repository and the consumer, who each may or may not be the same entity depending on the organisation.	*3.7*

4.c.1.1.2 Collection policy

Recommendation	Found in
A clear set of guidance documents such as the organization's mission and collection (selection) policies and guidelines will reduce long term cost defining the aim and direction of collections and services for more efficient decision making.	*3.7*

[1] Both were also involved with the RLG-OCLC report.
[2] The site states that "…during this interim period of establishing the new DPC website we are aware that the Handbook requires some further work. http://www.dpconline.org/advice/preservationhandbook

4.c.1.1.3 Criteria for evaluating data

Recommendation	*Found in*
N.A.	

4.c.1.2 Legal and regulatory framework

Recommendation	Found in
Use a legal advisor to guide your rights management policy and develop documents.	*3.4*
Develop model letters for rights clearance, model deposit agreements, model licences and clauses for preservation activities	*3.4*
If you are licensing material from third parties ensure they have addressed future access to subscribed material in the licence and have robust procedures to support this	*3.4*
Prepare reasoned arguments and explanations for your preservation activities; rights holders will need to be convinced of the need and persuaded that their interests will be safeguarded.	*3.4*
Keep detailed records of rights negotiations	*3.4*
Treat licences and rights correspondence as key institutional records to be retained in fireproof and secure environments	*3.4*
Make a schedule clearly identifying a list of materials deposited and covered by the licence	*3.4*

4.c.1.3 Funding and resource planning

Recommendation	Found in
Offering bit-level preservation, where the only undertaking is to guarantee storage and delivery of the sequence of bits, will have lower costs than a repository managing full migration paths or emulation solutions.	*3.7*
A clear set of guidance documents such as the organization's mission and collection (selection) policies and guidelines will reduce long term cost defining the aim and direction of collections and services for more efficient decision making (also mentioned under 1.1.2 Collection policy)	*3.7*
Consider engagement in greater collaboration within and between organizations in order effectively to confront and overcome the challenges of digital preservation. Collaboration forms that should be considered include a) Internal collaborations; b) External collaborations; and c) Informal arrangements.	*3.1*
Preservation strategies enacted early in the life cycle are likely to be more cost effective than salvage attempts left until technology has already moved on significantly	*3.7*
Develop an Information Strategy which integrates IT training with the overall mission of the institution	*3.5*

Identify, in consultation with key staff, a skills audit, to determine what specific competencies are required to meet organizational objectives	*3.5*
Establishing a balance between recruiting specific skills and effectively developing existing talent	*3.5*
Ensure staff have access to appropriate equipment	*3.5*
Ensure access to practical "hands on" training and practice	*3.5*
Encourage networking between colleagues in other institutions	*3.5*
Consider strategies such as short-term secondment to an institution which may have more experience in a specific area	*3.5*
Involving staff in designing training and development programmes	*3.5*
Facilitate effective multi-disciplinary communication	*3.5*
Take a broad view of what constitutes training and development (i.e. combination of formal courses, both generic and tailor-made, informal training within the organisation, skills transfer within the organisation, networking etc.).	*3.5*
IPR: A clause should be drafted to cover the following: • *Permissions needed for content.* • *Permissions needed for associated software.* • *Permissions needed for copying for the purposes of preservation.* • *Permissions needed for future migration of content to new formats for the purposes of preservation.* • *Permissions needed for emulation for the purposes of preservation.* • *Permissions in respect of copyright protection mechanisms.*	*3.5*
Statutory and Contractual Issues: A clause should be drafted to cover the following: • *Statutory permissions and legal deposit obligations in respect of electronic materials.* • *Grant and contractual obligations in respect of electronic materials.* • *Conditions, rights and appropriate interests of authors, publishers and other funders.* • *Confidential information and protection of the confidentiality of individuals and institutions.* • *Protecting the integrity and reputation of data creators or other stakeholders.*	*3.5*
Investment by the preservation agency: A clause should be drafted to cover the following: • *IPR in any value added by the preservation agency.* • *Withdrawal clauses (and associated fees).*	*3.5*

4.c.1.4 Long-Term Preservation Policy

Recommendation	Found in
being explicit about preservation commitments is a crucial element of preservation policy and procedures.	*4.2.1*
Policies to identify which digital resources should be stored online	*4.3.1*
Retention policies to determine at what stage (if ever) online storage of digital resources will be re-assessed	*4.3.1*

4.c.1.5 Access Policy

Recommendation	Found in
A clause should be drafted to cover permissions and conditions in respect of access to the material	*3.5*

4.c.2 Technological Environment

4.c.2.1 IT Architecture

Recommendation	*Found in*
N.A.	

4.c.2.2 Standards and Formats

Recommendation	Found in
Selection of cataloguing and documentation standard	*4.3.2*
Knowledge of all relevant standards for all categories of digital resources acquired by the institution.	*4.3.2*
Written guidelines on preferred and acceptable standards	*4.3.2*
Institutional strategies for outreach, collaboration, standards and best practice.	*4.3.2*
Technology watch on standards activities	*4.3.2*
Use "open" non-proprietary, well-documented file formats wherever possible.	*5.5*
Alternatively utilise file formats which are well-developed, have been widely adopted and are de facto standards in the marketplace.	*5.5*
Identify formats acceptable for the purposes of transfer, storage and distribution to users (these may be distinct).	*5.5*
Minimise the number of file formats to be managed as far as is feasible/desirable	*5.5*
Do not use encryption or compression for archival files if possible	*5.5*

4.c.2.3 Security and Risk Management / Media Monitoring and Refreshing Strategy

Recommendation	Found in
System and physical security policies and procedures should be in place to	*4.2.3*

ensure the care and integrity of items during accessioning. These should be developed from and reflect the institutional policies and procedures on security.	
Establish disaster recovery plan	*4.3.1*
Control access to storage facilities and processing areas. Store in separate, preferably lockable area	*4.3.1*
Ensure no unauthorised access	*4.3.1*
Design audit features into mass storage systems and computerised physical access controls. Undertake regular random checks if automated audits are not feasible	*4.3.1*
Establish procedures to ensure no deliberate or inadvertent changes can take place	*4.3.1*
Ensure all legal requirements are met	*4.3.1*
Establish procedures for ensuring authenticity	*4.3.1*
Use passwords and user ids, and other network security procedures	*4.3.1*
Define system and area access privileges for staff	*4.3.1*
Assign specific staff responsibilities for data security and storage facilities	*4.3.1*
Undertake a retrospective survey of digital holdings, a risk assessment and action plan	*5.5*
Implement a process of technology watch and/or implement procedures for standardisation and changes in technology in your IS strategy.	*5.5*
Maintain a list of hardware/software available within the institution and use this to flag implications for technology change and hardware/ software replacement/ retention.	*5.5*
Ensure you have good preservation metadata in a computerised catalogue which can form the basis for technology watch and monitoring.	*5.5*
Consider "digital archaeology" to retrieve access to data in obsolete formats.	*5.5*

4.c.3 Data Curation

4.c.3.1 Pre-Ingest Function

4.c.3.1.1 Information and guidance given to data producer

Recommendation	Found in
Provide documentation to guide and support transfer of digital resources from suppliers.	*4.2.4*

4.c.3.2 Ingest Function

4.c.3.2.1 Information and documentation from data producer

Recommendation	Found in
A minimum standard of information required for cataloguing	*4.2.4*

4.c.3.2.2 Quality assurance and data checking

Recommendation	Found in
Unique numbering of each item accessioned.	*4.2.4*
Marking and labelling procedures	*4.2.4*
Handling guidelines for different media	*4.2.4*
Validation procedures to check media, content, and structure	*4.2.4*

4.c.3.2.3 Data documentation and enhancement

Recommendation	Found in
Guidelines for retrospective documentation or catalogue enhancement	*4.2.4*

4.c.3.3 Archival Storage and Preservation

4.c.3.3.1 Physical data preservation and storage

Recommendation	Found in
Generating multiple copies of an item as part of an institution's storage and preservation policy	*4.2.4*
Policies for maintaining documents on central file server	*4.3.1*
Strategies for migrating to larger file server before full capacity is reached	*4.3.1*

4.c.3.3.2 Preservation strategy

Recommendation	Found in
Procedures to update collection management databases	*4.2.4*
Targets for accessioning tasks and timescales for their completion	*4.2.4*
Written policies and guidelines, including selection policy for materials to be migrated	*4.3.2*
Quality control procedures	*4.3.2*
Rigorous documentation of migration procedure	*4.3.2*
Preservation metadata and documentation	*4.3.2*
Migrate data whenever there is a software upgrade or a new software application is installed	*4.3.2*
Ensure the migration results in little or no loss in content or context	*4.3.2*
Employ strict quality control procedures that may include testing the migration programme with a sample of records or bit/byte or checksum comparisons of migrated and original data.	*4.3.2*
Retain copies of the digital resource in its original format whenever some information or presentation of the resource may be lost or modified in migration	*4.3.2*

4.c.3.3.3 Version control/change procedures

Recommendation	Found in
Procedures for updating, and managing versions or editions of an item.	*4.2.4*

4.c.3.4 Dissemination

4.c.3.4.1 Visibilty

Recommendation	*Found in*
N.A.	

4.c.3.4.2 Availability and accessibility

Recommendation	*Found in*
N.A.	

4.c.3.4.3 Tools and interfaces

Recommendation	*Found in*
N.A.	

4.c.3.4.4 Monitoring, review and feedback

Recommendation	*Found in*
N.A.	

4.d Appendix D DPE Platter

Planning Tool for Trusted Electronic Repositories (PLATTER)

The PLATTER is a framework, which provides a basis for a digital repository to plan the development of its goals, objectives and performance targets over the course of its lifetime in a manner which will contribute to the repository establishing trusted status amongst its stakeholders. PLATTER cannot be perceived as an audit or certification tool itself, but rather a framework that will allow new repositories to incorporate the goal of achieving trust into their planning from an early stage. A repository planned using PLATTER will find itself in a strong position when it subsequently comes to apply one of the existing auditing tools to confirm the adequacy of its procedures for maintaining the long term usability of and access to its material.

The PLATTER acknowledges the diversity of the organisations, which may be included under the term "digital repository" by requiring repositories to answer a questionnaire which characterises the repository relative to other repositories and which can be used to determine how and whether the goals and objectives we have identified are to be realised in a given organisation.

The PLATTER process is centred around a group of Strategic Objective Plans (SOPs) through which a repository specifies its current objectives, targets, or key performance indicators in those areas which have been identified as central to the process of establishing trust.

4.d.1 Organizational Framework

4.d.1.1 Purpose and Requirements

Recommendation	Found in
Determine the general functional type of the repository: *What is the source of the repository's mandate? Is the Repository for profit or non-profit? Does the Repository receive a significant proportion of its material from a legally mandated source? What is the operational status of the repository (not yet running, running but still under development, mature)*	*Recommendation 3.1.*

4.d.1.1.1 Scope and objectives

Recommendation	Found in
Define the overall scale of the repository, whether expressed in human, technical, or financial terms: *What is the amount of digital material you expect to archive? How many distinct digital objects do you expect to archive? How many fulltime-equivalent staff does the Repository expect to employ? How many distinct end-users are expected to access material in the Repository over the course of a calendar year?*	*Recommendation 3.2.*

Is the majority of the material in the Repository simple, moderately complex, or highly complex?	*Recommendation 3.3 Question Q3.2*
How specialized is the data in the Repository (low, medium or high)?	*Recommendation 3.3 Question Q3.3*
How sensitive is the most sensitive material in the Repository (low, medium, high)?	*Recommendation 3.3 Question Q3.4*
In which of the three access classes (open, restricted, closed) does the Repository have significant holdings?	*Recommendation 3.3 Question Q3.5*

4.d.1.1.2 Collection policy

Recommendation	Found in
Which of the three acquisition strategies (push, pull, self-creation) account for a significant portion of the total material in the Repository?	*Recommendation 3.3 Question Q3.1*
Acquire relevant material: specify quantitative targets for the material to be acquired by the repository	*Goal 2.1*
Negotiate deposit agreements: Scope of material, Delivery form, File formats, Accompanying metadata, etc.	*Goal 2.2*
Obtain Physical Control of Materials	*Goal 2.3*
It is essential that a repository have in place a monitoring system to determine that the required material is actually made available by the producer or depositors.	*Goal 2.4*
The repository should institute procedures to monitor the relevance of any deposit agreement, taking into account the same issues considered in the initial development of the agreement	*Goal 2.5*

4.d.1.1.3 Criteria for evaluating data

Recommendation	*Found in*
N.A.	

4.d.1.2 Legal and Regulatory Framework

Recommendation	*Found in*
N.A.	

4.d.1.3 Funding and Resource Planning

Recommendation	Found in
In order for the business plan to remain current, regular review is needed. As sources of income cannot normally be guaranteed for more than a few years, monitoring is needed to identify such shortages before they develop into a budget deficiency.	*Goal 1.1*
The repository must achieve an income sufficient for routine functioning.	*Goal 1.2*
Ensure contingency plans for financial cutbacks or emergencies: have an agreement with another repository about housing vital data; prioritise services to be retained in the case of financial problems	*Goal 1.3*
Define and maintain marketing and outreach plans suitable for the	*Goal 1.4*

repository's needs	

4.d.1.4 Long-Term Preservation Policy

Recommendation	Found in
The preservation tasks are ensured even beyond the existence of the digital repository by an agreement with other repositories to take over the collections. The very purpose of the succession plan(s) is to detail agreement(s) about who will inherit the digital data if the repository ceases to function.	*Goal 7.1*

4.d.1.5 Access Policy

Recommendation	Found in
Create, Maintain and Review a Mission Statement which reflects the Repository's Mandate	*Goal 4.1*
The definition of the designated community should be available on the repository website and the community should be very well aware of available delivery and access options. It is essential to monitor and reflect all changes inside the community over the time.	*Goal 4.2*

4.d.2 Technological Environment

4.d.2.1 IT Architecture

Recommendation	Found in
IT Infrastructure must be capable of coping with the scale of data storage, processing and transport appropriate for the repository.	*Goal 5.1*
IT infrastructure must be able to guarantee the integrity and security of the stored data.	*Goal 5.2*
The IT infrastructure must be able to guarantee that certain services remain available to the users. Identify which services of the repository are no longer required and can be closed and which are in great demand.	*Goal 5.3*

4.d.2.2 Standards and Formats

Recommendation	Found in
What are the main sources of bibliographic and descriptive metadata in the repository?	*Recommendation 3.4 Question Q4.1*
What interoperability standards are implemented in the Repository?	*Recommendation 3.4 Question Q4.2*
A data provider encodes his data in a package format (Submission Information Package) acceptable to the repository. The repository receives these SIPs, and repackages them for storage (Archival Information Package). A repository user requests the data, and the repository repackages it in a format appropriate for the user (Dissemination Information Package).	*Recommendation 5.6.*
For Submission Information Package: Specify the digital object formats and accompanying meta-data the repository will accept. Formats (potentially) containing encryption, or embedded objects and files, as well as commercially protected formats can sometimes prove impossible to convert.	*Goal 6.1*
For Archival Information Package: Specify the data format and metadata	*Goal 6.2*

content for archiving digital objects. Where the SIPs defined in the previous goal should strike some balance between commonly used formats, and formats useful for archiving, the AIPs do not need to.	
For Dissemination Information Package: Specify the data formats used for disseminating digital objects.	*Goal 6.3*
Specify the transformation from SIP to AIP and from AIP to DIP.	*Goal 6.4 and 6.5*

4.d.2.3 Security and Risk Management / Media Monitoring and Refreshing Strategy

Recommendation	Found in
The repositories must have strategies to deal with budget shortfalls. Common strategies include willing creditors to boost the repository for a period, or having savings. The repository should also have a prioritized list of the services it provides, beyond what is absolutely required of it, to ease in selecting which could be terminated, if income shortages arrive.	*Goal 8.1*
Political upheaval can come to seemingly stable societies. Repositories caught in such circumstances could face opposing demands to both open up their collections, and to close them down.	*Goal 8.1*
Many repositories are a part of a larger organization, which due to some changes could have problems seeing the purpose of the repository. The repository must actively seek to demonstrate its value to those with executive power over it.	*Goal 8.1*
The repository should, so far as possible, detail procedures for dealing with environmental disasters in a timely manner.	*Goal 8.1*
Loss of educated key staff: • *the internal workings and trade secrets of the repository could be revealed to outsiders – the repository should prepare press statements and other communications for when the secrets break* • *the services of the repository could cease to function - the best course of action would be to shut down unmaintainable services until new staff have been hired or trained.*	*Goal 8.1*
The physical security of the repository can also be threatened. The repository should identify likely threats, and address them, in addition to having plans in the event of these threats materializing.	*Goal 8.1*

4.d.3 Data Curation

4.d.3.1 Pre-Ingest Function

Recommendation	*Found in*
N.A.	

4.d.3.1.1 Information and guidance given to data producer

Recommendation	Found in
The repository should have some description for content providers about how to package data and metadata or representation information. The repository should have a policy regarding the completeness and correctness of new data and what action to take regarding invalid or incomplete data.	*Goal 6.1*
Specify sources and formats for bibliographic and descriptive metadata in	*Goal 6.1.1*

the received data. This is the basic information required to enable the object to be discovered (e.g. by a search engine) and interpreted.	
Specify technical metadata in the SIP	*Goal 6.1.2*

4.d.3.2 Ingest Function

Recommendation	*Found in*
N.A.	

4.d.3.2.1 Information and documentation from data producer

Recommendation	*Found in*
N.A.	

4.d.3.2.2 Quality assurance and data checking

Recommendation	*Found in*
N.A.	

4.d.3.2.3 Data documentation and enhancement

Recommendation	Found in
The repository should have a method for verifying the completeness and correctness for newly generated AIPs, as with SIPs.	*Goal 6.2*
The metadata from the SIP should be conserved, and elaborated, rather than being changed. A file format, that strikes a balance between being human readable and being machine parseable, such as XML could be used to encode the metadata.	*Goal 6.2.1*

4.d.3.3 Archival Storage and Preservation

Recommendation	*Found in*
N.A.	

4.d.3.3.1 Physical data preservation and storage

Recommendation	Found in
What strategy is used for storage? (in-house, external, in-house under external support)?	*Recommendation 3.4 Question Q4.3*

4.d.3.3.2 Preservation strategy

Recommendation	Found in
What strategy is used for software management? • *support by the software supplier* • *support by a third party* • *self-support (i.e. in-house)* • *support by a user and developer community*	*Recommendation 3.4 Question Q4.4*
Repository must maintain understanding of all structural (e.g. file encoding) standards and formats	*Goal 9.2*

Repository must maintain understanding of contemporary and emerging hardware, software and storage technologies	*Goal 9.1*
Repository must maintain understanding of identified user communities and their associated competences and knowledge base	*Goal 9.3*
Repository must maintain understanding of preservation requirements for each stored information asset or class of information	*Goal 9.4*
Repository must maintain, exercise and evaluate preservation strategies capable of meeting specific preservation targets	*Goal 9.5*
Repository must maintain and exercise appropriate appraisal policies to determine which information must continue to be preserved	*Goal 9.6*

4.d.3.3.3 Version control/change procedures

Recommendation	*Found in*
N.A.	

4.d.3.4 Dissemination

Recommendation	*Found in*
N.A.	

4.d.3.4.1 Visibility

Recommendation	*Found in*
N.A.	

4.d.3.4.2 Availability and accessibility

Recommendation	Found in
Create and Implement A Repository Access Policy taking into account: copyright law, law on dissemination of personal or commercially-sensitive data, National security, Libel, obscenity, hate-speech and blasphemy laws, Specific contractual restrictions imposed in deposit agreements	*Goal 4.3*
Specify and fulfil technical requirements for dissemination and access	*Goal 4.4*

4.d.3.4.3 Tools and interfaces

Recommendation	Found in
N.A.	

4.d.3.4.4 Monitoring, review and feedback

Recommendation	Found in
The repository should periodically evaluate usage statistics and watch for services going out of favour with the users. Giving the users a way to request new features could also help the repository staying current.	*Goal 8.1*
Technological changes can greatly affect the business model of a repository, and at times come about very quickly. Repositories need to watch the technological trends, evaluate which could be potentially harmful, and adapt.	*Goal 8.1*

4.e Appendix E Nestor

Catalogue of Criteria for Trusted Digital Repositories

The work of the Nestor working group "Trusted repositories" has the main focus on the challenge of earning trust and communicating this, be it to fulfill a legal mandate or simply to survive in the market. It identifies criteria which permit the trustworthiness of a digital repository to be evaluated, both at the organizational and technical levels.

4.e.1 Organizational Framework

4.e.1.1 Purpose and Requirements

Recommendation	Found in
N.A.	

4.e.1.1.1 Scope and objectives

Recommendation	Found in
The digital repository has defined its goals	*NESTOR Organisational Framework 1*
The digital repository has developed criteria for the selection of its digital objects	*NESTOR Organisational Framework 1.1*
The digital repository has defined its designated community/communities	*NESTOR Organisational Framework 1.3*

4.e.1.1.2 Collection policy

Recommendation	Found in
The digital repository ensures the integrity of the digital objects during all processing stages. These stages are Ingest, Archival storage, Access	*NESTOR Object management 6*

4.e.1.1.3 Criteria for evaluating data

Recommendation	*Found in*
The digital repository ensures the authenticity of the digital objects during all stages of processing. These stages are Ingest, Archival storage, Access	NESTOR Object management 7

4.e.1.2 Legal and Regulatory Framework

Recommendation	Found in
Legal contracts exist between producers and the digital repository.	*NESTOR Organisational Framework 3.1*
In carrying out its archiving tasks, the digital repository acts on the basis of legal arrangements.	*NESTOR Organisational Framework 3.2*
With regards to use, the digital repository acts on the basis of legal	*NESTOR Organisational*

arrangements.	*Framework 3.3*

4.e.1.3 Funding and Resource Planning

Recommendation	Found in
Adequate financing of the digital repository is secured.	*NESTOR Organisational Framework 4.1*
Sufficient numbers of appropriately qualified staff are available.	*NESTOR Organisational Framework 4.2*

4.e.1.4 Long-Term Preservation Policy

Recommendation	Found in
The digital repository assumes responsibility for long-term preservation of the information represented by the digital objects	*NESTOR Organisational Framework 1.2*
Appropriate organisational structures exist for the digital repository.	*NESTOR Organisational Framework 4.3*
The digital repository engages in long-term planning.	*NESTOR Organisational Framework 4.4*
The digital repository reacts to substantial changes	*NESTOR Organisational Framework 4.5*
Continuation of the preservation tasks is ensured even beyond the existence of the digital repository.	*NESTOR Organisational Framework 4.6*
The digital repository implements strategies for the long-term preservation of the archival information packages (AIPs).	*NESTOR Object management 10.4*

4.e.1.5 Access Policy

Recommendation	Found in
The digital repository ensures its designated community/ communities can access the digital objects.	*NESTOR Organisational Framework 2.1*
The digital repository ensures that the designated community/communities can interpret the digital objects.	*NESTOR Organisational Framework 2.2*
The digital repository permits usage of the digital objects based on defined criteria.	*NESTOR Object management 11*
The digital repository acquires adequate metadata to record the corresponding usage rights and conditions.	*NESTOR Object management 12.6*

4.e.2 Technological Environment

4.e.2.1 IT Architecture

Recommendation	Found in
Archival storage of the digital objects is undertaken to defined specifications.	*NESTOR Object management 10*
The data management system is capable of providing the necessary digital	*NESTOR Object*

repository functions.	*management 12*
The digital repository uniquely and persistently identifies its objects and their relationships.	*NESTOR Object management 12.1*
The IT infrastructure is appropriate.	*NESTOR Infrastructure and Security 13*
The IT infrastructure implements the object management requirements.	*NESTOR Infrastructure and Security 13.1*
The IT infrastructure implements the security requirements of the IT security system.	*NESTOR Infrastructure and Security 13.2*
The infrastructure protects the digital repository and its digital objects.	*NESTOR Infrastructure and Security 14*

4.e.2.2 Standards and Formats

Recommendation	Found in
The digital repository records adequate metadata for formal and contentbased description and identification of the digital objects.	*NESTOR Object management 12.2*
The digital repository records adequate metadata for structural description of the digital objects.	*NESTOR Object management 12.3*
The digital repository records adequate metadata to document all the changes made by the digital repository to the digital objects.	*NESTOR Object management 12.4*
The digital repository acquires adequate metadata for technical description of the digital objects.	*NESTOR Object management 12.5*
The package structure is preserved at all times.	*NESTOR Object management 12.7*

4.e.2.3 Security and Risk Management / Media Monitoring and Refreshing Strategy

Recommendation	*Found in*
N.A.	

4.e.3 Data Curation

4.e.3.1 Pre-Ingest Function

Recommendation	*Found in*

4.e.3.1.1 Information and guidance given to data producer

Recommendation	Found in
The digital repository accepts digital objects from the producers based on	*NESTOR Object*

defined criteria.	*management 9*
The digital repository specifies its submission information packages (SIPs).	*NESTOR Object management 9.1*
The digital repository identifies which characteristics of the digital objects are significant for information preservation.	*NESTOR Object management 9.2*

4.e.3.2 Ingest Function

Recommendation	Found in
N.A.	

4.e.3.2.1 Information and documentation from data producer

Recommendation	Found in
N.A.	

4.e.3.2.2 Quality assurance and data checking

Recommendation	Found in
All processes and responsibilities have been defined.	NESTOR Organisational Framework 5.1
The digital repository documents all its elements based on a defined process.	NESTOR Organisational Framework 5.2

4.e.3.2.3 Data documentation and enhancement

Recommendation	Found in
N.A.	

4.e.3.3 Archival Storage and Preservation

Recommendation	Found in
The digital repository defines its archival information packages (AIPs).	NESTOR Object management 10.1
The digital repository takes care of transforming the submission information packages (SIPs) into archival information packages (AIPs).	NESTOR Object management 10.2
The digital repository guarantees the storage and readability of the archival information packages (AIPs).	NESTOR Object management 10.3

4.e.3.3.1 Physical data preservation and storage

Recommendation	Found in
The digital repository has technical control of the digital objects in order to carry out long-term preservation measures.	NESTOR Object management 9.3

4.e.3.3.2 Preservation strategy

Recommendation	Found in
The digital repository has a strategic plan for its technical preservation measures (preservation planning).	NESTOR Object management 8

4.e.3.3.3 Version control/change procedures

Recommendation	Found in
N.A.	

4.e.3.4 Dissemination

Recommendation	Found in
The digital repository defines its dissemination information packages (DIPs).	*NESTOR Object management 11.1*
The digital repository ensures transformation of archival information packages (AIPs) into dissemination information packages (DIPs).	*NESTOR Object management 11.2*

4.e.3.4.1 Visibility

Recommendation	*Found in*
N.A.	

4.e.3.4.2 Availability and accessibility

Recommendation	*Found in*
N.A.	

4.e.3.4.3 Tools and interfaces

Recommendation	*Found in*
N.A.	

4.e.3.4.4 Monitoring, review and feedback

Recommendation	*Found in*
N.A.	

4.f Appendix F OAIS

Reference Model for an Open Archival Information System (OAIS)

OAIS is a technical recommendation of practice for providing long term preservation of digital information. An organization can establish an OAIS archive by adhering to the recommendations and standards stated by OAIS. The term 'open' in OAIS indicates, that these recommendations and standards are not final and are subject to discussion in open forums, it does not imply that the access to the particular archive is unrestricted.

http://public.ccsds.org/publications/archive/650x0m2.pdf

4.f.1 Organizational Framework

4.f.1.1 Purpose and Requirements

Recommendation	*Found in*
N.A.	

4.f.1.1.1 Scope and objectives

Recommendation	Found in
The repository shall have documented policies and procedures for preserving its AIPs, they should never be deleted unless allowed as part of an approved policy, there should be no ad-hoc deletions.	*3.2.5*
The producers and consumers communities should be provided with submission and dissemination standards, policies and procedures to support the preservation objectives of the OAIS.	*3.2.5*

4.f.1.1.2 Collection policy

Recommendation	Found in
The repository shall have a policy that the information to be preserved is understandable to the communities without the need of an expert.	*3.1*

4.f.1.1.3 Criteria for evaluating data

Recommendation	Found in
Archives may express an evaluation of authenticity of its holdings, based on community practise and recommendations.	*4.1.1.2*

4.f.1.2 Legal and Regulatory Framework

Recommendation	Found in
Archives will honour all applicable restrictions of information. They shall understand the intellectual property rights concepts, such as copyrights and any other applicable laws prior to accepting copyrighted materials into the OAIS.	*3.2.2*

4.f.1.3 Funding and Resource Planning

Recommendation	Found in
The archive should have a formal succession plan, contingency plans, and/or escrow arrangements in place, in case the archive ceases to operate or the governing or funding institution substantially changes its scope.	*3.2.5*
Management is often the primary source of funding for an OAIS and may provide for resource utilization (personnel, equipment, facilities).	*2.3.1*

4.f.1.4 Long-Term Preservation Policy

Recommendation	Found in
The repository shall negotiate and accept appropriate information from information producers, obtain sufficient control of the information, determine which communities are the Designated Communities, define the information provided accordingly, hence define a knowledge base, ensure that the information to be preserved is independently understandable to the designated community without the need of experts, follow documented policies and procedures to prevent data loss by deletion or any other means, disseminate information as copies to the designated communities.	*3.1*

4.f.1.5 Access Policy

Recommendation	Found in
Supporting tools, such as access and use mechanisms, i.e. Content Data Object (CDO) are determined to remain adequate for an extended period of time. Once this software will cease to function it shall be re-implemented, or alternatively emulated by new software.	*5.2*

4.f.2 Technological Environment

4.f.2.1 IT Architecture

Recommendation	Found in
OAIS associations can be categorized technically by both external and internal factors. External factors include characteristics of the Producer and Consumer communities. Internal factors could include common implementations of the information models presented in 4.2, or multi-Archive sharing of one or more of the functional areas presented in 4.1.	*6.1*

4.f.2.2 Standards and Formats

Recommendation	Found in
To preserve the meaning of an Information Object, its Representation Information must also be preserved. This is most easily accomplished when the Representation Information objects are expressed in forms that are easily understandable, such as text descriptions that use widely supported standards such as ASCII characters for electronic versions. One problem with the use of only text descriptions is that such descriptions can be ambiguous. This is addressed by the use of standardized, formal description languages containing well-defined constructs with which to describe data structures. These languages may need to be augmented with text descriptions to convey fully the semantics of the Representation Information.	*4.2.1.3.2*

4.f.2.3 Security and Risk Management / Media Monitoring and Refreshing Strategy

Recommendation	Found in
Media Decay: Digital media, over time, become increasingly unreliable as secure preservers of bits. Even those that are used with some level of error correction eventually need to be replaced. The net result of media decay is that AIP information must be moved to newer media.	*5.1.1*
Management will generally conduct some regular review process to evaluate the OAIS performance and progress toward Long Term goals, and assess the risks to which the OAIS and its holdings are exposed.	*2.3.1*
The Establish Standards and Policies function is responsible for establishing and maintaining the Archive system standards and policies. It receives budget information and policies such as the OAIS charter, scope, resource utilization guidelines, and pricing policies It provides Management with periodic reports. It receives from Management. Recommendations for Archive system enhancement, proposals for new Archive data standards, and periodic risk analysis reports from Preservation Planning. It will have to face risks from unforeseen events (unplanned down time due to network outage, software bugs, hardware failure, human error, disk crash, etc.) and make the appropriate decisions to minimize the risk of not fulfilling the Archive's commitments. It also receives performance information and Archive holding inventories from Manage System Configuration.	*4.1.1.5*

4.f.3 Data Curation

4.f.3.1 Pre-Ingest Function

Recommendation	Found in
The first contact between the OAIS and the Producer is a request that the OAIS preserve the data products created by the Producer. This contact may be initiated by the OAIS, the Producer or Management. The Producer establishes a Submission Agreement with the OAIS, which identifies the SIPs to be submitted and may span any length of time for this submission. Some Submission Agreements will reflect a mandatory requirement to provide information to the OAIS, while others will reflect a voluntary offering of information and others may reflect any payments which may be	*2.3.2*

involved. Even in the case where no formal Submission Agreement exists such as in the archiving of many World Wide Web sites, a virtual Submission Agreement may exist specifying the file formats and the general subject matter the OAIS will accept.	

4.f.3.1.1 Information and guidance given to data producer

Recommendation	Found in
The Receive Submission function provides the appropriate storage capability or devices to receive a SIP from the Producer (or from Administration). Digital SIPs may be delivered via electronic transfer (e.g., FTP), loaded from media submitted to the Archive, or simply mounted (e.g., CD-ROM) on the Archive file system for access. Non-digital SIPs would likely be delivered by conventional shipping procedures. The Receive Submission function may represent a legal transfer of custody for the Content Information in the SIP, and may require that special access controls be placed on the contents. This function provides a confirmation of receipt of a SIP to the Producer, which may include a request to resubmit a SIP in the case of errors resulting from the SIP submission.	*4.1.1.2*

4.f.3.2 Ingest Function

Recommendation	Found in
The Ingest Functional Entity (labeled 'Ingest' in the figures in this section) provides the services and functions to accept Submission Information Packages (SIPs) from Producers (or from internal elements under Administration control) and prepare the contents for storage and management within the Archive. Ingest functions include receiving SIPs, performing quality assurance on SIPs, generating an Archival Information Package (AIP) which complies with the Archive's data formatting and documentation standards, extracting Descriptive Information from the AIPs for inclusion in the Archive database, and coordinating updates to Archival Storage and Data Management.	*4.1*

4.f.3.2.1 Information and documentation from data producer

Recommendation	Found in
The data within the data Producer entity are private and may be in any format the Producer desires. However, when the decision is made to store the data in an OAIS, the Producer who is responsible for the data meets	*4.3.1*

with archivists to negotiate a Submission Agreement as [..] This agreement defines information such as the content, format, and scheduled arrival times of the Submission Information Package (SIP). The SIP is an Information Package that is provided to the OAIS by the Producer. The SIP consists of the Content Information plus the data that is necessary to assure that those data can be maintained by the OAIS and that the data can be interpreted and used by Consumers who withdraw them from the OAIS in the future.	

4.f.3.2.2 Quality assurance and data checking

Recommendation	Found in
The Quality Assurance function validates (QA results) the successful transfer of the SIP to the temporary storage area. For digital submissions, these mechanisms might include Cyclic *Redundancy Checks (CRCs) or checksums associated with each data file, or the use of system log files to record and identify any file transfer or media read/write errors.*	*4.1.1.2*

4.f.3.2.3 Data documentation and enhancement

Recommendation	Found in
The Generate AIP function transforms one or more SIPs into one or more AIPs that conform to the Archive's data formatting standards and documentation standards. This may involve file format conversions, gathering adequate Representation Information, data representation conversions or reorganization of the Content Information in the SIPs. The Generate AIP function may issue report requests to Data Management to obtain reports of information needed by the Generate AIP function to produce the Descriptive Information that completes the AIP. This function sends SIPs or AIPs for audit to the Audit Submission function in Administration, and receives back an audit report. As a result of the audit report for example, it may be necessary to gather further Representation Information to ensure that the Content Information is understandable and usable by the Designated Community.	*4.1.1.2*

4.f.3.3 Archival Storage and Preservation

Recommendation	Found in

It is expected that the Ingest Functional Entity will coordinate the updates between Data Management and Archival Storage and provide appropriate coordination and error recovery. The AIP should first be stored in Archival Storage. The confirmation of that operation will include a unique identification to retrieve that AIP from Storage. This identifier should be merged into the Package Description prior to the addition of the Collection Description to *Data Management.*	*4.3.2*

4.f.3.3.1 Physical data preservation and storage

Recommendation	Found in
The Receive Data function receives a storage request and an AIP from Ingest and moves the AIP to permanent storage within the Archive. The transfer request may need to indicate the *anticipated frequency of utilization of the Data Objects making up the AIP in order to allow the appropriate storage devices or media to be selected for storing the AIP. This function will select the media type, prepare the devices or volumes, and perform the physical transfer to the Archival Storage volumes. Upon completion of the transfer, this function sends a storage confirmation message to Ingest, including the storage identification of the AIPs.*	*4.1.1.3*

4.f.3.3.2 Preservation strategy

Recommendation	Found in
Fundamentally, approaches to information preservation in the face of changing technologies and Designated Community requirements require digital migration [...] Digital migration of an AIP can include: – *Copying Content Data Object or Representation Information bits to new media* – *Altering or adding to Content Data Object or Representation Information bits* – *Altering or adding to PDI bits* – *Altering or adding to operational software whose role is essential to Content Information preservation (i.e., it is part of Representation Information)* – *Altering or adding to the bits that make up the AIP's Packaging Information*	*5*

4.f.3.3.3 Version control/change procedures

Recommendation	Found in
Digital Migrations that require some changes to the Content Information or PDI are referred to as Transformations. These changes will be to some of the bits in the Content Information *or PDI with corresponding changes in the associated Representation Information. In all cases the intent is to provide maximum information preservation. The resulting AIP is intended to be a full replacement for the AIP that is undergoing Transformation. The new AIP qualifies as a new AIP Version of the previous AIP. The first version of the AIP is referred to as the original AIP and may be retained for verification of information preservation.*	*5.1.3.4*

4.f.3.4 Dissemination

Recommendation	Found in
The OAIS may choose to provide this API as an implementation *alternative to the production and delivery of a physical DIP for dissemination. This type of service allows the Consumer, as a client, to develop applications that appear to directly access the AIPs. This sort of access could be very useful for applications such as data mining where the creation and shipping of DIPs containing large AICs is impractical. This API could allow an application to virtually navigate through an AIC, deliver the bits of the Content Data Object of selected AIUs to the application and identify locations for obtaining associated Representation Information and PDI. However, as technology evolves, the OAIS moves to new hardware, new media, and new operating systems. If the OAIS wishes to maintain the same API for its Consumers, it will need to provide a 'wrapper' around part of its new infrastructure to match its services to the established API. The API will need to be adequately documented and tested to ensure it correctly delivers the AIU Content Information using this new Access Software. This approach should not result in any changes to software developed by the Consumer community. When the API is applicable across a wide range of AIUs in the OAIS or there are a significant number of Consumer applications based on the API, this wrapping approach is clearly feasible and may result in a favorable cost/benefit ratio to the OAIS and its Designated Community. The 'Layered Model of Information' presented in annex E of this document further describes some potentially standard APIs.*	*5.2.1*

4.f.3.4.1 Visibility

Recommendation	Found in
The Package Description is not required for the Long Term Preservation of the Content Information but is needed to provide visibility and access into the contents of an Archive. The contents of the Package Description are highly dependent on the structure of the Content Information and PDI it describes.	*4.2.2.3*

4.f.3.4.2 Availability and accessibility

Recommendation	Found in
The Coordinate Access Activities function provides one or more interfaces to the information holdings of the Archive. This interface will normally be via computer network or dial-up link to an on-line service, but might also be implemented in the form of a walk-in facility, printed catalog ordering service, or fax-back type service. Three categories of Consumer requests are distinguished: query requests, which are executed in Data Management and return immediate query responses for presentation to the user; report requests, which may require a number of queries and produce formatted reports for delivery to the Consumer; and orders, which may access either or both Data Management and Archival Storage to prepare a formal Dissemination Information Package (DIP) for on- or off-line delivery. The Generate DIP function accepts a dissemination request, retrieves the AIP from Archival Storage, and moves a copy of the data to a temporary storage area for further processing. This function also transmits a report request to Data Management to obtain Descriptive Information needed for the DIP. The Deliver Response function handles both on-line and off-line deliveries of responses (DIPs, query responses, reports and assistance) to Consumers. For on-line delivery, it accepts a response from Coordinate Access Activities and prepares it for on-line distribution in real time via communication links.	*4.1.1.7*

4.f.3.4.3 Tools and interfaces

Recommendation	Found in
The OAIS Consumer interface in Access provides one or more Content Information IDs, with associated name spaces, to assist in identifying a particular Content Information object of interest. One or more of these Content Information IDs will be included in the PDI Reference Information associated with that Content Information object. The Descriptive Information in Data Management will map each of these IDs to the same	*5.1.2*

AIP ID. The Access Function uses this information to obtain the AIP ID and gives it to Archival Storage to retrieve the associated AIP.	

4.f.3.4.4 Monitoring, review and feedback

Recommendation	Found in
The Preservation Planning Functional Entity […] provides the services and functions for monitoring the environment of *the OAIS, providing recommendations and preservation plans to ensure that the information stored in the OAIS remains accessible to, and understandable by, the Designated Community over the Long Term, even if the original computing environment becomes obsolete.* *Preservation Planning functions include evaluating the contents of the Archive and periodically recommending archival information updates, recommending the migration of current Archive holdings, developing recommendations for Archive standards and policies, providing periodic risk analysis reports, and monitoring changes in the technology environment and in the Designated Community's service requirements and Knowledge Base. Preservation Planning also designs Information Package templates and provides design assistance and review to specialize these templates into SIPs and AIPs for specific submissions. Preservation Planning also develops detailed Migration plans, software prototypes and test plans to enable implementation of Administration migration goals.*	*4.1*

4.g Appendix G Digiplanet

4.h Appendix H DSA

Data Seal of Approval

Data Seal of Approval is a set of 16 guidelines for data repositories or data archives. It regards three stakeholders of data: the data producer, the data consumer and the data archive. The main responsibility belongs to the data archive. The objectives are:

- Assurance for researchers that their research results will be stored in a reliable manner and can be reused
- Guarantee for sponsors that research results will remain available for reuse
- Reliable access to the repositories for researchers
- Efficient distribution and archiving of data
- A division in three groups is made: data producers, data repository and data consumer. The storage scheme is also similar. The data producer provides the data with an information about the data, consists of
- Descriptive metadata: Information required to find research data, add transparency and importance, such as data elements in the Dublin Core Element Set.
- Structural metadata: Describes how different components of the data relate to each other. If the data is coded, the description of how it is coded is attached.
- Administrative metadata: Ensures permanent access by storing description of intellectual property, conditions for use and access and the preservation metadata needed for durable archiving of the research data.

The guidelines are specified in the form of questions, to which applicants should response. All information an applicant provides during the assessment should be available online.

If an organization only provides access to data, it does not need to implement certain guidelines if it archives the data in a Trusted Digital Repository (TDR) that has the DSA logo or better.

4.h.1 Organizational Framework

4.h.1.1 Purpose and Requirements

Best practice	Found in
An explicit mission in digital archiving which is documented in a mission statement	*Guideline 4*
A document which outlines the way in which the mission statement is implemented	*Guideline 4*
A continuation plan for digital assets	*Guideline 4*

4.h.1.1.1 Scope and objectives

Best practice	Found in
An explicit mission in digital archiving which is documented in a mission statement	*Guideline 4*
A document which outlines the way in which the mission statement is implemented	*Guideline 4*
A continuation plan for digital assets	*Guideline 4*

4.h.1.1.2 Collection policy

Best practice	Found in
Provide information on Institutional Review Boards (IRBs) or other official approval for data collection	*Guideline 1*
Information on data collection method: e.g. a methodology report	*Guideline 1*
Solid reputation of data producer	*Guideline 1*

4.h.1.1.3 Criteria for evaluating data

Best practice	Found in
Quality control checks to ensure that the data producer adheres to the request for metadata	Guideline 3
Check the identities of depositors	Guideline 12

4.h.1.2. Legal and Regulatory Framework

Best practice	Found in
Repository needs to be a legal entity	*Guideline 5*
Model contracts with data producers and consumers	*Guideline 5*
Published uses of conditions	*Guideline 5*
Comply with national and international laws	*Guideline 5*
Defined measures if conditions are not complied with	*Guideline 5*
Special procedures to manage data with disclosure risk: reviewing disclosure risk, limited/secure access, conditions of distribution, data anonymization	*Guideline 5*
Staff training to manage data with disclosure risk	*Guideline 5*

4.h.1.3 Funding and Resource Planning

Best practice	Found in
Carry out promotional activities	*Guideline 4*

4.h.1.4 Long-Term Preservation Policy

Best practice	Found in
Provide a plan for long-term preservation for digital assets	*Guideline 7*
Take future obsolescence of file formats into account	*Guideline 7*
Take measures to ensure long-term data usability	*Guideline 7*

4.h.1.5 Access Policy

Best practice	Found in
Require the data consumer to comply with access regulations and generally accepted codes of conduct (e.g. End User Licences)	*Guideline 14*
Define measures which are taken if conditions are not complied with	*Guideline 14*
Special contracts to grant access to confidential data	*Guideline 14*

4.h.2 Technological Environment

4.h.2.1 IT Architecture

Best practice	Found in
Derive metadata elements from established metadata standards, registries or conventions	*Guideline 3*

4.h.2.2 Standards and Formats

Best practice	Found in
Provide a plan for infrastructural development	*Guideline 13*
Technical infrastructure should be based on internationally accepted archival standards like OAIS	*Guideline 13*
Provide data in formats used by the research community	*Guideline 10*
Publish a list of preferred formats	*Guideline 2*
Check that data are deposited in preferred formats	*Guideline 2*
Require detailed information about file formats from depositors	*Guideline 2*

4.h.2.3 Security and Risk Management / Media Monitoring and Refreshing Strategy

Best practice	Found in
Create a crisis management plan	*Guideline 9*

4.h.3 Data Curation

4.h.3.1 Pre-Ingest Function

4.h.3.1.1 Information and guidance given to data producer

Best practice	Found in
Provide information on the full package information that should be deposited	*Guideline 1*
Provide: • *User friendly ways to enter metadata* • *Deposit forms with resource discovery metadata* • *Tools to create metadata at file level*	*Guideline 3*

4.h.3.2 Ingest Function

4.h.3.2.1 Information and documentation from data producer

Best practice	Found in
Sufficient information to others to assess the scientific and scholarly quality of data	*Guideline 1*
References to publications	*Guideline 1*
Used methods and research techniques, including data collection	*Guideline 1*
Information on how research data obtained shape	*Guideline 1*

4.h.3.2.2 Quality assurance and data checking

Best practice	Found in
Quality controls to ensure data is deposited in preferred formats and that requested metadata are included	*Guideline 2*

4.h.3.2.3 Data documentation and enhancement

Best practice	Found in
Provide references to publications based on the research data	*Guideline 1*
Pay attention to relevance of metadata items to data consumers	*Guideline 3*
Use metadata elements which are derived from established metadata standards, registries or conventions	*Guideline 3*

4.h.3.3 Archival Storage and Preservation

4.h.3.3.1 Physical data preservation and storage

Best practice	Found in
Implement and monitor checksums	*Guideline 11*
Data repository implement the following: • *Preservation policy* • *Strategy for backup / multiple copies*	*Guideline 6*

• *Recovery provisions* • *Risk management techniques used to inform the strategy* • *Checks on the consistency of the archive* • *Defined security levels* • *Policy to handle and monitor deterioration of storage*	

4.h.3.3.2 Preservation strategy

Best practice	Found in
Data repository should have a preservation policy (see also 2.2.3.3.1)	*Guideline 6*
Provide procedural documentation for archiving data, including: • *Workflows* • *Decision making process* • *Skills of employees* • *Types of data* • *Selection process* • *Measures regarding data that fall outside the mission* • *Guarding privacy of subjects* • *Information to data producers about handling the data*	*Guideline 8*

4.h.3.3.3 Version control/change procedures

Best practice	Found in
Define a versioning strategy	*Guideline 11*
Define a strategy for data changes and make also data producers aware of this strategy	*Guideline 12*
Maintain provenance data and related audit trails	*Guideline 12*
Compare essential properties of different versions of a file	*Guideline 12*
Procedure to maintain links to metadata and other datasets	*Guideline 12*

4.h.3.4 Dissemination

4.h.3.4.1 Visibility

Best practice	Found in
Provide search facilities, OAIS harvesting, deep searching	*Guideline 10*
Carry out promotional activities	*Guideline 4*

4.h.3.4.2 Availability and accessibility

Best practice	Found in
Offer persistent identifiers	*Guideline 10*
Promote data sharing and reuse	*Guideline 1*

4.h.3.4.3 Tools and interfaces

Best practice	Found in
Use of deposit forms (see also 2.2.3.1.1)	*Guideline 3*

4.h.3.4.4 Monitoring, review and feedback

Best practice	Found in
Monitor availability of data	*Guideline 11*
Monitor chekcsums	*Guideline 11*

4.h.4 Source

Data Seal of Approval Guidelines version 1 June 1, 2010; Assessment Manual, available at: https://assessment.datasealofapproval.org/documentation/

4.i. Appendix I TRAC TDR

Audit and Certification of Trustworthy Digital Repositories. CCSDS Recommended Practice.

The document is designed to be used as a tool for (self-) audit and certification processes for digital repositories which seek to assess and demonstrate their trustworthiness. Building on the RLG-OCLC Report "Trusted Digital Repositories: Attributes and Responsibilities" (2002), it and presents high-level criteria to assess the organizational infrastructure, digital object management, and infrastructures and security risk management in digital repositories. Another important reference for this document is the OAIS reference model, whose terminology is adopted.

CCSDS (2011): Audit and Certfication of Trustworthy Digital Repositories. Recommended Practice. http://public.ccsds.org/publications/archive/652x0m1.pdf

4.i.1 Organizational Framework

4.i.1.1 Purpose and Requirements

Recommendation	Found in
The repository shall have identified and established the duties that it needs to perform and shall have appointed staff with adequate skills and experience to fulfil these duties.	*3.2.1, 3.2.1.1*

4.i.1.1.1 Scope and objectives

Recommendation	Found in
The repository shall have a mission statement that reflects a commitment to the preservation of, long term retention of, management of, and access to digital information.	*3.1.1*
The repository shall have defined its Designated Community and associated knowledge base(s) and shall have these definitions appropriately accessible.	*3.3.1*

4.i.1.1.2 Collection policy

Recommendation	Found in
The repository shall have a Collection Policy or other document that specifies the type of information it will preserve, retain, manage, and provide access to.	*3.1.3*

4.i.1.1.3 Criteria for evaluating data

Recommendation	*Found in*

4.i.1.2 Legal and Regulatory Framework

Recommendation	Found in
The repository shall have an appropriate succession plan, contingency plans, and/or escrow arrangements in place in case the repository ceases to operate or the governing or funding institution substantially changes its scope.	*3.1.2.1*
The repository shall commit to transparency and accountability in all actions supporting the operation and management of the repository that affect the preservation of digital content over time.	*3.3.4*
The repository shall have and maintain appropriate contracts or deposit agreements for digital materials that it manages, preserves, and/or to which it provides access. [These shall] specify and transfer all necessary preservation rights, and those rights transferred shall be documented.	*3.5.1, 3.5.1.1*
The repository shall have specified all appropriate aspects of acquisition, maintenance, access, and withdrawal in written agreements with depositors and other relevant parties […] to ensure that the respective roles of repository, producers, and contributors in the depositing of digital content and transfer of responsibility for preservation are understood and accepted by all parties.	*3.5.1.2*
The repository shall have policies in place to address liability and challenges to ownership/rights […] in order to minimize potential liability and challenges to the rights of the repository. *[These might include] a definition of rights, licenses, and permissions to be obtained from producers and contributors of digital content; citations to relevant laws and regulations; policy on responding to challenges […].*	*3.5.1.4*

4.i.1.3 Funding and Resource Planning

Recommendation	Found in
The repository shall have the appropriate number of staff to support all functions and services.	*3.2.1.2*
The repository shall have in place an active professional development program that provides staff with skills and expertise development opportunities.	*3.2.1.3*
The repository shall have short- and long-term business planning processes in place to sustain the repository over time.	*3.4.1*
The repository shall have financial practices and procedures which are transparent, compliant with relevant accounting standards and practices, and audited by third parties in accordance with territorial legal requirements.	*3.4.2*
The repository shall have an ongoing commitment to analyze and report on financial risk, benefit, investment, and expenditure (including assets, licenses, and liabilities).	*3.4.3*
The repository shall have procedures, commitment and funding to replace	*5.1.1.1.4, 5.1.1.1.8*

hardware [and software] when evaluation indicates the need to do so.	

4.i.1.4 Long-Term Preservation Policy

Recommendation	Found in
The repository shall have a Preservation Strategic Plan that defines the approach the repository will take in the long-term support of its mission […] to help the repository make administrative decisions, shape policies, and allocate resources in order to successfully preserve its holdings.	*3.1.2*
The repository shall have Preservation Policies in place to ensure its Preservation Strategic Plan will be met.	*3.3.2*
The repository shall have mechanisms for review, update, and ongoing development of its Preservation Policies as the repository grows and as technology and community practice evolve.	*3.3.2.1*
The repository shall identify the Content Information and the Information Properties that the repository will preserve […] in order to make it clear to funders, depositors, and users what responsibilities the repository is taking on and what aspects are excluded. It is also a necessary step in defining the information which is needed from the information producers or depositors.	*4.1.1*
The repository shall have a procedure(s) for identifying those Information Properties that it will preserve [in order] to establish a clear understanding with depositors, funders, and the repository's Designated Communities how the repository determines and checks what the characteristics and properties of preserved items will be over the long term. These procedures will be necessary to confirm authenticity or to identify erroneous claims of authenticity of the preserved digital record.	*4.1.1.1*
The repository shall have a record of the Content Information and the Information Properties that it will preserve […]. *This is necessary in order to identify in writing the Content Information of the records for which it has taken preservation responsibility and the Information Properties it has committed to preserve for those records based on their Content Information.*	*4.1.1.2*

4.i.1.5 Access Policy

Recommendation	Found in
The repository shall comply with Access Policies […] to ensure [it] has fully addressed all aspects of usage which might affect the trustworthiness of the repository, particularly with reference to support of the user community.	*4.6.1*

4.i.2 Technological Environment

4.i.2.1 IT Architecture

Recommendation	Found in
The system of [unique persistent] identifiers shall be adequate to fit the repository's current and foreseeable future requirements such as numbers of objects.	*4.2.4.1.5*
The repository shall identify and manage the risks to its preservation operations and goals associated with system infrastructure. *[This can, for example, be achieved by means of] infrastructure inventory of system components; periodic technology assessments; estimates of system component lifetime; export of authentic records to an independent system; use of strongly community supported software e.g., Apache, iRODS, Fedora); re-creation of archives from backups.*	*5.1.1*
The repository shall have hardware [and software] technologies appropriate to the services it provides to its designated communities.	*5.1.1.1.1, 5.1.1.1.5*

4.i.2.2 Standards and Formats

Recommendation	Found in
The repository shall have adequate specifications enabling recognition and parsing of the SIPs. *[This can, for example, take the form of] Packaging Information for the SIPs; Representation Information for the SIP Content Data, including documented file format specifications; published data standards; documentation of valid object construction.*	*4.1.3*
The repository shall have a description of how AIPs are constructed from SIPs.	*4.2.2*
The repository shall have and use a convention that generates persistent, unique identifiers for all AIPs [so as to] uniquely identify each AIP [and its components] within the repository.	*4.2.4, 4.2.4.1.1, 4.2.4.1.2*

4.i.2.3 Security and Risk Management / Media Monitoring and Refreshing Strategy

Recommendation	Found in
The repository shall have adequate hardware and software support for backup functionality sufficient for preserving the repository content and tracking repository functions. *[This can include, for example,] documentation of what is being backed up and how often; audit log/inventory of backups; validation of completed*	*5.1.1.2*

backups; disaster recovery plan, policy and documentation; fire drills; testing of backups; support contracts for hardware and software for backup mechanisms; demonstrated preservation of system metadata such as access controls, location of replicas, audit trails, checksum values.	
The repository shall have a process to record and react to the availability of new security updates based on a risk-benefit assessment.	*5.1.1.4*
The repository shall have defined processes for storage media and/or hardware change (e.g., refreshing, migration).	*5.1.1.5*
The repository shall have identified and documented critical processes that affect its ability to comply with its mandatory responsibilities […] to ensure that the critical processes can be monitored to ensure that they continue to meet the mandatory responsibilities and to ensure that any changes to those processes are examined and tested.	*5.1.1.6*
The repository shall have a documented change management process that identifies changes to critical processes that potentially affect the repository's ability to comply with its mandatory responsibilities.	*5.1.1.6.1*
The repository shall have a process for testing and evaluating the effect of changes to the repository's critical processes.	*5.1.1.6.2*
The repository shall maintain a systematic analysis of security risk factors associated with data, systems, personnel, and physical plant.	*5.2.1*
The repository shall have implemented controls to adequately address each of the defined security risks.	*5.2.2*
The repository staff shall have delineated roles, responsibilities, and authorizations related to implementing changes within the system.	*5.2.3*
The repository shall have suitable written disaster preparedness and recovery plan(s), including at least one off-site backup of all preserved information together with an offsite copy of the recovery plan(s).	*5.2.4*

4.i.3 Data Curation

4.i.3.1 Pre-Ingest Function

Recommendation	Found in
The repository shall have contemporaneous records of actions and administration processes that are relevant to content acquisition.	*4.1.8*

4.i.3.1.1 Information and guidance given to data producer

Recommendation	Found in
The repository shall have written policies that indicate when it accepts preservation responsibility for contents of each set of submitted data objects […] to avoid misunderstandings between the repository and producer/depositor as to when and how the transfer of responsibility for the digital content occurs. *[This can be demonstrated, for example, by] properly executed submission agreements, deposit agreements, and deeds of gift; confirmation receipt sent back to producer/depositor.*	*3.5.1.3*

The repository shall provide the producer/depositor with appropriate responses at agreed points during the ingest processes.	*4.1.7*

4.i.3.2 Ingest Function

Recommendation	Found in
The repository shall document the final disposition of all SIPs. In particular […,] the repository shall follow documented procedures if a SIP is not incorporated into an AIP or discarded and shall indicate why the SIP was not incorporated or discarded.	*4.2.3, 4.2.3.1*

4.i.3.2.1 Information and documentation from data producer

Recommendation	Found in
The repository shall clearly specify the information that needs to be associated with specific Content Information at the time of its deposit [to achieve] a clear understanding of what needs to be acquired from the Producer.	*4.1.2*

4.i.3.2.2 Quality assurance and data checking

Recommendation	Found in
The repository shall have mechanisms to appropriately verify the identity of the Producer of all materials […] to avoid providing erroneous provenance to the information which is preserved. *[This can be achieved, for example, with the help of] legally binding submission agreements/deposit agreements/deeds of gift, evidence of appropriate technological measures; logs from procedures and authentications.*	*4.1.4*
The repository shall have an ingest process which verifies each SIP for completeness and correctness. *[This can be achieved, for example, with the help of] system log files from system(s) performing ingest procedure(s); logs or registers of files received during the transfer and ingest process; documentation of standard operating procedures, detailed procedures, and/or workflows; format registries; definitions of completeness and correctness.*	*4.1.5*

4.i.3.2.3 Data documentation and enhancement

Recommendation	Found in
The repository shall have for each AIP or class of AIPs preserved by the repository an associated definition [including appropriate Packaging Information] that is adequate for parsing the AIP and fit for long-term preservation needs.	*4.2.1*
The repository shall be able to identify which definition applies to which AIP. [This can be achieved, for example, by generating and employing] documentation clearly linking each AIP, or class of AIPs, to its definition.	*4.2.1.1*
The repository shall have a definition of each AIP that is adequate for long-term preservation, enabling the identification and parsing of all the	*4.2.1.2*

required components within that AIP. *[The definitions shall be used] to extract Content Information and PDI (Provenance, Access Rights, Context, Reference, and Fixity Information) from AIPs.*	
The repository shall have access to necessary tools and resources to provide authoritative Representation Information for all of the digital objects it contains. *[This includes] tools or methods to identify the file type of all submitted Data Objects [and] to determine what Representation Information is necessary to make each Data Object understandable to the Designated Community.*	*4.2.5, 4.2.5.1, 4.2.5.2*
The repository shall have access to the requisite Representation Information.	*4.2.5.3*
The repository shall have tools or methods to ensure that the requisite Representation Information is persistently associated with the relevant Data Objects.	*4.2.5.4*
The repository shall have documented processes for acquiring Preservation Description Information (PDI) for its associated Content Information and acquire PDI in accordance with the documented processes.	*4.2.6, 4.2.6.1, 4.2.6.2*
The repository shall ensure that the PDI is persistently associated with the relevant Content Information.	*4.2.6.3*
The repository shall capture or create minimum descriptive information and ensure that it is associated with the AIP.	*4.5.2*

4.i.3.3 Archival Storage and Preservation

Recommendation	Found in
The repository shall obtain sufficient [legal and physical] control over the Digital Objects to preserve them.	*4.1.6*
The repository shall be able to provide a complete list of all [unique persistent] identifiers and do spot checks for duplications.	*4.2.4.1.4*
The repository shall verify each AIP for completeness and correctness at the point it is created.	*4.2.8*
The repository shall have contemporaneous records of actions and administration processes that are relevant to AIP creation.	*4.2.10*
The repository shall preserve the Content Information of AIPs.	*4.4.1.1*
The repository shall have contemporaneous records of actions and administration processes that are relevant to storage and preservation of the AIPs.	*4.4.2*
The repository shall maintain bi-directional linkage between each AIP and its descriptive information [over time] to ensure that all AIPs can be located and retrieved.	*4.5.3, 4.5.3.1*

4.i.3.3.1 Physical data preservation and storage

Recommendation	Found in
The repository shall define, collect, track, and appropriately provide its information integrity measurements.	*3.3.5*

The repository shall have a system of reliable linking/resolution services in order to find the uniquely identified object, regardless of its physical location [e.g. by means of naming conventions].	*4.2.4.2*
The repository shall provide an independent mechanism for verifying the integrity of the repository collection/content [in order] to enable the audit of the integrity of the collection as a whole.	*4.2.9*
The repository shall have specifications for how the AIPs are stored down to the bit level.	*4.4.1*
The repository shall actively monitor the integrity of AIPs [e.g. with the help of checksums].	*4.4.1.2*
The repository shall have effective mechanisms to detect bit corruption or loss.	*5.1.1.3*
The repository shall record and report to its administration all incidents of data corruption or loss, and steps shall be taken to repair/replace corrupt or lost data.	*5.1.1.3.1*
The repository shall manage the number and location of copies of all digital objects […] in order to assert that the repository is providing an authentic copy of a particular digital object.	*5.1.2*

4.i.3.3.2 Preservation strategy

Recommendation	Found in
The repository shall have documented preservation strategies relevant to its holdings.	*4.3.1*
The repository shall have mechanisms to change its preservation plans as a result of its monitoring activities.	*4.3.3*

4.i.3.3.3 Version control/change procedures

Recommendation	Found in
The repository shall have a documented history of the changes to its operations, procedures, software, and hardware […] in order to provide an 'audit trail' through which stakeholders can identify and trace decisions made by the repository.	*3.3.3*
Documentation shall describe any processes used for changes to [unique persistent] identifiers.	*4.2.4.1.3*
The repository shall have procedures for all actions taken on AIPs […] in order to ensure that any actions performed against an AIP do not alter the AIP information in a manner unacceptable to its Designated Communities.	*4.4.2.1*
The repository shall be able to demonstrate that any actions taken on AIPs were compliant with the specification of those actions.	*4.4.2.2*
The repository shall have mechanisms in place to ensure any/multiple copies of digital objects are synchronized.	*5.1.2.1*

4.i.3.4 Dissemination

Recommendation	Found in
The repository shall specify minimum information requirements to enable the Designated Community to discover and identify material of interest.	*4.5.1*
The repository shall follow policies and procedures that enable the dissemination of digital objects that are traceable to the originals, with evidence supporting their authenticity […]. This is necessary to establish an auditable chain of authenticity from the AIP to disseminated digital objects.	*4.6.2*

4.i.3.4.1 Visibility

Recommendation	*Found in*

4.i.3.4.2 Availability and accessibility

Recommendation	*Found in*

4.i.3.4.3 Tools and interfaces

Recommendation	*Found in*

4.i.3.4.4 Monitoring, review and feedback

Recommendation	Found in
The repository shall monitor its organizational environment to determine when to execute its succession plan, contingency plans, and/or escrow arrangements.	*3.1.2.2*
The repository shall commit to a regular schedule of self-assessment and external certification.	*3.3.6*
The repository shall track and manage intellectual property rights and restrictions on use of repository content as required by deposit agreement, contract, or license.	*3.5.2*
The repository shall ensure that the Content Information of the AIPs is understandable for their Designated Community at the time of creation of the AIP. In particular, the […] repository shall have [and execute] a documented process for testing understandability for their Designated Communities of the Content Information of the AIPs at their creation.	*4.2.7, 4.2.7.1, 4.2.7.2*
The repository shall bring the Content Information of the AIP up to the required level of understandability if it fails the understandability testing.	*4.2.7.3*
The repository shall have mechanisms in place for monitoring its preservation environment […] so that the repository can react to changes and thereby ensure that the preserved information remains understandable and usable by the Designated Community.	*4.3.2*
The repository shall have mechanisms in place for monitoring and notification when Representation Information is inadequate for the Designated Community to understand the data holdings.	*4.3.2.1*

The repository shall have mechanisms for creating, identifying or gathering any extra Representation Information required.	*4.3.3.1*
The repository shall provide evidence of the effectiveness of its preservation activities […] to assure the Designated Community that the repository will be able to make the information available and usable over the mid-to-long-term.	*4.3.4*
The repository shall log and review all access management failures and anomalies.	*4.6.1.1*
The repository shall record and act upon problem reports about errors in data or responses from users.	*4.6.2.1*
The repository shall employ technology watches or other technology monitoring notification systems.	*5.1.1.1*
The repository shall have procedures in place to monitor and receive notifications when hardware technology [or software] changes are needed.	*5.1.1.1.2, 5.1.1.1.6*
The repository shall have procedures in place to evaluate when changes are needed to current hardware [or software].	*5.1.1.1.3, 5.1.1.1.7*

4.j Appendix J

Planning Tool for Trusted Electronic Repositories (PLATTER)

The PLATTER is a framework, which provides a basis for a digital repository to plan the development of its goals, objectives and performance targets over the course of its lifetime in a manner which will contribute to the repository establishing trusted status amongst its stakeholders. PLATTER cannot be perceived as an audit or certification tool itself, but rather a framework that will allow new repositories to incorporate the goal of achieving trust into their planning from an early stage. A repository planned using PLATTER will find itself in a strong position when it subsequently comes to apply one of the existing auditing tools to confirm the adequacy of its procedures for maintaining the long term usability of and access to its material.

The PLATTER acknowledges the diversity of the organisations, which may be included under the term "digital repository" by requiring repositories to answer a questionnaire which characterises the repository relative to other repositories and which can be used to determine how and whether the goals and objectives we have identified are to be realised in a given organisation.

The PLATTER process is centred around a group of Strategic Objective Plans (SOPs) through which a repository specifies its current objectives, targets, or key performance indicators in those areas which have been identified as central to the process of establishing trust.

4.j.1. Organizational Framework

4.j.1.1 Purpose and Requirements

Recommendation	Found in
Determine the general functional type of the repository: *What is the source of the repository's mandate? Is the Repository for profit or non-profit? Does the Repository receive a significant proportion of its material from a legally mandated source? What is the operational status of the repository (not yet running, running but still under development, mature)*	*Recommendation 3.1.*

4.j.1.1.1 Scope and objectives

Recommendation	Found in
Define the overall scale of the repository, whether expressed in human, technical, or financial terms: *What is the amount of digital material you expect to archive? How many distinct digital objects do you expect to archive? How many fulltime-equivalent staff does the Repository expect to employ? How many distinct end-users are expected to access material in the Repository over the course of a calendar year?*	*Recommendation 3.2.*

Is the majority of the material in the Repository simple, moderately complex, or highly complex?	*Recommendation 3.3 Question Q3.2*
How specialized is the data in the Repository (low, medium or high)?	*Recommendation 3.3 Question Q3.3*
How sensitive is the most sensitive material in the Repository (low, medium, high)?	*Recommendation 3.3 Question Q3.4*
In which of the three access classes (open, restricted, closed) does the Repository have significant holdings?	*Recommendation 3.3 Question Q3.5*

4.j.1.1.2 Collection policy

Recommendation	Found in
Which of the three acquisition strategies (push, pull, self-creation) account for a significant portion of the total material in the Repository?	*Recommendation 3.3 Question Q3.1*
Acquire relevant material: specify quantitative targets for the material to be acquired by the repository	*Goal 2.1*
Negotiate deposit agreements: Scope of material, Delivery form, File formats, Accompanying metadata, etc.	*Goal 2.2*
Obtain Physical Control of Materials	*Goal 2.3*
It is essential that a repository have in place a monitoring system to determine that the required material is actually made available by the producer or depositors.	*Goal 2.4*
The repository should institute procedures to monitor the relevance of any deposit agreement, taking into account the same issues considered in the initial development of the agreement	*Goal 2.5*

4.j.1.1.3 Criteria for evaluating data

Recommendation	*Found in*
N.A.	

4.j.1.2 Legal and Regulatory Framework

Recommendation	*Found in*
N.A.	

4.j.1.3 Funding and Resource Planning

Recommendation	Found in
In order for the business plan to remain current, regular review is needed. As sources of income cannot normally be guaranteed for more than a few years, monitoring is needed to identify such shortages before they develop into a budget deficiency.	*Goal 1.1*
The repository must achieve an income sufficient for routine functioning.	*Goal 1.2*
Ensure contingency plans for financial cutbacks or emergencies: have an agreement with another repository about housing vital data; prioritise services to be retained in the case of financial problems	*Goal 1.3*
Define and maintain marketing and outreach plans suitable for the	*Goal 1.4*

repository's needs	

4.j.1.4 Long-Term Preservation Policy

Recommendation	Found in
The preservation tasks are ensured even beyond the existence of the digital repository by an agreement with other repositories to take over the collections. The very purpose of the succession plan(s) is to detail agreement(s) about who will inherit the digital data if the repository ceases to function.	*Goal 7.1*

4.j.1.5 Access Policy

Recommendation	Found in
Create, Maintain and Review a Mission Statement which reflects the Repository's Mandate	*Goal 4.1*
The definition of the designated community should be available on the repository website and the community should be very well aware of available delivery and access options. It is essential to monitor and reflect all changes inside the community over the time.	*Goal 4.2*

4.j.2. Technological Environment

4.j.2.1 IT Architecture

Recommendation	Found in
IT Infrastructure must be capable of coping with the scale of data storage, processing and transport appropriate for the repository.	*Goal 5.1*
IT infrastructure must be able to guarantee the integrity and security of the stored data.	*Goal 5.2*
The IT infrastructure must be able to guarantee that certain services remain available to the users. Identify which services of the repository are no longer required and can be closed and which are in great demand.	*Goal 5.3*

4.j.2.2 Standards and Formats

Recommendation	Found in
What are the main sources of bibliographic and descriptive metadata in the repository?	*Recommendation 3.4 Question Q4.1*
What interoperability standards are implemented in the Repository?	*Recommendation 3.4 Question Q4.2*
A data provider encodes his data in a package format (Submission Information Package) acceptable to the repository. The repository receives these SIPs, and repackages them for storage (Archival Information Package). A repository user requests the data, and the repository repackages it in a format appropriate for the user (Dissemination Information Package).	*Recommendation 5.6.*
For Submission Information Package: Specify the digital object formats and accompanying meta-data the repository will accept. Formats (potentially) containing encryption, or embedded objects and files, as well as commercially protected formats can sometimes prove impossible to convert.	*Goal 6.1*
For Archival Information Package: Specify the data format and metadata content for archiving digital objects. Where the SIPs defined in the previous	*Goal 6.2*

goal should strike some balance between commonly used formats, and formats useful for archiving, the AIPs do not need to.	
For Dissemination Information Package: Specify the data formats used for disseminating digital objects.	*Goal 6.3*
Specify the transformation from SIP to AIP and from AIP to DIP.	*Goal 6.4 and 6.5*

4.j.2.3 Security and Risk Management / Media Monitoring and Refreshing Strategy

Recommendation	Found in
The repositories must have strategies to deal with budget shortfalls. Common strategies include willing creditors to boost the repository for a period, or having savings. The repository should also have a prioritized list of the services it provides, beyond what is absolutely required of it, to ease in selecting which could be terminated, if income shortages arrive.	*Goal 8.1*
Political upheaval can come to seemingly stable societies. Repositories caught in such circumstances could face opposing demands to both open up their collections, and to close them down.	*Goal 8.1*
Many repositories are a part of a larger organization, which due to some changes could have problems seeing the purpose of the repository. The repository must actively seek to demonstrate its value to those with executive power over it.	*Goal 8.1*
The repository should, so far as possible, detail procedures for dealing with environmental disasters in a timely manner.	*Goal 8.1*
Loss of educated key staff: • *the internal workings and trade secrets of the repository could be revealed to outsiders – the repository should prepare press statements and other communications for when the secrets break* • *the services of the repository could cease to function - the best course of action would be to shut down unmaintainable services until new staff have been hired or trained.*	*Goal 8.1*
The physical security of the repository can also be threatened. The repository should identify likely threats, and address them, in addition to having plans in the event of these threats materializing.	*Goal 8.1*

4.j.3. Data Curation

4.j.3.1 Pre-Ingest Function

Recommendation	*Found in*
N.A.	

4.j.3.1.1 Information and guidance given to data producer

Recommendation	Found in
The repository should have some description for content providers about how to package data and metadata or representation information. The repository should have a policy regarding the completeness and correctness of new data and what action to take regarding invalid or incomplete data.	*Goal 6.1*
Specify sources and formats for bibliographic and descriptive metadata in	*Goal 6.1.1*

the received data. This is the basic information required to enable the object to be discovered (e.g. by a search engine) and interpreted.	
Specify technical metadata in the SIP	*Goal 6.1.2*

4.j.3.2 Ingest Function

Recommendation	*Found in*
N.A.	

4.j.3.2.1 Information and documentation from data producer

Recommendation	*Found in*
N.A.	

4.j.3.2.2 Quality assurance and data checking

Recommendation	*Found in*

4.j.3.2.3 Data documentation and enhancement

Recommendation	Found in
The repository should have a method for verifying the completeness and correctness for newly generated AIPs, as with SIPs.	*Goal 6.2*
The metadata from the SIP should be conserved, and elaborated, rather than being changed. A file format, that strikes a balance between being human readable and being machine parseable, such as XML could be used to encode the metadata.	*Goal 6.2.1*

4.j.3.3 Archival Storage and Preservation

Recommendation	*Found in*
N.A.	

4.j.3.3.1 Physical data preservation and storage

Recommendation	Found in
What strategy is used for storage? (in-house, external, in-house under external support)?	*Recommendation 3.4* *Question Q4.3*

4.j.3.3.2 Preservation strategy

Recommendation	Found in
What strategy is used for software management? • *support by the software supplier* • *support by a third party* • *self-support (i.e. in-house)* • *support by a user and developer community*	*Recommendation 3.4* *Question Q4.4*
Repository must maintain understanding of all structural (e.g. file encoding) standards and formats	*Goal 9.2*
Repository must maintain understanding of contemporary and emerging hardware, software and storage technologies	*Goal 9.1*

Repository must maintain understanding of identified user communities and their associated competences and knowledge base	*Goal 9.3*
Repository must maintain understanding of preservation requirements for each stored information asset or class of information	*Goal 9.4*
Repository must maintain, exercise and evaluate preservation strategies capable of meeting specific preservation targets	*Goal 9.5*
Repository must maintain and exercise appropriate appraisal policies to determine which information must continue to be preserved	*Goal 9.6*

4.j.3.3.3 Version control/change procedures

Recommendation	Found in
N.A.	

4.j.3.4 Dissemination

Recommendation	Found in
N.A.	

4.j.3.4.1 Visibility

Recommendation	Found in
N.A.	

4.j.3.4.2 Availability and accessibility

Recommendation	Found in
Create and Implement A Repository Access Policy taking into account: copyright law, law on dissemination of personal or commercially-sensitive data, National security, Libel, obscenity, hate-speech and blasphemy laws, Specific contractual restrictions imposed in deposit agreements	*Goal 4.3*
Specify and fulfil technical requirements for dissemination and access	*Goal 4.4*

4.j.3.4.3 Tools and interfaces

Recommendation	Found in
N.A.	

4.j.3.4.4 Monitoring, review and feedback

Recommendation	Found in
The repository should periodically evaluate usage statistics and watch for services going out of favour with the users. Giving the users a way to request new features could also help the repository staying current.	*Goal 8.1*
Technological changes can greatly affect the business model of a repository, and at times come about very quickly. Repositories need to watch the technological trends, evaluate which could be potentially harmful, and adapt.	*Goal 8.1*

4.K Appendix K DigiPlanet

The Blue Ribbon Task Force on Sustainable Digital Preservation and Access

The Blue Ribbon Task Force on Sustainable Digital Preservation and Access was created in late 2007 to address the following issues:

- the assumption that key information will be there when we want it, where we want it, and for the foreseeable future
- how will we ensure the long-term preservation and access to our digital information, growing exponentially with each passing day?; how will we successfully migrate data as technology moves from one preservation medium to the next?; who should determine which digital data should be saved, and what criteria will be used to make those decisions?
- economic sustainability - what is the cost to preserve valuable data and who will pay for it?

In February 2010, the Taskforce published its Final Report, called "Sustainable Economics for a Digital Planet: Ensuring Long-Term Access to Digital Information." The report provides general principles and actions to support long-term economic sustainability; context-specific recommendations tailored to specific scenarios analyzed in the report; and an agenda for priority actions and next steps, organized according to the type of decision maker best suited to carry that action forward. It is based on an investigation into preservation in the United States and the United Kingdom.

The Task Force has developed a general framework for analysing digital preservation as an economic problem. It has also identified problems intrinsic to all preserved digital materials e.g. long-term horizons, diffused stakeholders, misaligned or weak incentives and lack of clarity about roles and responsibilities amongst stakeholders. Four key digital preservation contexts are considered:

- *Scholarly discourse:* the published output of scholarly inquiry
- *Research data:* the primary inputs into research, as well as the first-order results of the data
- *Commercially owned cultural content:* culturally significant digital content that is owned by a private entity and is under copyright protection
- *Collectively produced Web content:* Web content that is created interactively, the result of collaboration and contributions by consumers.

Finally, the Task Force propose action points for stakeholders to meet challenges to sustainability.

4.k.1 Organizational Framework

4.k.1.1 Purpose and Requirements

Recommendation	*Found in*
N.A.	

4.k.1.1.1 Scope and objectives

Recommendation	Found in
N.A.	

4.k.1.1.2 Collection policy

Recommendation	*Found in*
N.A.	

4.k.1.1.3 Criteria for evaluating data

Recommendation	*Found in*
N.A.	

4.k.1.2 Legal and Regulatory Framework

Recommendation	Found in
One area of priority for near-term action in achieving sustainability concerns public policy action (including modifying copyright laws and to enable digital preservation; creating incentives and requirements for private entities to preserve on behalf of the public; sponsoring public-private partnerships; clarifying rights issues associated with Web-based materials; empowering stewardship organisations to protect digital orphans form unacceptable loss).	*Executive Summary (p.5)*
Issues surrounding copyright laws for digital data need careful	*3.2.2 (p.40-41)*

consideration when choosing preservations strategies. *"The US Copyright office and the Library of Congress have made recommendations3 to change the law [of copyrighting digital materials] in light of digital technology realities." The DigiPlanet team "urge Congress to take up this matter expeditiously".*	
An important choice available to decision makers in designing suitable preservation strategies is to determine who owns the digital materials.	*2.2.2 (p.33)*

4.k.1.3 Funding and Resource Planning

Recommendation	Found in
Define roles and responsibilities among stakeholders to ensure an ongoing and efficient flow of resources to preservation throughout the digital lifecycle.	*Executive summary (p.2-3); 2.1.3 (p.22)*
Funding models should be tailored to the norms and expectations of anticipated users.	*Executive Summary (p.3)*
One area of priority for near-term action in achieving sustainability concerns organisational action (including developing public-private partnerships; ensuring that organisations have access to skilled personnel, from domain experts to legal and business specialists; creating and sustaining secure chains of stewardship between organisations over time; achieving economies of scale and scope; addressing the free-rider problem).	*Executive summary (p.4)*
The selection of an appropriate funding model requires an in-depth knowledge of the circumstances surrounding the effort, preservation goals, the stakeholder community etc.	*Section 3.2.4 (p. 44)*
Stimulating growth of capacity and funding to meet the demand [of preservation] is crucial.	*2.1.3 (p.23)*

[3] "One of the recommendations of The Section 108 Study Group Report by the United States Copyright Office and the National Digital Information Infrastructure and Preservation Program of the Library of Congress is to qualify museums for archival exception for the first time. See:

http://www.section108.gov/docs/Sec108StudyGroupReport.pdf" (page 40).

Preservation works best when the interests and actions of users, owners, and archives can be aligned in an economic strategy and operationalized in a business model.	*2.2.1 (p.29-30)*
An important choice available to decision makers in designing suitable preservation strategies is to determine who preserves the digital materials.	*2.2.2 (p.33)*
An important choice available to decision makers in designing suitable preservation strategies is to determine who pays for the preservation of the digital materials.	*2.2.2 (p.34)*
Advice on what to do when roles and responsibilities among stakeholders and actors are unclear.	*3.2.3 (p.43)*
There is no single "best" funding model for digital preservation. Common funding models for digital preservation outlined as well as strategies to reduce costs.	*3.2.4 & 3.2.6 (p.44-47)*
Condition of sustainability: appropriate organisation and governance of preservation activities. Principle of action: Roles and responsibilities among stakeholders must be clear, transparent, and well integrated; and handoffs between responsible parties must be ensured at key moments of risk in the digital lifecycle. Actions: Create effective governance mechanisms to aggregate and rationalize collective preservation interests and costs. Create mechanisms to address free-rider problems in the provision of preservation.	*5.1 (p.78)*
Condition of sustainability: ongoing and efficient allocation of resources to preservation. Principle of action: Funding models must reflect community norms, be flexible to adjust to disruptions over time, and leverage economies of scale and scope as appropriate. Actions: Choose funding models according to norms and expectations of anticipated users; leverage economies of scale and scope; lower costs of preservation overall.	*5.1 (p.78-79)*
Action agenda for National and International Archives: *1. Create mechanisms for public-private partnerships to align or reconcile benefits that accrue to commercial and cultural entities. These agencies can play a critical role in convening stakeholders, sponsoring cooperation and collaboration, and ensuring representation of all stakeholders.* *2. Convene expert communities to address the selection and preservation needs of materials of particular interest to the public for which there is no stewardship (Web materials, digital orphans).*	*5.3 (p.82)*

3. Act expeditiously to reform national and international copyright legislation to address digital preservation needs. *4. Create financial incentives to encourage private entities to preserve digital materials on the public behalf.*	
Action agenda for Funders and Sponsors of Data creation: *1. Create preservation mandates when possible, ensuring that they adhere to community selection criteria, and specifying roles and responsibilities of individuals and organizations.* *2. Invest in building capacity throughout the system. The Library of Congress, the National Archives and Records Administration, the National Science Foundation, and JISC have set important precedents for supporting capacity building within specific communities of practice. Seeding stewardship capacity and developing sustainable funding models should, however, be a high priority for all funders.* *3. Provide leadership in training and education for 21st century preservation, including domain expertise and core competencies in STEM. Such organizations as the National Archives, Library of Congress, National Library of Medicine, National Agricultural Library, National Science Foundation, Smithsonian Institution, Institute of Museum and Library Services, National Endowment for the Arts, and National Endowment for the Humanities in the United States; and the British Library, National Archives, JISC, Digital Curation Centre, and Digital Preservation Coalition in the United Kingdom each have a remit for promoting digital preservation skills.* *4. Fund the modeling and testing of domain-specific preservation strategies. This would entail developing domain-specific requirements for lifecycle management to create a timeline of predictable risks, strategies to meet them, and triggering mechanisms to address them.*	*5.3 (p.82)*
Action agenda for Stakeholder organisations: *1. Secure preservation of high-value institutional materials by making explicit roles and responsibilities across organizational boundaries.* *2. Develop preservation strategies that assign responsibilities for achieving outcomes. Service-level agreements and MOUs with third-party archives should include contingency plans for handoffs and clauses for putting internal monitoring systems in place.*	*5.3 (p.83)*

3. Leverage resources; create economies of scope and economies of scale by partnering with related organizations and industry professional associations. *4. Work with domain and preservation experts to ensure that personnel are fully equipped with the technical skills needed for selecting, curating, and preserving materials.* *5. Fund internal preservation and access activities as core infrastructure.*	
Action agenda for individuals: *1. Provide nonexclusive rights to preserve content they create and to distribute this content through publicly accessible venues.* *2. Partner with preservation experts early in the lifecycle of one's own digital data, to ensure that data are ready to hand off to an archive in forms that will be useful over the long term.* *3. Actively participate in professional societies and relevant organizations in developing stewardship best practices and selection priorities.*	*5.3 (p.83)*
Flexibility in preservation decision making is necessary as investment in digital preservation is not necessarily once-and-for-all, or all-or-nothing. Making clear that this flexibility exists when an investment is first considered may make the scope of the perceived economic commitment less daunting, and correspondingly increase decision makers' willingness to make this commitment.	*Appendix 6 (p. 100)*
Define roles and responsibilities among stakeholders to ensure an ongoing and efficient flow of resources to preservation throughout the digital lifecycle.	*Executive summary (p.2-3); 2.1.3 (p.22)*

4.k.1.4 Long-Term Preservation Policy – general advice for sustainable digital preservation

Recommendation	Found in
Consider the broader economic setting in which preservation decisions are made as these are fundamental in shaping the prospects for achieving long-term sustainability.	*Preface (p.6)*
Articulate a compelling value proposition– e.g. when making the case for preservation, make the case for use.	*Executive Summary (p.1); 5.1 (p.75-76)*

Condition of sustainability: recognition of the benefits of preservation by decision makers. Principle: To make a case for preservation, make the case for use. Actions: Articulate the value proposition in terms of use cases; identify at-risk materials whose damage or loss is unacceptable; use proxy organizations to aggregate diffuse demand; use option strategies in cases of uncertain value.	
Provide clear incentives to preserve in the public interest. Condition of sustainability: incentives for decision makers to act in the public interest. Principle of action: Incentives must be strengthened when they are weak; aligned when they diverge among different stakeholder communities; and created where there are none. Actions: When there are insufficient incentives, use preservation mandates as appropriate. Provide financial incentives for private owners to preserve on behalf of the public. Bring copyright law and mandatory deposit requirements up to date for digital preservation. Remove barriers to creating efficient decentralized stewardship mechanisms by use of nonexclusive licenses granting preservation rights to third parties.	*Executive Summary (p.2); 5.1 (p.77)*
One area of priority for near-term action in achieving sustainability concerns Education and Public Outreach action (including: promoting education and training for 21st century digital preservation and raising awareness of the urgency to take timely preservation actions).	*Executive Summary (p.5)*
Five conditions required for economic sustainability of digital preservation are: - *recognition of the benefits of preservation by decision makers;* - *a process for selecting digital materials with long-term value;* - *incentives for decision makers to preserve in the public interest;* - *appropriate organization and governance of digital preservation activities; and* - *mechanisms to secure an ongoing, efficient allocation of resources to digital preservation activities.*	*1.2.1 (p.12); 5.1 (p.74)*
A sustainable preservation strategy must be flexible enough to span generations of data formats, access platforms, owners and users.	*2.0 (p.17)*
Demand for preservation must be articulated well enough to ensure there is sufficient supply.	*2.1 (p.18)*
It is important to frame the demand for preservation services as a demand for the product of those services (the digital materials) for future use.	*2.2.1 (p.25)*
The decision to preserve will ultimately be based on the perceived value associated with the digital materials over time. We must make ongoing	*2.2.1 (p.25-26)*

investments in their maintenance to sustain their value-creating capacity over time.	
Preserving digital orphans is a priority for public institutions. ('Digital orphans' is the name given to materials on the Web whose ownership or provenance are uncertain).	*3.2.2 (p.42)*
Condition of sustainability: selection of materials with long-term value *Principle of action: Priorities must be made for selecting materials that have the greatest promise of returning value to users over time.* *Action: Make decisions about priorities among collections competing for scarce resources based on projected future use.*	*5.1 (p.76)*
Policy interventions for digital preservation can be necessary when the level of investment in preservation by private decision-makers to meet their own needs falls short of what is optimal for society.	*Appendix 7 (p.101)*
Consider the broader economic setting in which preservation decisions are made as these are fundamental in shaping the prospects for achieving long-term sustainability.	*Preface (p.6)*
Articulate a compelling value proposition– e.g. when making the case for preservation, make the case for use. *Condition of sustainability: recognition of the benefits of preservation by decision makers. Principle: To make a case for preservation, make the case for use. Actions: Articulate the value proposition in terms of use cases; identify at-risk materials whose damage or loss is unacceptable; use proxy organizations to aggregate diffuse demand; use option strategies in cases of uncertain value.*	*Executive Summary (p.1); 5.1 (p.75-76)*

4.k.1.5 Access Policy

Recommendation	Found in
One way to strengthen preservation incentives is for stakeholders to aggregate and leverage their demand-side power to negotiate preservation requirements as part of access arrangements. An example is the NESLi2 Model License for Journals in the United Kingdom, used by the Joint Information Systems Committee (JISC) in negotiating e-journal licenses on behalf of UK higher education institutions (See	3.2.2 (p.41)

http://www.nesli2.ac.uk)	

4.k.2 Technological Environment

4.k.2.1 IT Architecture

Recommendation	Found in
One area of priority for near-term action in achieving sustainability concerns technical action (including building capacity to support stewardship in all areas; lowering the cost of preservation overall; determining the optimal level of technical curation to operationalise an option strategy for all types of digital material.	*Executive Summary (p.4-5)*

4.k.2.2 Standards and Formats

Recommendation	*Found in*
N.A.	

4.k.2.3 Security and Risk Management / Media Monitoring and Refreshing Strategy

Recommendation	Found in
Advice on what to do when the benefits from assets no longer outweigh the costs of maintaining them e.g. the best practice for all institutions is to have a plan for accessing long-term value periodically, at which point they decide whether to renew, handoff, or abandon a preservation commitment.	*3.2.4 (p.43-44)*

4.k.3 Data Curation

4.k.3.1 Pre-Ingest Function

Recommendation	*Found in*
N.A.	

4.k.3.1.1 Information and guidance given to data producer

Recommendation	*Found in*
N.A.	

4.k.3.2 Ingest Function

Recommendation	*Found in*
N.A.	

4.k.3.2.1 Information and documentation from data producer

Recommendation	*Found in*
N.A.	

4.k.3.2.2 Quality assurance and data checking

Recommendation	*Found in*
N.A.	

4.k.3.2.3 Data documentation and enhancement

Recommendation	Found in
N.A.	

4.k.3.3 Archival Storage and Preservation

Recommendation	Found in
The path-dependent nature of preservation decision making means that decisions made at any time shape future conditions and determine the range of future choices. Diagram showing traditional preservation path v. Digital Preservation Lifestyle can be found in 2.2.1 (p.29).	*2.2.1 (p.29-30)*
Strategies for overcoming problems with supply of preserved digital materials include lowering barriers to preserve; giving incentives to private parties to preserve in the public interest and imposing mandates to preserve.	*3.2.1 & 3.2.2 (p.38-41)*
There is an urgent need in all sectors of digital creation—public and private, cultural and scientific—for support in the near-term to model and test robust preservation strategies. All stakeholder communities must provide leadership and accept responsibility for the development of a common digital preservation infrastructure that is sustainable for generations to come.	*5.3 (p.81)*
The path-dependent nature of preservation decision making means that decisions made at any time shape future conditions and determine the range of future choices. Diagram showing traditional preservation path v. Digital Preservation Lifestyle can be found in 2.2.1 (p.29).	*2.2.1 (p.29-30)*

4.k.3.3.1 Physical data preservation and storage

Recommendation	Found in
One of the choices available to decision makers in designing suitable preservation strategies is to determine who selects what to preserve.	*2.2.2 (p.32)*

4.k.3.3.2 Preservation strategy – for sustainable digital preservation specific to the four 'content domains' listed on p.2 of this summary

Recommendation	Found in
Actions recommended for sustainability – Scholarly discourse -	*Executive Summary (p.3)*

There are particular needs to align preservation incentives among commercial and non-profit providers; ensure handoffs between commercial publishers and stewardship organizations in the interest of long-term preservation of the scholarly record; and address the free-rider problem. Clarification of the long-term value of emerging genres of digital scholarship, such as academic blogs and grey literature, is a high priority. Research and education institutions, professional societies, publishers, libraries, and scholars all have leading roles to play in creating sustainable preservation strategies for the materials that are valuable to them.	
Actions recommended for sustainability – Research data - there are few robust systems for making decisions about what to preserve; and there is often a lack of coordination of roles, responsibilities, and funding sources among those best positioned to preserve data (researchers) and the preservation infrastructure (curation and archiving services) that should support them. Research and education institutions, professional societies, archives, researchers, and the funding agencies that support data creation all have leading roles to play in creating sustainable preservation strategies.	*Executive Summary (p.3)*
Actions recommended for sustainability – commercially owned cultural content - There are well-established preservation and access strategies undergoing fundamental changes as a result of new information technologies. This includes the creation, distribution, and consumption of cultural content, most evident in the emergence of interactive genres such as games and the creation of a long tail of use and reuse. As a result, there may be two forms of benefits— commercial and cultural, or private and public—that compete with one another. When that occurs, proxy organizations must step in to represent the public interest. Leading players in preserving this content include private creators, owners, and trade associations, stewardship organizations, regulatory authorities, and leading national and international institutions that can sponsor public-private partnerships to ensure the long-term access to our digital cultural heritage.	*Executive Summary (p.4)*
Actions recommended for sustainability – collectively created Web content - The Web environment is marked by great dynamism, uncertainty about long-term value of digital content, and obscure ownership and rights issues for many collectively produced Web assets. The priority here is for stewardship organizations, content creators, hosting sites, platform providers, and users to model and test preservation strategies, and to provide clarification about long-term value and selection criteria.	*Executive Summary (p.4)*

Engaging context experts to identify context-specific attributes is crucial for developing a sound preservation strategy.* **features specific to a data type or user community that constrain choices for preservation.*	*2.2.1 (p.30-31)*
Scholarly discourse – recommendation 1- libraries, scholars, and professional societies should develop selection criteria for emerging digital genres in scholarly discourse, and prototype preservation and access strategies to support them.	*4.1.1 (p.50-52)*
Scholarly discourse – recommendation 2-publishers reserving the right to preserve should partner with third-party archives libraries to ensure long-term preservation.	*4.1.2 (p.52-53)*
Scholarly discourse – recommendation 3-Scholars should consider granting nonexclusive rights to publish and preserve, to enable decentralised and distributed preservation of emerging scholarly discourse.	*4.1.2 (p.52-53)*
Scholarly discourse – recommendation 4 – Libraries should create a mechanism to organize and clarify their governance issues and responsibilities to preserve monographs and emerging scholarly discourse along lines similar to those for e-journals.	*4.1.3 (p.53-54)*
Scholarly discourse – recommendation 5 – All open-access strategies that assume the persistence of information over time must consider provisions for the funding of preservation.	*4.1.3 (p.54-55)*
Research data – recommendation 1- Each domain, through professional societies or other consensus-making bodies, should set priorities for data selection, level of curation, and length of retention.	*4.2.1 (p.56-58)*
Research data - recommendation 2- Funders should impose preservation mandates, when appropriate. When mandates are imposed, funders should also specify selection criteria, funds to be used, and responsible organizations to provide archiving.	*4.2.2 (p.58-59)*
Research data - recommendation 3 - Funding agencies should explicitly recognize "data under stewardship" as a core indicator of scientific effort and include this information in standard reporting mechanisms.	*4.2.2 (p. 58-59)*
Research data - recommendation 4 - Preservation services should reduce curation and archiving costs by leveraging economies of scale when possible.	*4.2.3 (p.59)*

Research data - recommendation 5 - Agreements with third-party archives should stipulate processes, outcomes, retention periods, and handoff triggers.	*4.2.3 (p.59)*
Commercially owned cultural content - recommendation 1 - Leading cultural organizations should convene expert communities to address the selection and preservation needs of commercially owned cultural content and digital orphans.	*4.3.1 (p.60-61)*
Commercially owned cultural content - recommendation 2 - Regulatory authorities should bring current requirements for mandatory copyright deposit into harmony with the demands of digital preservation and access.	*4.3.2 (p.62-63)*
Commercially owned cultural content - recommendation 3 - Regulatory authorities should provide financial and other incentives to preserve privately held cultural content in the public interest.	*4.3.2 (p.63-64)*
Commercially owned cultural content - recommendation 4 - Leading stewardship organizations should model and test mechanisms to ensure flexible long-term public-private partnerships that foster cooperative preservation of privately held materials in the public interest.	*4.3.3 (p.64-65)*
Collectively produced web content- recommendation 1 - Leading stewardship organizations should convene stakeholders and experts to address the selection and preservation needs of collectively produced Web content.	*4.4.1 (p.66-68)*
Collectively produced web content- recommendation 2 - Creators, contributors, and host sites could lower barriers to third-party archiving by using a default license to grant nonexclusive rights for archiving.	*4.4.2 (p.68)*
Collectively produced web content- recommendation 3 - Regulatory authorities should create incentives, such as preservation subsidies, for host sites to preserve their own content or seek third party archives as preservation partners.	*4.4.2 (p.68-69)*
Collectively produced web content- recommendation 4 - Regulatory authorities should take expeditious action to reform legislation to grant authority to stewardship institutions to preserve at-risk Web content.	*4.4.2 (p.69)*
Collectively produced web content- recommendation 5 - Leading stewardship organizations should develop partnerships with one or more major content providers to explore the technical, legal, and financial dimensions of long-term preservation.	*4.4.3 (p.70)*

Action agenda for Scholarly Discourse – *1. Libraries, scholars, and professional societies should develop selection criteria for emerging genres in scholarly discourse, and prototype preservation and access strategies to support them.* *2. Publishers reserving the right to preserve should partner with third-party archives or libraries to ensure long-term preservation.* *3. Scholars should consider granting nonexclusive rights to publish and preserve, to enable decentralized and distributed preservation of emerging scholarly discourse.* *4. Libraries should create a mechanism to organize and clarify their governance issues and responsibilities to preserve monographs and emerging scholarly discourse along lines similar to those for e-journals.* *5. All open-access strategies that assume the persistence of information over time must consider provisions for the funding of preservation.*	*5.3 (p.84)*
Action agenda for Research Data – *1. Each domain, through professional societies or other consensus-making bodies, should set priorities for data selection, level of curation, and length of retention.* *2. Funders should impose preservation mandates, when appropriate. When mandates are imposed, funders should also specify selection criteria, funds to be used, and responsible organizations to provide archiving.* *3. Funding agencies should explicitly recognize "data under stewardship" as a core indicator of scientific effort and include this information in standard reporting mechanisms.* *4. Preservation services should reduce curation and archiving costs by leveraging economies of scale when possible.* *5. Agreements with third-party archives should stipulate processes, outcomes, retention periods, and handoff triggers.*	*5.3 (p.84)*
Action agenda for Commercially owned cultural content – *1. Leading cultural organizations should convene expert communities to address the selection and preservation needs of commercially owned cultural content and digital orphans.*	*5.3 (p.84)*

2. Regulatory authorities should bring current requirements for mandatory copyright deposit into harmony with the demands of digital preservation and access. *3. Regulatory authorities should provide financial and other incentives to preserve privately held cultural content in the public interest.* *4. Leading stewardship organizations should model and test mechanisms to ensure flexible long-term public-private partnerships that foster cooperative preservation of privately held materials in the public interest.*	
Action agenda for collectively produced web content – *1. Leading stewardship organizations should convene stakeholders and experts to address the selection and preservation needs of collectively produced Web content.* *2. Creators, contributors, and host sites could lower barriers to third-party archiving by using a default license to grant nonexclusive rights for archiving.* *3. Regulatory authorities should create incentives, such as preservation subsidies, for host sites to preserve their own content or seek third-party archives as preservation partners.* *4. Regulatory authorities should take expeditious action to reform legislation to grant authority to stewardship institutions to preserve at-risk Web content.* *5. Leading stewardship organizations should develop partnerships with one or more major content providers to explore the technical, legal, and financial dimensions of long-term preservation.*	*5.3 (p.85)*

4.k.3.3.3 Version control/change procedures

Recommendation	*Found in*
N.A.	

4.k.3.4 Dissemination

Recommendation	Found in
Strategies for overcoming problems with demand for preserved digital materials including use or formulation of proxy organisations; benefits for	*3.1.1& 3.1.2 (p.36-38);*

future users & avoiding irreversible loss.	*Appendix 5 (p.96-97)*

4.k.3.4.1 Visibility

Recommendation	Found in
N.A.	

4.k.3.4.2 Availability and accessibility

Recommendation	Found in
Several choices need to be made by decision makers when designing suitable preservation strategies including: - *Who benefits from use of the preserved asset?* - *Who selects what to preserve?* - *Who owns the asset?* - *Who preserves the asset?* - *Who pays?*	*2.2.2 (p.32-34)*
Because a significant portion of digital assets with long-term value are privately created and owned, strong and robust mechanisms for partnering between public and private sectors are necessary to ensure the long-term public interest in private assets.	*3.2.2 (p.41)*
Organizations can help dilute the risk of free riding. A commonly used mechanism for avoiding free-rider consumption is to make access to or use of the goods exclusive to a group, usually a fee-paying group. Tiered access is a standard remedy for leveling the costs of access to institutions with different financial means without prejudicing access by the community at large. Excluding access obviously does not work for those assets that communities wish to share freely.	*3.2.4 (p.44)*
Condition for sustainability: timely actions to ensure access *Principle of action: Preservation requires a series of decisions to be made over the lifecycle of digital assets. Action: Take preservation steps early in the digital lifecycle; create and codify contingency plans; make and implement plans for handoffs to address economic risks over the digital lifecycle.*	*5.1 (p.74-75)*

4.k.3.4.3 Tools and interfaces

Recommendation	*Found in*
N.A.	

4.k.3.4.4 Monitoring, review and feedback

Recommendation	Found in
Regular review of preservation priorities is important.	*5.1 (p. 75)*

4.k.4 Other information

Recommendation/Comment	Found in
"Sustainable economics for digital preservation is not just about finding more funds. It is about building an economic activity firmly rooted in a compelling value proposition, clear incentives to act, and well-defined preservation roles and responsibilities.	*Executive Summary (p.7)*

4.k.5 Links to further resources (Found in: Chapter 1.0, page 9)

Leadership organisations have engaged a broad spectrum of stakeholder communities to address digital preservation challenges from organisational, policy and technical perspectives. For example:

Ensuring the Integrity, Accessibility, and Stewardship of Research Data in the Digital Age by the

National Academy of Sciences, the National Academy of Engineering, and the Institute of Medicine of the National Academies. Available at http://www.nap.edu/catalog.php?record_id=12615

Preserving Our Digital Heritage: The National Digital Information Infrastructure and Preservation Program 2010 Report. A Collaborative Initiative of the Library of Congress, forthcoming from the Library of Congress. More information on the Library's digital preservation efforts can be found at http://www.digitalpreservation.gov/

The Digital Dilemma, by The Science and Technology Council of the Academy of Motion Picture Arts and Science. Available at http://www.oscars.org/sciencetechnology/council/projects/digitaldilemma/index.html

Sustaining the Digital Investment: Issues and Challenges of Economically Sustainable Digital Preservation (http://brtf.sdsc.edu/biblio/BRTF_Interim_Report.pdf)