Project no.507618

DELOS

A Network of Excellence on Digital Libraries

Instrument: Network of Excellence

Thematic Priority: IST-2002-2.3.1.12
Technology-enhanced Learning and Access to Cultural Heritage

<DELIVERABLE WP6,D6.1.1
Framework for Testbed for digital preservation experiments

Due date of deliverable: 30 November 2004
Actual submission date: 11 November 2004

Start Date of Project: 01 January 2004
Duration: 48 Months

Organisation Name of Lead Contractor for this Deliverable
National Archives of the Netherlands (Hans Hofman and Remco Verdegem)

Revision [Version 1.2]

Project co-funded by the European Commission within the Sixth Framework Programme (2002-2006)

Dissemination Level: PU

Digital Preservation Testbed: Research Framework

Date : 15 November 2004
Version : final

Table of Contents

EXECUTIVE SUMMARY

This research framework provides a basis for establishing a Digital Preservation Testbed Environment and potential experiments for identifying appropriate preservation strategies and methods, defines the scope of the project, expresses the research questions, delineates the experiment process, describes the Testbed environment, and presents the products of the research. This framework may be used in different contexts, such as libraries, archives, audio-visual, scientific data, multimedia, and should be customised and adapted for each research project. It will reflect relevant developments and findings from both the Testbed and external preservation research and initiatives. A Digital Preservation Testbed Environment should establish collaborative relationships with other digital preservation research projects, arrange regular reviews by external experts, and co-ordinate the active participation of relevant partners, such as publishers or Government agencies to insure that the results are relevant, appropriate to the research questions, and supportable by accepted empirical standards.

The objectives of a testbed research project, which will be further addressed by the research questions in section 2.3, are to provide insights into issues such as:

- authenticity features of digital objects
- management processes and activities required to capture, generate and maintain metadata that support the ingestion and preservation of, and long-term access to authentic digital objects
- cost factors for storing, preserving, and managing digital objects and associated metadata
- technical solutions for the preservation of authentic digital objects
- the effectiveness of current and potential preservation approaches.

The scope of a testbed research environment can be defined by:

- **Object Types**: e.g. text documents, audio, video, e-mail messages, multimedia, spreadsheets, databases, data sets
- **Preservation Approaches**: e.g. migration, emulation
- **Preservation Metadata**: defining a required subset through experimentation
- **Digital object Attributes**: content, context, structure, appearance and behaviour.

The digital laboratory, a hardware and software infrastructure, that is at the centre of a Testbed environment consists of supporting technology and two databases. A digital preservation Testbed should produce experiment documentation and results and research products. The Experiments Database, the core of a Testbed, will contain all of the experiment process documentation and experiment results of the Testbed experiments. The research database, the context for the experiment database, will contain the research products, including all supporting research and analysis.

Experiment results could include prototypes, specialised software, and digital objects that demonstrate the success or failure of a preservation approach in meeting specified requirements. Ongoing analysis and a comprehensive final review and analysis will generate research products, such as recommendations on preservation

procedures, policies and practices, cost models, and evaluations of implementation options for preservation approaches.

1. OVERVIEW

1.1 Introduction

1.1.1 Objectives of this report

This report is a deliverable within the preservation cluster of the DELOS project (2004-2006). It relates to task 1 of WP6 and has the following objectives:

- Establish a framework of a digital preservation testbed environment.
- Produce metrics for testing and validating digital preservation strategies.
- Establish mechanisms for ensuring comparability between testbed environment including a testbed test data set (which might include programmes as well as data).

In the following chapters a framework for developing and managing testbeds for digital preservation is provided (first objective), including an introduction to measurement of validity of preservation methods through metrics. This framework will also be the basis enabling comparison between testbed environments.

The origin of the testbed framework is to be found in the Dutch programme 'Digital Longevity' that was initiated in 1996 by the Ministry of the Interior, the Ministry of Education, Culture and Science, and the National Archives. The objectives of this programme were to develop appropriate strategies for managing digital records within and across government in order to ensure their preservation and access over time. This lead among other activities in 1999 to a study carried out by RAND Europe to identify and define a strategy and framework for the long-term management and preservation of digital records in government. One of the recommendations was the establishment of a testbed environment that would enable experimental and practical research to test available preservation strategies. This 'Testbed Digital Bewaring' functioned from 2001-2003 and produced several reports with practical advice on preserving simple document types and databases.[1]

1.1.2 A short overview of a testbed environment

The objectives of a testbed environment are to test and compare different preservation strategies and to provide insights in:

- authenticity features of digital objects
- management processes and activities required to capture, generate and maintain metadata that support the ingestion of, preservation of, and access to authentic digital objects
- cost factors for storing, preserving, and managing digital objects and associated metadata
- technical solutions for the preservation of authentic digital objects
- the effectiveness of current and potential preservation approaches

This Testbed environment framework will enable and monitor relevant developments from other testbeds or research projects, thereby avoiding duplication while

[1] Jeff Rothenberg and Tora Bikson (RAND Europe) , 'Carrying Authentic, Understandable and Usable Digital records Thorugh Time', The Hague 1999. Website: www.digitaleduurzaamheid.nl under Testbed (also in English).

maximising resources. As a collaborative partnership, it will facilitate sharing results and findings through an active communication program. Section 2.1 of the framework discusses relevant research projects and section 3 covers the scope of the Testbed research.

It will be necessary to carefully identify the targets for experiments. The Testbed research questions, continually supplemented by input from creating organisations and an Evaluation Group, should provide the basis for focus of the experiments and direction of the research. In addition, the experiment preparation stage (see section 6.2), in particular, and the research process as a whole should provide a comprehensive and current assessment of the state of the art for each preservation approach. Section 2.3 defines the research questions and discusses the central role of the research questions in the experiment and research processes.

A Testbed should enable construction and use of a digital laboratory environment to support digital preservation experiments that will investigate the best alternatives for preserving authentic digital objects. The Testbed environment should provide not only the infrastructure for conducting the research and the experiments, but will provide the basis for studying long-term maintenance issues for digital repositories. The Testbed itself should not be preserved, but the content of the Testbed must be documented, maintained and utilised over time. Section 4 of the framework discusses the structure and maintenance of the Testbed environment.

The Testbed experiments must address real preservation problems and concerns about digital objects. The experiments may require the support and co-operation from partners, such as government agencies, publishers or others. They will provide an understanding of existing practices and problems, the basis for identifying priorities and concerns, and the objects to be tested in experiments. The Testbed should address the identified issues and concerns, evaluate possible solutions, identify relevant tools and techniques, and recommend implementation strategies. The collaboration with partners is further discussed in sections 1.3 and 2.1, and throughout section 6 of the framework.

The research method for a Testbed defines an explicit and documented experiment process for targeted experiments on digital preservation problems. The experiments should be repeatable and adaptable to future problems and types of objects. The results should be accessible for further analysis and comparison. The main objective of the experiments is to produce practical results and recommendations that should provide the basis for building and implementing preservation solutions, as well as the procedures and policies that will enable the successful preservation of digital objects. Section 6 of the framework describes the full experiment process and context in detail.

The research should provide experiment process documentation, experiment results and research products. Experiment results should include the identification of authenticity requirements for digital objects, a proposed and tested preservation metadata standard, prototypes of preservation functionality, supporting tools and applications that are developed or adapted by the Testbed, and demonstrations of the application of preservation approaches to meet authenticity requirements. Research products should include analyses of empirical data to provide recommendations on identifying appropriate solutions to meet requirements, the effective implementation of

preservation approaches, and cost models for identifying factors and considerations for funding preservation over time. Section 6.3 identifies the experiment process documentation and Section 7 describes the experiment results and research products. The diagram below illustrates the Testbed research process cycle. It identifies the basic stages, the inputs and outputs, and the interaction between the Testbed project environment and external research participants.

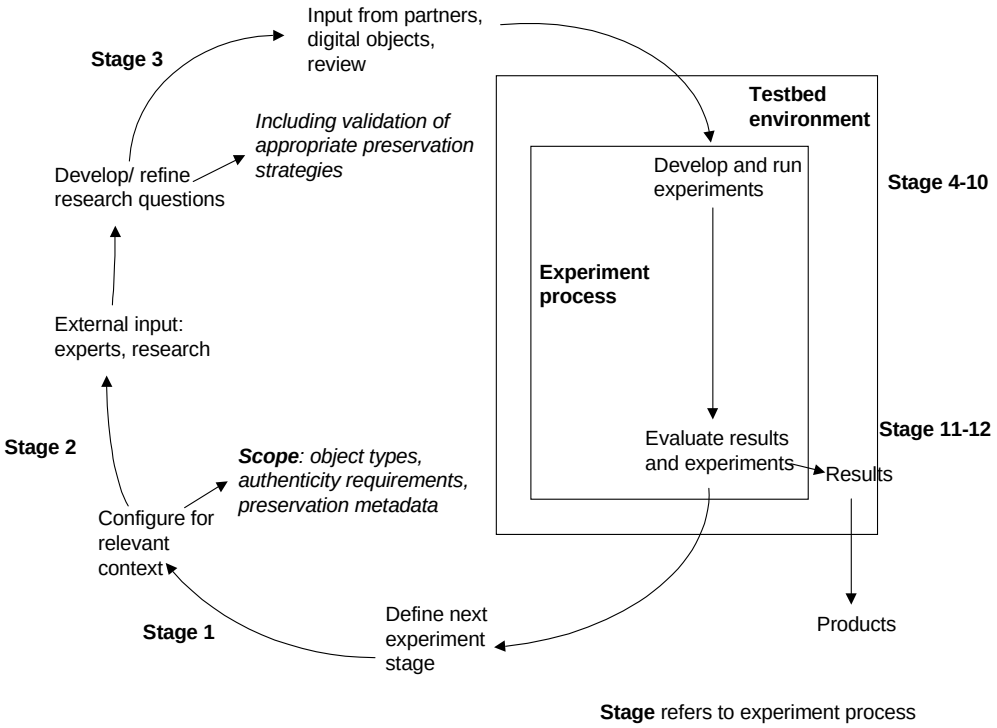


Figure 1 Testbed research process cycle

1.2 The Testbed application environment

The organisational structure of the Testbed consists of Principals, a Feedback Group, interested and experienced representatives from publishers, government, industry and academia (depending on the focus of the Testbed); external experts, from relevant projects and initiatives both national and international; the Testbed Team; and representatives from the identified partners. The Testbed Manager heads the Testbed Team that consists of an Evaluation Group (EG), which will identify the requirements, and run and evaluate experiments; a Research Group (RG), which will develop and maintain the research framework and design the Testbed experiments; and an Implementation Group (IG), which will design and build Testbed environment and develop and test experiments.

The participation of information creating or owning organisations (i.e. partners) in the Testbed is of crucial importance to its success. The Evaluation Group of the Testbed, as discussed below, is leading this area. The partners will have several important roles: to work with the EG to provide information on practices, concerns and

requirements; to identify and prepare selected objects for use by the Testbed (see section 4.3); to participate in running and evaluating experiments; and to serve as external experts whenever possible. The input from partners will contribute substantially to the research and to formulating guidelines and recommendations using Testbed results. Some representatives may also participate as external experts for the project. The Testbed should actively identify a range of external experts to evaluate the experiments designs, the experiment results, and the research products; to participate in designing, developing and evaluating selected experiments; and to contribute to the ongoing identification of technical trends, developments and findings that are significant for the digital preservation topics that the Testbed is addressing.

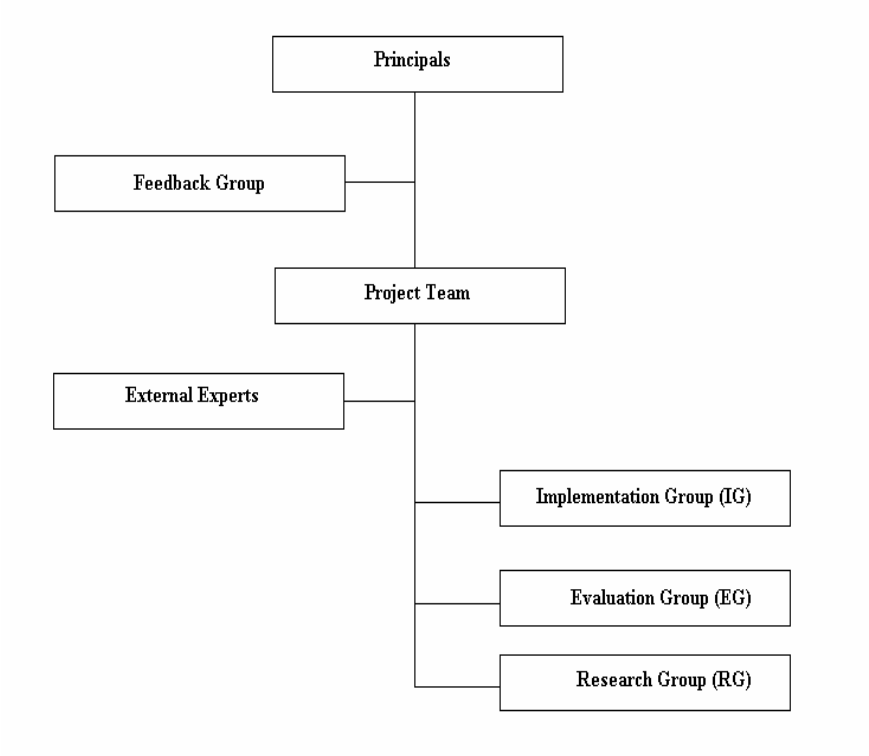


Figure 2 Structure of a Testbed Team

2. CONTEXT FOR THE RESEARCH

2.1 The testbed approach in research

Experimental tools like benchmarking and testbed development have become increasingly important for evaluation in computer science. Testbeds are controlled environments that enable evaluation through the generation of metrics that enable the comparison of systems or strategies (Hanks, Pollack & Cohen, 1993).

Perhaps the most influential models for the use of testbeds in the evaluation of information systems have come from the research domain of information retrieval (IR) and, in particular, from the series of Text REtrieval Conferences known as TREC (http://trec.nist.gov/). This international initiative is co-sponsored by the US National Institute of Standards and Technology (NIST) and the US Department of Defense's Advanced Research and Development Activity (ARDA). The TREC programme was initiated in 1992 in response to the perceived need for collections of test data that could be used to evaluate retrieval algorithms and to compare results across systems. Vorhees (2003) defines a test collection as "an abstraction of an operational retrieval environment that provides the means for researchers to explore the relative benefits of different retrieval strategies in a laboratory setting." In TREC, these collections comprise a set of test documents, a set of information needs (or topics), and relevance judgements. The standard tasks evaluated in TREC are searching for information on a given topic at a particular moment of time (known as ad hoc retrieval), known-item searches, and filtering from changing collections (known as routing). Since TREC-3 (1994), there have been a number of subsidiary 'tracks' that enable focused research into particular research problems or types of object. For example, TREC 2003 (Vorhees, 2003) contained six tracks, including three that were continuations of previous conferences (novelty, question answering and the Web) and three new ones (genomics, high-accuracy retrieval from documents, and robust retrieval). In 2004, a new track will be introduced to investigate ad hoc retrieval in terabyte-scale collections.TREC has inspired the development of a number of other evaluation testbeds in the information retrieval research domain. One specific area that is becoming more important is that of cross-language information retrieval. Growing out of a track first held in TREC-6 (1997), the Cross-Language Evaluation Forum (CLEF) is now an independent project supported by the European Commission (Braschler & Peters, 2004). CLEF is currently a key activity in the evaluation cluster activity of the DELOS Network of Excellence (http://dlib.ionio.gr/wp7/). In a similar way, the NTCIR workshops address the cross-language evaluation of Asian languages (http://research.nii.ac.jp/ntcir/index-en.html).

Testbed development in the digital library research field has largely focused on the need to develop systematic methodologies for the development and evaluation of digital libraries. The development of testbeds was a key component of the two phases of the US Digital Libraries Initiative (DLI). Arms (1998) proposed the creation of large test collections that would be representative of the major categories of material in digital libraries. One early response was the development of the D-Lib Test Suite (http://www.dlib.org/test-suite/research.html), a project supported by the US Defense Advanced Research Projects Agency (DARPA) and co-ordinated by the Coalition for Networked Resource initiatives (CNRI). The Test Suite was developed to address the need for test digital library collections that would be freely available for researchers to

use and which could facilitate comparative research and interoperability experiments (Larsen, 2002). The suite was made of six large test collections covering a wide range of different content types. The different collections covered video segments, computer science technical reports, images, databases, and scanned documents, maps and geographical information, scientific and engineering journal articles, software and numerical databases. The D-Lib Forum also sponsored a Metrics Working Group to consider evaluation issues from the point of view of systems, users and content. The working group defined a framework for the evaluation of the query and retrieval phases of information seeking, including the factors of timeliness, sufficiency, correctness and effort. For each of these, various evaluation metrics were identified and defined (Larsen, 2002).

More recent digital library testbed initiatives include a collection of video content made available by the Open Video project (http://www.open-video.org/project_info.php). The Open Video Digital Library provides an open source testbed of digital video content that can be used for evaluation, the comparison of retrieval approaches, and for use in teaching and learning contexts (Marchionini & Geisler, 2002). More recently, Coleman and Su (2004) have suggested that the US National Science Digital Library (NSDL) could itself become an infrastructure for research into the use of digital libraries in learning contexts.

The DELOS Digital Preservation Testbed has a different purpose to most of these other developments. To adapt Voorhees (2003) definition of TREC test collections, the DELOS testbed is an abstraction of an operational retrieval environment that provides the means for researchers to explore the relative benefits of different preservation strategies in a laboratory setting. It is built upon the experiences of the Dutch digital preservation test-bed, a project that enabled researchers "to ascertain the effects of undertaken preservation action on archival records" (Verdegem & Slats, 2004). While at least one of the D-Lib Test Suite collections was used to test the transformation of varied SGML formats into well-formed XML (Cole, et al., 2001) the Dutch and DELOS test-beds are the first serious attempts to develop a generic framework for the evaluation of digital preservation strategies.

References

Arms, W. Y. (1998). "Replication of results and the need for test suites." Position paper, D-Lib Working Group on Digital Library Metrics, January. Retrieved October 26, 2004 from http://www.dlib.org/metrics/public/PositionPapers/arms.html

Braschler, M., & Peters, C. (2004). "Cross-Language Evaluation Forum: objectives, results, achievements." Information Retrieval, 7, 7-31.

Cole, T. W., Mischo. W. H., Ferrer, R., & Habing, T. G. (2001). "Using XML and XSLT to process and render online journals." Library Hi Tech, 19(3), 210-222.

Coleman, A., & Su, Y. (2004). "The NSDL as a test bed for digital library learning research." In: Project Kaleidoscope volume IV: what works, what matters, what lasts. Retrieved October 26, 2004 from http://www.pkal.org/template2.cfm?c_id=1414

Hanks, S., Pollack, M. E., & Cohen, P. R. (1993). "Benchmarks, test beds, controlled experimentation, and the design of agent architectures." AI Magazine, 14(4), 17-42.

Larsen, R. L. (2002). "The DLib Test Suite and Metrics Working Group: harvesting the experience from the Digital Library Initiative." Fourth DELOS Workshop: Evaluation of Digital Libraries, Budapest, Hungary, 6-7 June. Retrieved October 26, 2004 from http://www.dlib.org/metrics/public/papers/The_Dlib_Test_Suite_and_Metrics.pdf

Marchionini G., & Geisler, G. (2002). "The Open Video Digital Library." D-Lib Magazine, 8(12), December 2002. Retrieved October 26, 2004 from http://www.dlib.org/dlib/december02/marchionini/12marchionini.html

Verdegem, R., & Slats, J. (2004). "Practical experiences of the Dutch digital preservation test-bed." Vine, 34(2), 56-65.

Voorhees, E. M. (2003). "Overview of TREC 2003." In: NIST Special Publication 500-255: the Twelfth Text Retrieval Conference (TREC 2003). Retrieved October 26, 2004 from http://trec.nist.gov/pubs/trec12/papers/OVERVIEW.12.pdf

2.2 Summary of Requirements and Priorities

The Evaluation Group (EG) is the liaison between the Testbed project and the stakeholders, creating and owning organisations, within a community. This will be context sensitive, because different communities will have different requirements. The Evaluation Group will contact the organisations within that community to establish and co-ordinate their participation in the Testbed project, to identify objects that will be used in Testbed experiments, to define the authenticity requirements for the selected objects, and to insure that objects are prepared for Testbed in accordance with established requirements. Working with representatives, the Evaluation Group will identify specific requirements that should be met in preserving the objects that are selected for testing and establish success criteria for Testbed experiments to measure the results of the experiments. The Evaluation Group is preparing the checklist and information package, which will be used for working with information creating organisations.

2.3 Influences

The Testbed should be aware of work going on elsewhere and therefore should actively identify and monitor relevant research projects to insure that the most current information is available to the Testbed, to establish co-operative or collaborative arrangements with other research projects when feasible, to provide information about the Testbed, to seek advice and comment when appropriate, and to provide the full context for the research issues. The Testbed will actively seek other academic, corporate and government programs and initiatives that are addressing relevant digital preservation topics.

2.4 Research questions

There are some fundamental research questions that should be addressed in a Digital Preservation Testbed Environment, such as:

- What are the advantages and disadvantages of implementing each of the specified preservation approaches?

- What are the factors that affect the effectiveness or appropriateness of a particular preservation approach: Cost? Digital object type? Authenticity requirements? Retention period? Preservation management regime? Supporting resources, e.g., staff, equipment?
- How can the effectiveness of a preservation approach be measured and/or demonstrated?
- When, ideally, should the selection of a preservation approach occur?
- What are the basic requirements for each preservation function?
- How are the requirements for preservation functions affected by: preservation approach? Metadata implementation options? Digital object type? Specific authenticity requirements?

Such research questions have to be identified and put into context, so the results will understandable and applicable. In order to be successful in answering these questions it is very important to work out how the Testbed will be working (calibrating the Testbed).

Defining essential preservation metadata will be a priority of a Testbed project. The subset of fundamental research questions that are core metadata research questions include:

- What is the basic set of preservation metadata that is required to insure the preservation of authentic digital objects?
- What factors affect the metadata required for preservation: digital object type? Digital object attribute(s)? Preservation approach? Software? Preservation function requirements?
- What factors affect automatic capture of required preservation metadata: digital object type? Digital object attribute? Software? Preservation approach?
- What are the options for associating metadata with digital objects? How to keep them associated?
- What factors affect when and how metadata are associated and managed: digital object type? Implementation? Preservation approach? Software? Preservation function implementation?

The role and significance of digital object attributes in preservation is another priority research area. The subset of digital object attribute questions include:

- What are the options for preserving each digital object attribute?
- What factors affect the preservation of digital object attributes: digital object type? Software? Preservation approach? Preservation function implementation? Software? Metadata implementation?
- Is it possible, useful or necessary to define retention periods for digital object attributes[2]?
- What is the relationship between the preservation of specific digital object attributes and the cost of preservation?

Each experiment will address one or more of the fundamental research questions as well as research questions that are defined in the experiment design and are specific to an individual experiment or to a category of experiments. For example, research

[2] The preservation of a particular digital object attribute is problematic and/or only required for a specified time period. Preservation strategies have tended to take an all or nothing approach. By assessing preservation effectiveness at the digital object attribute-level, it is possible to consider alternatives, such as identifying methods for documenting that attributes existed at creation and could not be or were not preserved.

questions in the experiment design may address a specific digital object type, digital object attribute, metadata element or category, authenticity requirement, preservation approach, type of software, preservation function, or a specified combination of these.

Examples of some specific research questions are included under sample objects (see section 4.3) and preservation functions (see section 6.4). The research questions are embedded in the experiment process and made explicit in the experiment design. The place of the research questions in the experiments is addressed in the discussion of the experiment process throughout section 6.2.

Any Testbed framework should explicitly define which topics and issues in relation to the types of objects that will be investigated, will be out of scope. These should be included as they are identified.

An example may be: *Storage media*
While storage media management is a core preservation function, there are accepted practices for this function. Many institutions, in particular national archives and universities, have included storage media management programs as a component of preservation management since the late 1960s. These programs include the periodic recopying of files from one medium to another, annual samples to insure media readability, and environmental control of storage areas. While institutions must identify acceptable storage media for preservation programs, the selection and maintenance of storage media for preservation is not a focus of this research. Any incidental observations that pertain to storage media limitations, reliability, or suitability will be documented in the experiment results and research database.

2.5 The framework and how it can be applied

As mentioned previously, digital preservation research so far has not produced hard data on the implementation of available preservation approaches.
A Testbed should be broadly significant for digital preservation, because the research should:

- provide specific results on the requirements for implementing preservation approaches for common object types
- define a metadata set to meet the requirements, such as on authenticity, usability
- analyse these requirements in relation to digital object attributes.

Testbeds will be used in different communities and contexts (libraries, archives, science, audio-visual community). Comparison between the results of those different testbeds will identify differences and commonalities and as such help to understand possible underlying issues.

Though there has been great interest in digital preservation, there is no universally accepted preservation approach. Preservers accepted the concept of migration/conversion as a viable approach without testing the results, or at least without doing so in a manner than can be or has been disseminated. Preservers agree that metadata are the key for preservation, but there is no universally accepted standard nor are there results from trying to implement any of the existing standards. The authenticity of objects is another area where preservers agree that it is necessary, but not on how to insure the preservation of authentic objects. The Testbed will contribute extensively in each of these areas.

Both individual experiment results and the ongoing analysis of the accumulated results should contribute to formulating new or supplementing existing policies, procedures and practices. The Testbed should establish a foundation of information on preservation requirements and approaches, and a laboratory for exploring new problems and requirements as they arise. The efforts of the Evaluation Group should provide needed information on the extent of existing preservation practices and problems.

2.6 Terminology

Many terms are used in different ways within and between professions. For example, digital object, file, object, archiving, and metadata are terms that generally have different meanings for preservers than for information technology professionals or other communities.
There may also be concepts that are specific to the Testbed, such as Stored Object and Experiment Plan. Annex 1 of the Research Framework, *Testbed Terminology: Working Definitions*, contains a initial list of those terms and definitions. The terminology used in the OAIS reference model is the basis for this framework. Terms may be added and definitions refined during the existence of the Testbed as needed.

3. SCOPE OF THE TESTBED RESEARCH

The Testbed should address requirements for preserving authentic digital objects. The possible scope of digital preservation research is very broad. The Testbed experiments should narrow the scope by focusing on the most important and problematic preservation issues that are identified by the Evaluation and Research Groups. Research questions (see section 2.3) address the scope of the research as a whole as well as the scope of individual experiments.

The following sections identify the ways in which the possible scope can been narrowed. For each topic discussed in the following sections, the Testbed should prepare a full discussion document that provides a comprehensive compilation and analysis of current research, existing practices, and relevant issues. This should address an ongoing problem for identifying digital preservation solutions, a fundamental lack of information about the technology, the possible approaches and techniques, and the implications of those implementation options.

3.1 Object Types

In establishing a Testbed environment it is necessary to define the subject and scope of the Testbed experiments, that is the object types that will be experimented with.
The choice of these types will influence the research questions, the technical infrastructure needed for the Testbed, the procedures for testing, the ingest tools, the metadata required for preservation, the method of measuring results (metrics) and probably the preservation strategy to be tested.

3.2 Preservation Approaches

Evaluating the effectiveness of different preservation approaches may be the main objective of Testbed research. Specifically, for each approach, the research should consider the limitations, the potential costs and risks, the appropriate uses, the requirements for the approach, and the resources needed. Additional approaches to investigate may be added or identified in the course of the research.

To insure that the Testbed experiments are pertinent, effective and cost-efficient, it should be important throughout the project to track the results of digital preservation research projects, the progress of digital preservation programs and the development of relevant technology. This tracking should be particularly important for Stage 1, Define Exploration Area, and Stage 2, Prepare for the Experiment, of the Testbed experiment process (see section 6.2). Preservation Approach Trials are discussed in section 6.3.

3.3 Role of metadata

Metadata are needed to capture information about the original format of objects, to document any transformations of the objects that are required or undertaken for preservation or other purposes, and to support the long-term management of and access to the objects. The categories of metadata that are identified in the OAIS provide a basic starting point for the Testbed. There are numerous research projects on metadata standards for (archival) digital objects for instance, but there is currently no single, universally accepted preservation metadata standard. For example, the OAIS standard provides an information model, but not an explicit metadata definition.[3]
The Testbed should also document appropriately all experiments, the results and the framework in which they were carried out. Therefore it should define, capture and rely upon metadata about experiments, requirements, digital objects, the intellectual objects, computer files, and preservation approaches.
While preservation metadata is a primary focus of the research, substantive recommendations on a broader set of possibly required metadata may not be.

The research will determine the metadata needed to document preservation actions and to enable long-term management of objects. It will distinguish metadata that can be captured from metadata that must be generated, and will identify options and mechanisms for associating metadata with digital objects. This research assumes that metadata are required to preserve objects and will identify an essential set of preservation metadata. The OAIS model identifies categories of metadata that are needed for preservation management in a digital repository. The Testbed should consider these categories, as well as existing OAIS-compliant metadata sets, and the source and role of specific metadata elements. The Evaluation Group will map authenticity requirements to a required set of metadata. The Testbed should consider relevant metadata sets and define its own set of metadata that will be refined during the project.

3.4 Attributes of objects

[3] A recent publication by the Planning Committee of the OCLC/RLG Working Group on Preservation Metadata demonstrated the absence of a single standard. It compares three standards that are significantly different from each other but compliant with OAIS (ISO 14721:2003) The metadata research projects are working towards, but have not achieved, convergence.

The research will consider the five quintessential characteristics of digital objects: content, context, structure, appearance, and behaviour. Not all of the attributes may be relevant for all types of documents. These objects have both conceptual and technical characteristics. The role of an object is determined by the context in which it is created and will be used. That (business) context will influence the extent to which certain characteristics of a digital object have to be preserved, e.g. in relation to authenticity.

While these categories of characteristics have been widely accepted by librarians, archivists and records managers, the terms have been neither clearly defined nor consistently used. Based on the results of the experiments, these categories may be revised. Also, new types of digital objects may require other categories of characteristics.

There is general agreement that appearance and behaviour are potentially more difficult to preserve, because these two attributes tend to be more integrally tied to the software that created the digital objects than the other three attributes are. The authenticity requirements for preserving appearance and behaviour may be for example a central focus of research.

The Evaluation Group should develop a set of characteristic-based requirements. The mapping of authenticity requirements and object characteristics, the development of model objects (see section 4.3.2), and the preparation of test objects (see section 4.3.1) will provide opportunities to define the nature of and requirements associated with each characteristic.
An example of how to analyse digital objects may be the use of templates produced in other projects (see section 2.2.4).

The Testbed experiments should consider the options, and the costs and effectiveness of these options, for addressing attribute-based authenticity requirements.

3.5 Attributes of processes

For evaluating preservation processes, the research group considers the following four attributes: Trustworthiness, Stability, Scalability and Usability. All of them can be further detailed as described in Chapter 2 of Deliverable 6.4.1 'Functionality and Behaviour Attributes and Verification Meetings'. [4]

As already mentioned in the previous chapter, one major topic of a preservation process is trustworthiness. It covers two aspects of a preservation solution: First, whether changes are allowed possible and second, if such changes can be traced back, so that the original state of an object can be restored and the process can be repeated. This characteristic comes close to a file characteristic, but since memorising and saving changes in a document is not part of the document itself, the attribute belongs to the process characteristics.

The second attribute, stability, treats a very wide area including the long-term stability of digital objects. It ensures and enables a preservation process to be carried out

[4] Published separately within Delos Workpackage 6.

consistently to the extent that the results are what was expected within a certain framework. Stability furthermore deals with the computer platform (dependency on infrastructure), the acquisition of data (dependency on data acquisition), the storage medium and the storage software, as well as with long-term support of the file format itself.

The third attribute scalability describes, to what extent a preservation process can be extended to a wider or smaller number of files. This criterion is of importance, if the final size, but also if the final range of file formats of a preservation process is not defined.

Finally, usability is the last criterion to describe a preservation process. With usability the handling and satisfaction of the users working throughout and in the process is described. That includes complexity, reaction speed, or the time to user, but also more complex aspects, such as the frequency of migration cycles.

4. TESTBED ENVIRONMENT

The core components of a Testbed environment are supporting technology and two databases operated within an institutional framework. The Experiments Database will contain all of the experiment process documentation and experiment results of the Testbed experiments. The Research Database, the context for the experiment database, should contain the research products, including all supporting research and analysis, the comparative analysis of experiment results, and recommendations and other products that are based upon the results. The hardware and software provide the infrastructure for the Testbed environment. The tools, prototypes, and specialised software that are developed for or used by Testbed experiments will be registered in the Testbed and available for further development or reuse.

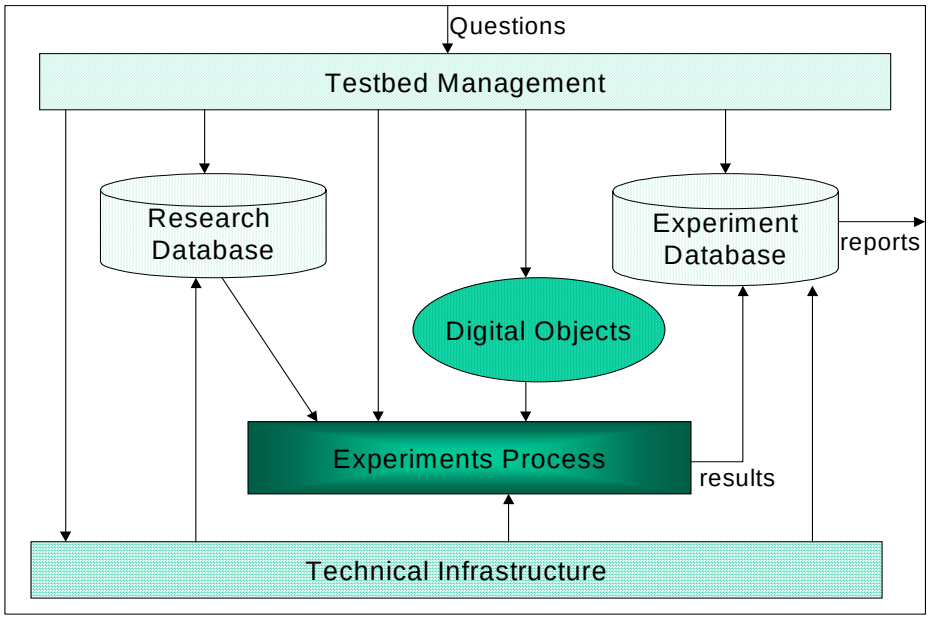


Figure 3 Testbed Environment

4.1 Sample Objects

It is useful to have two categories of sample objects in a Testbed: test objects that are provided by participating organisations as examples of objects out of a real working or business environment, and model objects that are created by the Testbed for experiments. Test objects and model objects will be temporarily stored until the objects are registered. Registration will associate a basic set of metadata with the test and model objects for the purposes of identification and control. The metadata will consist of the metadata defined by the Evaluation Group as well as additional metadata that are specific to the Testbed. Registration will produce Submitted Digital Objects (SDO). Experiments will be run on a subset of one or more SDOs. All SDOs will be stored within a designated area of the experiment database and will be available for use in any experiment. Experiments will generate Tested Digital Objects, a version of an SDO that has been processed by an experiment. TDOs will contain the SDO plus additional preservation metadata as well as versions of the objects, e.g., original formats, transformed formats, or new formats, as dictated by the experiment. TDOs will be the direct result of one experiment.

4.1.1 Test Objects

The Evaluation Group will work with organisations or stakeholders involved to identify representative test objects and to establish co-operative agreements with those organisations to insure that the experiments reflect the real-world preservation problems. The Evaluation Group will work with involved parties to identify the specific authenticity requirements for the selected objects and to prepare the objects for Testbed experiments. The organisational representatives will then participate in the definition, development, running and evaluation of experiments. It will be useful, and in some cases necessary, to analyse the selected objects within the context of their original environment, to analyse the business process that created or used the objects, to identify the authenticity requirements, to consider the capture requirements for the objects, to determine the metadata that are available for capture or must be generated at capture or ingest, and to identify the elements from the original environment that will be needed to enable particular preservation approaches.

4.1.1.1 *Selection and Preparation of Test Objects*

The Evaluation Group will identify objects to be used in Testbed experiments. Regardless of where objects are created and maintained objects must be securely stored and readily accessible. There are three critical issues to consider:

1. the file format for capture, e.g., the original format and the format required for transfer to a repository, e.g., XML or PDF. The Testbed will consider the necessary steps if the objects are not already stored in one of those formats.
2. the requirements that must be met, such as, authenticity requirements, preservation requirements, and requirements that may be specific to a preservation approach, a digital object type, a software package, or a file format

3. the metadata that will be needed to meet the requirements. Is it possible to capture the necessary metadata or will the metadata have to be generated?

The research will further consider the following questions through the preparation of the objects:

- What are the critical steps that must occur at capture to insure the effective preservation of authentic objects?
- Are there steps that should occur at capture to enable a particular preservation approach?
- In what file format are the digital objects stored (original format)?
- What steps are needed to convert the files to XML or PDF?
- Should the original file(s) be retained? Is this affected/determined by the preservation approach being tested?
- What steps should/can an organisation perform to prepare digital objects? What are the options for performing these steps?
- What steps are needed to prepare the files containing the digital objects for the relevant preservation approaches?
- What metadata are present and available for capture in the original environment?
- What required metadata are missing in the original environment?
- How can the missing metadata be captured or generated?
- Is it possible/useful to create a template for capture to provide an active process rather than passive guidelines?

4.1.2 Model Objects

The Research Group and the Implementation Group will work together to create and document a set of model objects with a well described, comprehensive and representative set of characteristics and features. It will serve as the testbed test data set. The Evaluation Group will identify requirements for model objects that are representative of preservation issues. These objects and their characteristics need to be described with adequate metadata. The Testbed needs to establish this metadata set.

The model objects will be used to:

- develop and test experiments
- identify the possible set of components, captured metadata, and attributes of digital documents
- analyse the capabilities and limitations of relevant software packages in the creation of digital documents
- provide a control group for experiments on test objects
- insure the continuity of the research in the event of any delay in obtaining test objects and participation from organisations involved

The creation of model objects could contribute to the development of guidelines on the creation of preservation-compliant objects.

4.1.3 Submitted Digital Objects (SDO)

Model objects and test objects that are stored in the Testbed will be stored as Submitted Digital Objects (SDO) with metadata that are associated with the objects to identify the source, definition and uses of the objects. SDOs

will be accessible to all Testbed experiments. The Experiment plan (see stage 7 of the experiment process in section 6.2) will identify the SDOs that are used in each Testbed experiment. Note: an Experiment Design may be used for multiple experiments and so cannot identify specific objects that are used in an experiment that used the design.

4.1.4 Tested Digital Objects (TDO)

A transformed digital object (TDO) is the representation of a digital object with all of the components that are needed to support and enable its preservation within the context of a certain chosen preservation strategy. These components include the association to the computer files that contain the digital object, the authenticity and preservation requirements that apply to the digital object, the preservation actions that have been carried out on the digital object, all of the metadata that have been defined as required for preserving the digital object, and the means to support the selected preservation approach. Preserved objects will be stored as TDOs.

The research will determine and refine the components of the TDO. The definition of TDOs will be based upon the OAIS information and data models.

The following diagram illustrates the types of Testbed objects and the relationships between them.

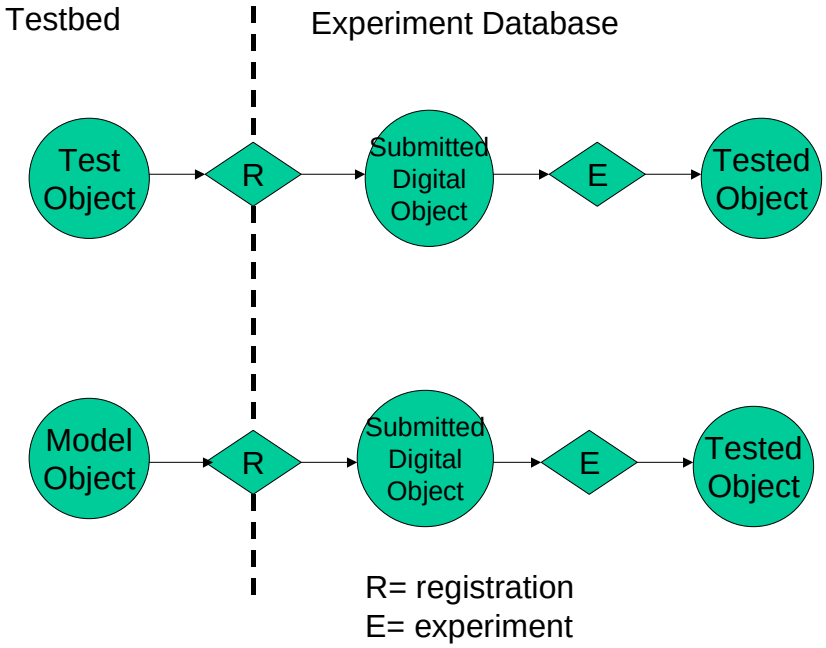


Figure 3 Testbed Objects

4.2 Experiment Database

Each Testbed experiment will be registered in the research database using the experiment checklist, which will track each experiment through all of the experiment

stages. The experiment database will contain all of the products of experiments. The experiment stages are identified in section 6.2, the experiment process documentation in section 6.3, and the experiment results and research products in Section 7.

The Experiment Checklist in the research database will provide access to the content of the experiment database. The experiment database entries, once finalised, will give read-only access to insure that there is an accurate digital object of each experiment.

The products of an experiment will be stored in the experiment database even if there is a decision not proceed at stage 6 or 9 of the experiment process, the Go/No Go Decisions. These products will be available to be reactivated at a later date, redeveloped for future experiments, or for further analysis during or after the project.

4.3 Research Database

The research database will:
- support the ongoing research by documenting research citations and observations, the design and development of experiments, test objects, model objects, and all relevant research information
- register and track experiments to allow for comparison, analysis and reuse
- enable the results and insights from the project to be shared and fully utilised, both during the project and beyond.

The Research Database will contain at least the following object types:
- Research Citations: information on all relevant sources used by the research project
- Research Observations: information about specific experiments, preservation approaches, digital object types, file formats, or any relevant topic that might be useful in advancing the research
- Experiment Registration: capture of the status of experiments at all stages in the experiment process.

And compiled references or white papers on:
- Preservation Approaches, e.g., migration, emulation
- Object Classes, e.g., e-mail messages, simple spreadsheets, multimedia, audio, video, and databases
- Object Attributes, e.g., content, context, structure, appearance, behaviour
- Software Types, e.g., word processing, e-mail, spreadsheet, database, imaging.
- Relevant external preservation projects either existing or finished, e.g, InterPARES, NEDLIB, CAMILeON, Dutch Testbed Digitale Bewaring, Digital Curation Centre (DCC).

4.4 Testbed Management

The Testbed requires an ongoing process to manage the hardware and software that support the Testbed environment. The Testbed infrastructure will provide the opportunity to identify and consider long-term maintenance issues for digital repositories. The Implementation Group will produce a complete set of Testbed

design documents. See the Use Case Document and Software Requirements Document for detailed specifications.

The Implementation Group will establish a full maintenance plan for the Testbed. The software tools that are used by (including those developed by) the project should be registered and available for evaluation and reuse. The Testbed should insure that experiment results are both securely captured and easily accessible. Testbed users will be registered, and access to the Testbed should be controlled. Implementation of the maintenance plan will be monitored and documented.

5. METRICS FOR TESTING

For testing and comparing various preservation strategies, a metric is introduced in this chapter. The metric, which is based on the concept of 'Utility Analysis', can be adapted to the individual requirements of collections and is not limited to any preservation strategies.

The metric follows a process, as illustrated in Figure 5, starting with the definition of project objectives. In this first step an objective tree is created, where many different criteria, concerning the file itself (containing among many others the previously mentioned attributes of content, context, structure, appearance and behaviour, which can be all further detailed), concerning the preservation process and preservation framework and finally the costs, arising through the preservation method. In the second step measurable effects, such as mm, seconds or EURO or the possibility of categorising the criterion, are assigned to the objective.

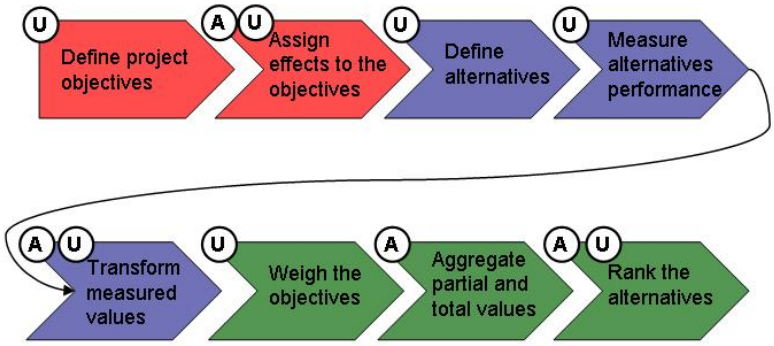


Figure 4 Process of Utility Analysis

In the second block of steps, all various alternatives, which are to be evaluated, are defined by the Research Group. These are then implemented with the help of representative files, as described in Chapter 4.1.1.1. The Evaluation Group then measures the performance of these implementations regarding the previously defined objectives. In case a categorizing measure is attached to a criterion, the Evaluation Group sets the value according to its impression.
In the fifth step, a table is defined, which transforms the heterogeneous values of the outcomes into comparable numbers, which is subsequently done. Having the comparable numbers, the objectives and the branches of the objective tree are weighted according to the user's preferences and requirements. Then the comparable values are aggregated with respect to these weights and summed up to a single value per evaluated preservation strategy. Finally as a last step a sensitivity analysis can be done, stressing the outcomes reaction on minimal weight changes and the alternatives are ranked according to their total values, which shows a clear and rational preference for one or several solutions. These final values not only describe a ranking among the preservation solutions, but also show the degree, to which a solution fulfills the criteria of a perfect preservation strategy for the individual requirements.
The work concerning the evaluation metric and 'Utility Analysis' was done in context with the description of the characteristics of attributes of an object.[5] In the report of

[5] As part of the Delos project, workpackage 6 Preservation, Task 4. Report written by Carl Rauch and Andreas Rauber published separately.

this task a more detailed description of the process and two exemplary implementations, one with text-documents another with audio-files are presented.

6. APPLICATION OF FRAMEWORK

A Testbed experiment will address the authenticity requirements that are defined by the Evaluation Group for a set of objects. An experiment may address one or more of the preservation functions that are discussed below: ingest, storage and access to consider its implications for the preservation process or to consider the implications of that stage for a particular preservation approach. The preservation approach trials are the primary focus of the experiments.

The total number of experiments that are run during the research will be significantly higher than the number of experiment process cycles that are completed during the research. The experiment process produces one experiment design. One experiment design will be used to run one or more experiments using different sets of test objects, modest variations of the experiment design, or alternative sequencing of the designed experiment steps, any of which might produce significant results. The research questions section below and Stage 1 of the experiment process that is described in full in section 6.2.1 addresses the focus and scope of experiments.

6.1 Scenario for applying framework

A scenario for a Testbed experiment process is descried in this section. It consists of twelve possible stages. There are two points (Go/No Go decisions) at which experiments may be halted permanently or temporarily. The products of each stage will be captured and available for analysis, reuse or further development. The following discussion provides a comprehensive description of each stage. A diagram of the full experiment process follows the discussion.

6.1.1 Stages of an experiment

Stage 1 Define Exploration Area

What This is a fairly informal stage to select the area for exploration in accordance with priorities and sequences agreed upon by the Project Team and with the results of previous experiments. The Team will investigate more than one area concurrently to make the best use of time and resources. Initially the area to be explored will be defined first by the digital object type and second by the preservation approach. Later in the research, experiments may address more than one digital object type if the application of a preservation approach can be generalised for all or for multiple digital object types. This stage defines the parameters of the area to be explored, identifies the relevant research questions, and references relevant research, practices, and developments. This stage should also set the basic timeframe for the experiment to make the best use of time and resources.

Why There is a vast scope of possible experiments. It is important to identify the areas that will produce the most significant results.

Who RG with active input from EG and IG and agreement by the PT

When At the start of each experiment

Where Within the Testbed environment using all available resources
How Consider what is desirable and what is feasible
Input Background research and current information on relevant areas compiled by RG. Information gathered by EG from agencies on requirements, concerns, etc. Technical considerations from IG on potential or desirable research targets.
Output Statement on the scope of the experiment (RG)
Then… Proceed to stage 2

Stage 2 Prepare for Experiment

What Once the exploration area is agreed, preparation includes any needed background study of the technical properties of the object type, the preservation approach, the storage formats, i.e. PDF or XML, and any other relevant topics that are identified. This stage will entail further discussion of the research questions to be addressed by the experiment. The information gathered will be captured and accessible within the research database of the Testbed as background for the experiment developed and as supporting data for future experiments. More than one experiment can be prepared concurrently. Preparation may include small demonstrators to explain or consider an aspect of an experiment. These are informal prototypes that are commonly used by software developers rather than the formal preservation process prototypes (PPP) referred to in section 6.5.
Why Each experiment needs to start with a solid base with the most current information on preservation approaches, digital object types, software, etc. The preparation for each experiment will be somewhat different for each experiment but preparation will build on the information gathered for earlier experiments.
Who RG with input from the PT
When After stage 1, once the exploration is agreed
Where Within the Testbed environment using all available resources
How The best sources for the preparation will be websites as the most current information on research initiatives and software and technology developments tend to be made available on the internet; information provided to the Testbed project by relevant digital preservation projects, and documentation and other publications on the relevant topics.
Comparison with the results of other testbed initiatives is another important source of information.
Input Statement of scope from stage 1 and additional active input from PT
Output Analysis and compilation of the background on and preparation for the experiment (RG). [Compilation based on new and existing experiment observations and results, citations, discussion papers, white papers, etc.]
Then… Proceed to stage 3

Stage 3 Define Requirements

What This stage will determine the authenticity requirements for the experiment as defined by the Evaluation Group, i.e. the success

criteria for the experiment upon which the evaluation will be based. These requirements will be based upon legislation or regulations, authenticity requirements identified by the creating organisations, the critical attributes of the type of object that need to be preserved, requirements for future access to the objects, etc. The definition of the requirements for an experiment will consider the relevant research questions and provide recommendations for addressing the questions. Success criteria for the authenticity requirements will be defined. Metrics to evaluate these criteria will be developed.

Why These requirements define the objectives of the experiment. The degree to which application of the preservation approach in the experiment can meet the defined requirements will determine the success of the experiment.

Who EG with input from the PT and agency representatives

When For universal requirements that affect all experiments or all experiments on specific object types or preservation approaches, requirements will be developed from the start of the project and as a result of experiments. For requirements that may be specific to experiments, requirements will be defined once the exploration area is agreed in stage 1 and the preparation completed in stage 2.

Where Within the Testbed environment using agency input

How The Evaluation Group will analyse the information provided, the criteria identified by the creating organisations, consider the pertinent legislative requirements, etc. These requirements may be reconsidered during the experiment design.

Input Statement of scope from stage 1 and compilation and analysis of background and preparation from stage 2.

Output Requirements definition (EG), basic evaluation checklist (EG)

Then… Proceed to stage 4

Stage 4 Develop Experiment Design

What The experiment design will define:

- the purpose of the experiment
- the research questions that will be addressed, both general and specific
- the type of objects that will be the focus of the experiment
- their characteristics properties
- the steps needed to run the experiment
- the procedures for setting up and running the experiment
- the metadata needed to run the experiment and that will result from the experiment
- the requirements of the software to be developed or used
- the manner in which the requirements will be addressed
- the validation to be performed
- the type of results expected (e.g., processed files, metadata).

The first experiment designs may require more preliminary testing than later experiments to identify validation criteria and methods, to identify needed and generated metadata, to establish the capture mechanisms for results, and to establish the sequence of

events for the experiment. An experiment design might be used for multiple experiments using different sets of objects, testing options defined in the design, rerunning experiments for comparative results, etc. (see Annex 3.6 for the Experiment Design template).

Why The experiment design defines the research questions that will be addressed and provides the framework for each experiment, the method for running the experiment, and some of the core documentation for the experiment.

Who RG in collaboration with the IG with active input from and review by the EG and approval by the PM before proceeding to the next stage

When After the requirements are defined (a basic set of requirements will be developed for the first experiments and refined and expanded through successive experiments) in stage 3 and after potential revisions are identified during development (stage 7), testing (stage 8)or running the experiment (stage 10)

Where Within the Testbed environment

How Each experiment design will address key research questions. Some of these will be general questions for all experiments and some will be specific to the object type, the preservation approach, or the experiment. The experiment design will consider the requirements, the critical attributes of the object type, the steps required for the preservation approach. The experiment designs will build upon previous experiments.

Input Statement of scope from stage 1, background and analysis from stage 2, and requirements definition from stage 3.

Output Experiment Design (RG), further refinement of the Evaluation Checklist (EG)

Then… If PM approves, proceed to stage 5. If not revise Experiment Design or possibly return to stage 3 to reconsider and/or refine the requirements.

Stage 5 Specify Resources

What This stage will provide a detailed estimate of the time and resources needed to build, test and perform the experiment. The estimates will include possible options for building, testing and performing the experiments and the implications of each.

Why The recommended preservation approaches and the application of these approaches must consider the feasibility of these decisions over time. The experiments should produce the most significant results and make the best use of time and resources.

Who IG with input from and review by the PT

When Once the PM okays the Experiment Design (stage 4) or after fixes or additional developments are identified during testing (stage 8) or running the experiment (stage 10)

Where Within the Testbed environment

How The IG will analyse the requirements and factor the time needed to complete each stage. Options for minimal to full compliance, particularly for costly and complex experiments, will be laid out. The estimates will be broken down by phases and steps to

provide a basis for the Go/No Go decision.

Input Experiment Design from stage 4

Output Resource specification (IG)

Then… If the IG needs more information or input return to stage 4 to refine the Experiment Design. If the resources are considered to be high, return to stage 3 and/or 4. Otherwise, proceed to stage 6.

Stage 6 Go/No Go Decision

What This stage considers the experiment design and the resource definition to determine if the experiment should proceed as planned (Go), if revisions to the design or the statement are needed before the experiment can proceed (provisional Go), if the experiment should be delayed for a specified period or until a specified event occurs (deferred Go) or if the experiment should be terminated at his stage (No Go). If the decision is No Go, the PM/PT will prepare a full justification for the decision as part of the documentation for the experiments. The end result of this stage will be an approved experiment design (once any revisions are completed) or a full justification not to proceed. The final design will include a proposed schedule for the experiment.

Why An experiment may be too costly in time or resources either as an experiment or as an ongoing practice. The results may be interesting but the cost may be too high to undertake the experiment. Documenting this decision will be important for the results of the Testbed project and for future experiments, policies, and decisions.

Who PM with input from the PT as requested, possibly external advice

When After the experiment design has been completed (stage 4) and the time and resources estimated (stage 5) and after revisions or additional developments are identified while testing (stage 8) or running the experiment (stage 10)

Where Within the Testbed environment

How Consider the cost and significance of the experiment within the overall context and the parameters of the budget for time and resources.

Input Experiment Design from stage 4, Resource Specification from stage 5 and other input as requested

Output Declaration of decision and justification when needed (PM)

Then… If Go, proceed to stage 7. If provisional Go, return to stage 4 and redo stages as needed. If No Go, capture decision and justification, and retain documents from previous stages for possible future use.

Stage 7 Develop Experiment

What This stage will produce a specific development plan for each experiment that will be run using the experiment design, as well as a test plan and experiment plan, any specialised software, procedures and other experiment components needed to prepare the Testbed for the experiment. (The plans and all other experiment documentation are defined and discussed in section

6.2) All of the items needed for the experiment, as defined in the design, will be developed and/or installed and tested, including copies of all the objects needed for the experiment, software packages and programs needed, and mechanisms for capturing the results and the evaluation.

Why The development of the experiment will determine the success of the experiment, the nature of the results, the completeness of the documentation, the ease of running and demonstrating the experiment, and the potential reusability of the experiment modules.

Who IG with input and support from the RG and EG, for procedures, explanations, additional input, etc. Review and approval by PM.

When After the Go/No GO decision (stage 6) and after revisions or additional developments are identified while testing (stage 8) or running the experiment (stage 10). Development will be done using the agreed schedule

Where Within the Testbed environment.

How The Implementation Group will use standard methodologies to develop the software for the experiments.

Input Experiment Design, stage 4 and Resource Specification, stage 5

Output Development plan (IG), test plan (IG) and experiment plan (IG).

Then… If approved, proceed to stage 8.

Stage 8 Test Experiment

What This phase will systematically test running experiments and review the captured results. The tests may identify aspects of the experiment that need to be corrected or additional development that is required. The tests will generally use model objects to test experiments, though test objects provided by the involved organisations may also be used.

Why The experiments will be run by or for representatives of organisations that are participating. It is possible that running the experiments may identify problems with the design or development of the experiment, but, to the extent possible, the experiments should be fully functional and ready to run.

Who EA with input from the EG

When After the experiment development has been completed (stage 7) and after any additional fixes/developments identified during testing and running (stage 10)

Where Within the Testbed environment

How The experiments will be tested in accordance with defined test plans. The test plans and results (including metrics about them) will be captured as part of the documentation of the experiments.

Input Test Plan with reference to Development Plan and Experiment Plan as well as input from EG and organisational representatives

Output Test results (IG), final evaluation checklist (EG)

Then… If approved by EG (and organisational representatives), proceed to 10. If not, return to stage 4 or stage 7, as appropriate, and redo stages as needed.

Stage 9 Go/No Go Decision

What This stage entails an analysis of the development plan, experiment plan and test results to determine if the experiment should proceed as planned (Go), if revisions to the design or the statement are needed before the experiment can proceed (provisional Go), if the experiment should be delayed for a specified period or until a specified event occurs (deferred Go) or if the experiment should be terminated at this stage (No Go). If the decision is No Go, the PM/PT will prepare a full justification for the decision as part of the documentation for the experiments. The end result of this stage will be an experiment that is ready to run or a full justification not to proceed.

Why The test may indicate that the experiment will not provide the expected or desired results, may take too long to prepare and run, or some logistical consideration may cause a delay or termination of the experiment. Documenting this decision will be important for the results of the Testbed and for future experiments, policies, and decisions.

Who PM with input from the PT as requested

When After the experiment development has been completed (stage 7) and the experiment has been tested (stage 8) and after revisions or additional developments are identified while testing (stage 8) or running the experiment (stage 10)

Where Within the Testbed environment

How Consider the cost and significance of the experiment within the overall context and the parameters of the budget for time and resources.

Input Development plan, test plan and results, experiment plan, and other input as requested

Output Declaration of decision and justification when needed

Then… If Go, proceed to stage 10. If provisional Go, return to stage 4 or 7 per PM decision and redo stages as needed. If No Go, capture decision and justification, and retain documents from previous stages for possible future use.

Stage 10 Run Experiment

What An experiment will test one or more aspects of applying a preservation approach to a defined set of objects. Running an experiment will produce results, e.g. converted computer files, revised metadata, etc., that will be evaluated in stage 11. More than one experiment may be run using an experiment design on different sets of objects.

Why The experiments and the results generated by the experiments are the core of the research.

Who EG with partner organisations and support from the PT as needed

When After the experiments have been tested (stage 8) and approved (stage 9); possibly after the experiments have been revised or enhanced (stage 7)

Where Within the Testbed environment

How The EG will run the experiments following the steps identified in

the experiment plans defined during the experiment development. The EG will consider the number of experiments to run based on sets of objects that fit the scope of the experiment design. Any problems encountered or observations made while running the experiments will be captured as part of the experiment documentation. Running the experiments may identify fixes or further developments.

Input Experiment plan (possibly development plan and test plan and results) plus all materials needed to run experiment

Output Experiment results (automatically captured or observed by EG and agency representatives).

Then… The EG may recommend rerunning the experiment to check results or proceed to stage 11.

Stage 11 Evaluate Experiment

What The results of an experiment will be evaluated to determine how successfully the requirements were met, to consider if more experiments are needed on more sets of objects, and to consider what secondary results can be developed from the results of one or more experiments, such as proposed guidelines, policies, procedures, discussion papers, etc. The evaluation will be done with help of metrics developed in stage 3. An evaluation could compare the results of one or more experiments.

Why All of the possible and desirable experiments cannot be developed and run. It is important to make the most effective use of the results and the resources, and to carefully consider the experiments that are needed.

Who EG with input from the PT

When After an experiment is run (stage 10) and after several experiments with comparable results have been completed.

Where Within the Testbed environment.

How The results of the experiments will be evaluated against the defined requirements.

Input Experiment results

Output Completed evaluation checklist (EG and organisational representatives)

Then… Based on determination by EG (and organisational representatives) rerun experiment, recommend revisions at stages 4 or 7, and/or proceed to stage 12.

Stage 12 Consider Results

What All of the stages of the experiment(s) will be considered to make recommendations for the refinement and enhancement of future experiments, to propose further experiments, and to provide input into the evaluation of the Testbed.

Why The evaluation in stage 11 should focus on specific experiments and form part of the experiment results. The review in stage 12 should consider the broader context of the results of one or more experiments for the development of the Testbed.

Who PT

When After one or more experiments have been run (stage 10) and

evaluated (stage 11).

Where Within the Testbed environment.

How The PT will look at cumulative experiment results, consider the range of topics that the Testbed needs to address, and review the experiment process and the Testbed as a whole.

Input All experiment products from stages 1-11 and additional input from whole PT

Output Recommendations on further experiments, additional analysis of one or more experiments, Testbed evaluations, white papers when appropriate

Then… Bring input to stage 1

6.1.2 Experiment Process Considerations

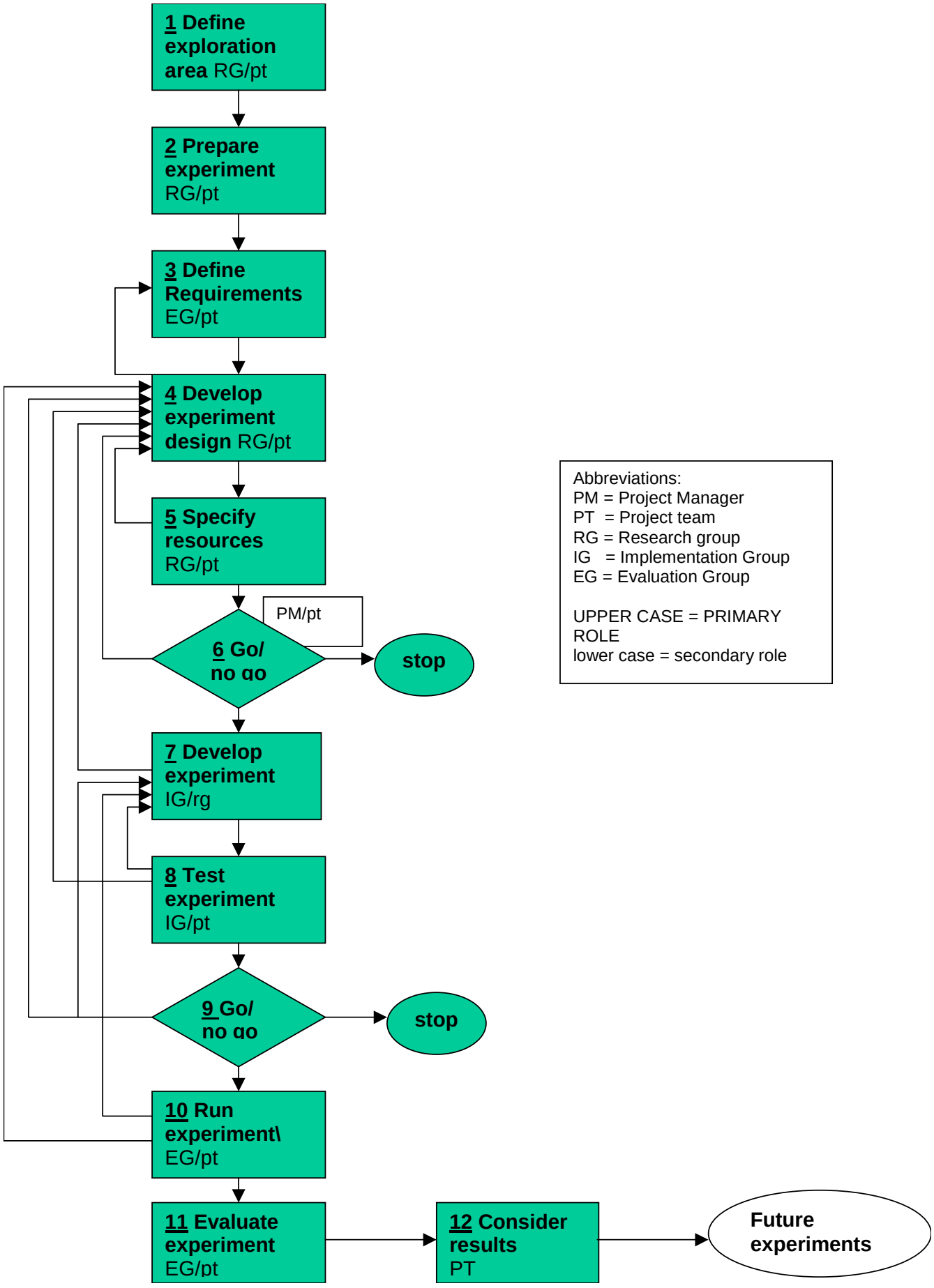
Experiment designs and results should be evaluated by a panel of experts. Establishing the panel and considering the best means for conducting the evaluation may take some time. Perhaps the best way to proceed is to develop and run two or more experiments to provide a basis for the evaluation and to avoid delaying the start of experiments.

The experiments, particularly the specification of resources in stage 5, will produce input for the cost models.

Evaluation of the Testbed is to some extent covered in stage 12 but is not really within the scope of the experiments. The evaluation of the experiments will provide input for the evaluation of the Testbed. External advisors should provide an evaluation of the Testbed as well.

The following diagram illustrates the stages of the experiment process and the points at which the process loops back to a previous stage for revision and reconsideration.

Figure 5 Digital Preservation Testbed Experiment Scenario



6.2 Experiment Process Documentation

Each stage of the experiment will produce or be captured by an explicit piece or set of documentation. All of the documentation will be captured in the experiment database.

6.2.1 Statement of Experiment Scope

The Research Group has primary responsibility for the Statement of Experiment Scope that is produced in stage 1 of the experiment process. This document makes an explicit statement about the intended scope of the experiment, making references to other experiments as appropriate and providing a concise description of the purpose and specific focus. See Annex 2.2.

6.2.2 Analysis/compilation of background and preparation for the experiment

The Research Group has primary responsibility for this analysis that is produced in stage 2 of the experiment process. This document highlights the most relevant references and the information that pertains to the scope of the experiment. It also compiles pertinent results, observations, and other information that will serve as substantive input for the experiment design. See Annex 2.3.

6.2.3 Requirements Definition

The Evaluation Group has primary responsibility for the Requirements Definition that is produced in stage 3 of the experiment process. This document defines the authenticity requirements identified by the organisation involved and additional requirements identified by the Evaluation Group as appropriate. See Annex 2.4.

6.2.4 Basic Evaluation Checklist

The Evaluation Group has primary responsibility for Basic Evaluation Checklist that is produced in stage 3 of the experiment process. This document will match the organisation-identified requirements and other requirements to measures of experiment results to produce a checklist that will form the basis for evaluating experiments. It will be refined during the experiment process and finalised prior to stage 10, run experiment. See Annex 2.5.

6.2.5 Experiment Design

The Research Group has primary responsibility for Experiment Design that is produced in stage 4 of the experiment process. This document will define: the purpose of the experiment; the digital object types that will be the focus of the experiment; the steps needed to run the experiment; the procedures for setting up and running the experiment; the metadata needed to run the experiment; the metadata that will result from the experiment; the requirements of the software to be developed for or used in the experiment; the ways in which the requirements will be addressed by the experiment; the validation to be performed on the results; and the kind of results expected, e.g., converted computer files, metadata. An Experiment Design might be used for multiple experiments using different sets of objects, testing options

defined in the design, rerunning experiments for comparative results, etc. See Annex 2.6.

6.2.6 Resource Specification

The Implementation Group has primary responsibility for Resource Specification that is produced in stage 5 of the experiment process. This document will identify the time and resources that will be needed to develop, test and run the experiment. The Implementation Group will note any contingencies that should be addressed, such as the purchase of software or other materials for the experiment(s) or factors that will affect the sequence in which experiments need to be developed. See Annex 2.7.

6.2.7 Declaration of Go/No Go decision

The Project Manager has primary responsibility for Declaration of a Go/No Go decision, and the associated justification, that is produced in stages 6 and 9 of the experiment process. This document might be very brief, but explanations of No Go decisions that might be used to continue the experiment process in the future, or of provisional Go decisions, will be captured. See Annex 2.8.

6.2.8 Test Plan

The Implementation Group has primary responsibility for the Test Plan that is produced in stage 7 of the experiment process. This document will provide an explicit plan for testing experiments and define mechanisms for testing or capturing, as appropriate. Stage 8, Test Experiment, will produce test results based on the Test Plan. See Annex 2.9.

6.2.9 Development Plan

The Implementation Group has primary responsibility for the Development Plan that is produced in stage 7 of the experiment process. The experiment development includes all of the steps and activities needed to implement the experiment design. The Implementation Group will define a development plan for experiments that will be developed using each Experiment Design. The plan will include an indication of steps needed to build and test the specified software, a definition of the procedures required to run the experiment, and the provisions needed to prepare the Testbed environment for the experiment. See Annex 2.10.

6.2.10 Experiment Plan

The Implementation Group has primary responsibility for the Experiment Plan that is produced in stage 7 of the experiment process. This document defines the sequence of activities required to run an experiment and serves as a checklist for running a specific experiment. The content of the plan is outlined in the experiment design and completed during the experiment development. The plan will include any activities required to capture the experiment results of an experiment, particularly when the capture is automated. There will be one Experiment Plan for each experiment, so possibly more than one Experiment Plan for each Experiment Design. See Annex 2.11.

6.2.11 Final Evaluation Checklist

The Evaluation Group has primary responsibility for the Final Evaluation Checklist that is finalised by stage 8 of the experiment process. See section 6.3.4 on the Basic Evaluation Checklist and Annex 2.12.

6.2.12 Experiment Results

The Evaluation Group has primary responsibility for the Experiment Results that are produced in stage 10 of the experiment process, though some of the results may be captured automatically during the experiment. Each experiment will produce experiment results that are captured and available for evaluation, comparative analysis, and for future experiments. Experiment results will include converted files, captured and generated metadata, experiment log files of the sequence and completion of activities performed in an experiment, and the successful completion of required tasks. The experiment evaluations will weigh the success of an experiment against specified requirements and criteria, and provide the basis for recommendations and guidance based on the experiments as secondary results. See a further discussion of experiment results in section 6.3 and Annex 2.13.

6.2.13 Experiment Evaluation

The Evaluation Group has primary responsibility for the Experiment Evaluation that is produced in stage 11 of the experiment process. The Evaluation Group will conduct the evaluation with agency representatives and input from the other members of the Project Team as appropriate. The evaluation will consider results that were observed, automatically captured during, or produced by running the experiment. The Evaluation Checklist will be the basis for the evaluation but any additional information will be captured to fully document each experiment and to provide input for developing future experiments. See Annex 2.14.

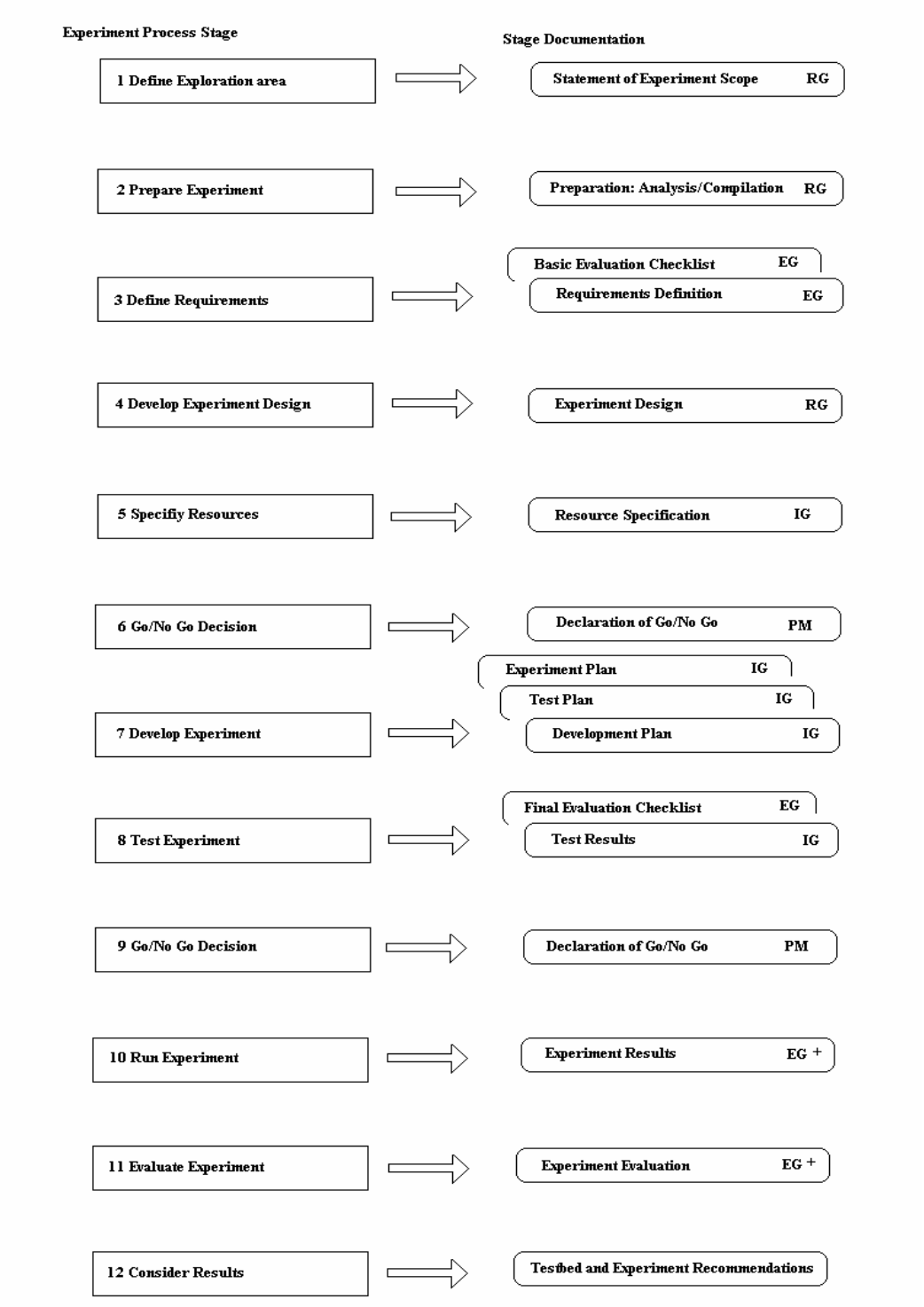


Figure 4 Experiment Process Documentation

6.3 Preservation Approach Trials

The research will develop experiments that apply a specified preservation approach to a set of captured and accepted objects. The evaluation of the trials will assess how well the approach meets the authenticity requirements that are defined by the Evaluation Group. Each trial will be designed, developed, run and evaluated to determine the technical requirements for the approach, the appropriate steps, the resources needed, and the contributing factors to the success or failure of the approach. For example, an approach may work well for some object types but not for others. A preservation approach trial will probably consist of multiple experiments. The cumulative results of a set of experiments would provide the basis for the evaluation of an approach.

What are the factors that determine the success of a preservation approach?
Is a preservation approach more successful in preserving specific object types, software packages, or file formats?
Is a preservation approach more effective in meeting some requirements than others?
What are the limitations, costs, most effective applications, potential risks, and other considerations of a preservation approach?
What are the triggers for determining when to apply a preservation approach?
What are the key factors in applying a preservation approach that must be present (e.g. specific metadata or original formats), accessible (e.g. access to the original environment), known, or avoided to insure its success?
Other questions.

6.4 Preservation Functions

The OAIS reference model provides a useful framework to identify the different functions enabling preservation. Regardless of where the objects are to be preserved or by whom, these functions may enable or inhibit preservation and may affect the success of preservation approaches. Whether or not a preservation function will be effective should be evaluated by metrics (see section 5). The experiments will address these functions.

6.4.1 Ingest

A process is needed to verify and accept the results of capture. Ingest could be considered to be the last step in capture, i.e., the objects are not successfully captured until the objects are verified and accepted, or as a separate stage. Regardless, capture and ingest are integrally related. Initially, ingest will be addressed as a separate stage. The results of the experiments should provide a basis for determining when and how ingest should/could occur.

When should ingest be performed, by whom and how?
What acceptance criteria are relevant or required?
What factors determine the acceptance criteria: requirements, digital object type, software, file format, etc.?
Are there steps that should occur at ingest to enable a particular preservation approach?
What steps are needed to ingest/accept metadata and digital objects?

How can metadata be associated with the digital objects? What is the best option? What factors determine the best option?
Should/can the metadata be associated with the digital objects at capture or at ingest?
The results of ingest tests will probably influence capture requirements.

6.4.2 *Storage*

Captured files will be stored temporarily. Accepted files will be stored and registered. Preservation approaches will be applied over time to stored files and the resulting files that have been converted, modified, refreshed, etc. will be stored again.
What requirements for registering and storing files affect locating and presenting digital objects?
Which of these requirements need to be considered at capture, during ingest, or in applying a preservation approach?
When should a preservation approach be applied to stored files? How should/could a preservation approach be initiated? What factors could/should determine when and how?

6.4.3 *Access*

Successful preservation requires that each stored and registered digital object can be found using defined criteria and presented in accordance with defined requirements that are stipulated at ingest. Like capture and ingest, storage and access are closely related. Access tests the success of registration and storage and could be the final step in storage.
What are the factors that affect the successful presentation of stored objects?
What effect does the preservation approach have on access?
What is the most effective method for using metadata to provide access?

6.5 Preservation Process Prototypes

Preservation Process Prototypes (PPP) should be developed as part of experiments to replicate a particular preservation function or to consider the implications of applying a particular preservation approach. The research should also consider the factors that affect using PPPs for specific preservation approaches, object types, software packages, and preservation formats to explore the possibility of generic PPPs. The Experiment Design should identify relevant PPPs for an experiment.

7 TESTBED EXPERIMENT RESULTS AND RESEARCH PRODUCTS

Each Testbed experiment should produce experiment results. Additionally, ongoing analysis and a comprehensive final analysis of the experiment results should generate research products. See sections 1.2, 6.2 and 6.3.12. The Testbed environment should consist of two databases that will store the experiment results and research products. See sections 4.1 and 4.2.

7.1 Research Database

The research database was discussed in some detail in section 4.1, which lists and describes the content of the database. The research database will be the primary

source of information about the research both within the testbed environment, and for external use. The Experiment Process Checklist (see Annex 3.1) will be stored in the research database and will maintain the current status of each completed and ongoing experiment.

The Project Manager will determine what portions of the research database can be released outside the testbed, when and to whom. Sharing information about the research is a primary objective of the Testbed.

7.2 Experiment Database

The experiment database will contain one copy of all of the experiment documentation and results for each experiment. In addition, the experiment database will contain all Submitted Digital Objects (SDO) that are used by experiments and Tested Digital Objects (TDO) that are generated by experiments. See section 4.2 on the experiment database.

7.3 Experiment Results

Experiment results will include any software developed and adapted for the experiment; the TDOs that were produced by the experiment including the full content of the SDO, any converted or otherwise transformed computer files, and additional metadata captured or generated during the experiment; the automatically captured results of the experiment, such as performance logs and audit trails; and the observations of the Evaluation Group and organisational representative or other participants. Each experiment should have a defined set of expected results.

7.4 Research Products

Research products should be generated during the preparation for or the evaluation of an experiment, or for a set of experiments; or as a result of ongoing analysis of the results and findings of the research. The final stage or any other suitable moment during the testbed may entail a comprehensive compilation and analysis of all of the research results. The types of research products that are anticipated are:

1) research papers and discussion drafts on core preservation issues;
2) external research findings, or technical developments;
3) comparative analyses of experiments to identify patterns, trends and new focus areas for experiments;
4) recommendations based upon the results in the form of guidance, procedures and best practices;
5) specifications and examples of software tools and applications to support or enable preservation; and
6) cost models that identify possible preservation cost scenarios.

8. RESEARCH FRAMEWORK FOR TESTBEDS - ANNEXES

Annex 1. Testbed Terminology: Working Definitions

Annex 2. Experiment Process Documentation: Templates, Examples, Explanations

Annex 1. Testbed Terminology: Working Definitions

Author :
Date :
Version :

NOTE: This is a preliminary list of terms, which are used differently in current literature. When reviewing the terms, please consider:

- terms from the OAIS reference model are considered applicable
- suggestions to improve the scope, accuracy, and/or clarity of the definitions
- additional terms that should/could be defined
- alternate names for terms that are unique to the Testbed if the proposed names are not desirable

The list may be adapted depending on the domain of research involved.

Terms that are defined in this document are in *Italics*. The draft definitions provided reflect the way in which terms will be used by the Testbed rather than universal definitions.

Authenticity: an authentic object is an object that can be proven to be what it purports to be and to have been created (or sent) by the person purported to have created (or sent) it.

Critical Attributes: The *critical attributes* of a *digital object* are the aspects of the object that must be preserved for its successful preservation. Some attributes will refer to intellectual aspects of the digital object, others to the whole process of reproduction. Synonyms are *significant properties, essential characteristics,* and *characteristic properties*. The *authenticity* requirements define the critical attributes that must be preserved for a particular object. Preservation of the attributes should be achieved by capturing *metadata* that describe the attributes, using one of the *preservation approaches*, that meets the *requirements*.
Five types of attributes for a digital object have to be considered:

> **Content:** the comprehensive set of information contained in a object.
> **Context:** the origin of the digital objects, including the organisation that created it, the business function, legal mandate and business process. The context of provenance makes an digital object meaningful.
> **Structure:** the logical hierarchy of and relationship between the parts of a digital object.
> **Appearance:** the presentation of a digital object, the way a digital object looked.
> **Behaviour:** the interactive properties of a digital object.

Digital Object: An object composed of a set of bit sequences.
A *digital object* has to be processed (reproduced) to provide the conceptual representation of an entity that has been selected for preservation. For this reproduction more than one computer file may be necessary (for example in the case of multimedia).

Emulation:
The process of using mechanisms (whether software or not) to enable a software application designed to be run on a particular class of software and hardware to be

run on another class of software or hardware without any loss of behaviour, performance or functionality.

Experiment: a scientific test that is carried out in order to study what happens and to gain new knowledge (Oxford Advanced Learned Dictionary).

Experiment Design: An *experiment design* will define the purpose of the *experiment*, the *object types* that will be the focus of the *experiment*, the steps needed to run the *experiment*, the procedures for setting up and running the *experiment*, the *metadata* needed to run the *experiment* and that will result from the experiment, the *requirements* of the software to be developed or used, the ways in which the *requirements* will be addressed, the validation to be performed, and the kind of results expected (e.g., converted computer files, metadata). An experiment design might be used for multiple experiments using different sets of objects, testing options defined in the design, rerunning experiments for comparative results, etc. The Experiment designs are a product of the Testbed.

Experiment Development: The *experiment development* includes all of the steps and activities needed to implement the *experiment design*. The Implementation Group will define a development plan for each *experiment* that will include building and testing the software needed, defining all of the procedures required to run the experiment, and preparing the *Testbed environment* for the experiment.

Experiment Plan: The *experiment plan* defines the sequence of activities to follow to run an *experiment* and serves as a checklist for running an experiment. The content of the plan is outlined in the *experiment design* and completed during *the experiment development*. The plan will include any activities required to capture the *experiment results* of an experiment, particularly when the capture is automated.

Experiment Process: The *experiment process* is the eleven-stage sequence for defining, designing, developing, running and evaluating an *experiment*. See the full description of the experiment process in the research framework.

Experiment Results: Each experiment will produce *experiment results* that are captured and available for evaluation, comparative analysis, and for future experiments. Experiment results will include converted files, captured and generated *metadata*, experiment log files of the sequence and completion of activities performed in an experiment, and the successful completion of required tasks. The experiment evaluations will weigh the success of an experiment against set requirements and criteria, and provide recommendations and guidance based on the experiments as secondary results.

Migration: the process of moving objects from environment (whether software, operating system, or hardware) to another within the framework of specific requirements.

Model objects: objects with a well described, comprehensive and representative set of characteristics and features, used as a control group for experiments on test objects. See *test objects*.

Object Type: group of objects sharing similar characteristics, functionalities and behaviours., e.g. audio, video, multimedia, spreadsheets, email, databases, images.

Original Environment: The (business) context in which *objects* were created and/or used.

Preservation Digital Object (PDO): the representation of a *digital object* with all of the components that are needed to support and enable its preservation.

Test Objects: The actual objects that are run through the experiment process.

Annex 2 Experiment Process Documentation: Templates, Examples, Explanations

2.1. Experiment Process Checklist

2.2. Stage 1. Statement of Scope of Experiment [RG]

2.3. Stage 2. Preparation: Analysis/compilation [RG]

2.4. Stage 3. Requirements Definition [EG]

2.5. Stage 3. Basic Evaluation Checklist [EG]

2.6. Stage 4. Experiment Design [RG]

2.7. Stage 5. Resource Specification [IG]

2.8. Stage 6. Declaration of Go/No Go decision and justification as needed [PM]
[Note: same for Stage 9]

2.9. Stage 7. Test Plan [IG]
[Note: test Plan results recorded in Stage 8 by IG]

2.10. Stage 7. Development Plan [IG]

2.11. Stage 7. Experiment Plan [IG]

2.12. Stage 8 Test Results [IG]

2.13 by Stage 8. Final Evaluation Checklist [EG]

2.14 Stage 10. Experiment Results
[some captured automatically, some documented by EG]

2.15. Stage 11. Experiment Evaluation [EG]

2.16. Stage 12. Consider Results [PT]

Annex 2.1 Experiment Process Checklist template

Experiment Process Checklist

Experiment Number :
Experiment Process Number :
Preservation Approach :
Object Type :
Objects used in Experiment :

Experiment Stage	Description/Status
1 Define Exploration Area	
2 Prepare for Experiment	
3 Define Requirements	
4 Develop Experiment Design	
5 Specify Resources	
6 Go/No Go Decision (1)	
7 Develop Experiment	
8 Test Experiment	
9 Go/No Go Decision (2)	
10 Run Experiment	
11 Evaluate Experiment	
12 Consider Results	

Annex 2.2 Statement of Experiment Scope template

Stage 1. Statement of Experiment Scope
Experiment Process [number]

Author :
Date :
Version :

Preparing the Statement of Experiment Scope is stage 1 in the Experiment Process. The scope of the experiment may be further refined in stages 2-5 but the Statement should provide enough information to support later steps.

---

1. Scope of experiment

Definition of the scope, which may be similar to other experiment processes, but is in some way unique. This is the definition of the area to be explored by the experiments that result from the experiment process.

1.1 Object Type

1.2 Preservation Approach Type

2. Considerations

Summary of the factors that determined or influenced the definition of the scope.

3. Consensus

Statement about the date on which consensus was reached on the scope of the experiment. The Research Group is responsible for the Statement of Experiment Scope with input from the Testbed Team.

Annex 2.3 Experiment Preparation - Compilation/Analysis template

Stage 2. Experiment Preparation - Compilation/Analysis Experiment Process [number]

Author :
Date :
Version :

Compiling and analysing pertinent information in preparation for the experiment is stage 2 in the Experiment Process. The resulting document should provide a summary of current available information sources for use throughout the experiment process.

---

1. Overview

Summary of the step 1

2. Background

Summary of current information on preservation approach, object type and related research in areas that are relevant to the experiment scope.

3. Relevant Sources

List of sources that relevant to and may have influenced the definition of the experiment process.

Annex 2.4 Define Requirements – Requirements Definition template

Stage 3. Define Requirements - Requirements definition Experiment Process [number]

Author :
Date :
Version :

This document defines the authenticity requirements identified by the involved organisation and additional requirements identified by the Evaluation Group as appropriate. This document will form the success criteria for the experiment upon which the evaluation will be based and is also the source for the Basic Evaluation Checklist.

---

1. Background

Brief discussion of specific issues under consideration for this Experiment Process

2. Scope and Specific Considerations

Not to be specific to any particular experiment, generic to the Experiment Process. Should illustrate how this document caters for the specific issues.

3. Authenticity Requirements

Broken down into the following sections. Introductory paragraph should introduce authenticity requirements considered pertinent to the object type applicable to this Experiment Process. Authenticity Requirements identified should be introduced within each of the attributes (content, context, structure, appearance and behaviour) below in an introductory paragraph, then followed by bullet points summing them up. See the definition of Authenticity for an explanation of the difference between authentication and integrity.

3.1 Authentication

What we mean by authentication

3.1.1 Context

3.2 Integrity

What we mean by integrity

3.2.1 Content

3.2.2 Structure

3.2.3 Appearance

3.2.4 Behaviour

3.3 General Requirements

Any requirement considered necessary by the author which cannot be catered for in

the above classifications.

4. Research Questions Considered

Research Questions applicable to authenticity concerns only - research questions are discussed in more detail and with the scope of the whole process in mind, in the Stage 4 document 'Experiment Design'

5. Recommendations for Addressing Research Questions

Make recommendations for the ways in which the research questions identified above can be answered, or begin to be answered. This is good preparation for the end of the experiment process, in which evaluation is required.

Annex 2.5 Define Requirements – Basic Evaluation Checklist template

Stage 3. Define Requirements - Basic Evaluation Checklist Experiment Process [number]

Author :
Date :
Version :

This document will match the identified requirements and other requirements to measures of experiment results to produce a checklist that will form the basis for evaluating experiments. It will be refined during the experiment process and finalises prior to stage 10: running the experiment.

---

1. Background

Identify any specific issues under consideration for this experiment process

2. Checklist

Each of the following attributes should now be broken down into a running list of elements considered pertinent to the authenticity of the digital object in this experiment process. Authenticity requirements come from the Stage 3 document 'Requirements definition'. They should be peculiar to the digital object type under consideration. The checklist does not require completion at this stage.

2.1 Content

2.2 Context

2.3 Structure

2.4 Appearance

2.5 Behaviour

Annex 2.6 Experiment Design template

Stage 4. Experiment Design
Experiment Process [number]

Author :
Date :
Version :

Developing the Experiment Design is stage 4 in the Experiment Process. The Experiment Design should provide enough information to specify the resources needed (stage 5). The Experiment Design and the Resource Specification should enable the Go/No Go decision (stage 6) and the development of the experiment (stage 7). One Experiment Design may be used by more than one experiment.

---

1. Overview

Summary of the purpose of the experiment

2. Background

Summary of stage 1, Define Exploration Area, and stage 2, Prepare Experiment, plus relationship to other experiments.

Link(s): relevant documents resulting from stages 1 and 2

3. Scope

Defined in Stage 1 and refined as needed through stage 4.

3.1 Object Type

Link(s): entries relating to object types

3.2 Preservation Approach Type

Link(s): entries relating to preservation approach types

3.3 Proposed experiments using experiment design

Link(s): relevant experiment plans defined in stage 7, Develop Experiment

4. Research Questions Addressed

4.1 General

Link(s): relevant documents on research questions

4.2 Specific to experiment

5. Basic Stages and Procedures

6. Metadata

6.1 Provided for experiment

6.2 Produced by experiment

The definition of metadata that is required to complete the experiment or that will result from the experiment, whether captured automatically or entered by a user. The metadata will be determined or influenced by the object type, preservation approach and requirements.

7. Validation

Success criteria based upon user requirements defined in stage 3, Define Requirements, by the EG, and additional technical requirements defined by the Testbed Team.

Link(s): relevant requirements definition

8. Expected Results and Products

Validation reports, Preservation Digital Objects, etc.

Link(s): relevant experiment results

9. Summary of Experiment Development Requirements

Basis for estimating stage 5, Specify Resources, and completing stage 7, Develop Experiment.

Link(s): relevant test plans and development plans

10. Prerequisites

Definition of limitations or logistical considerations based upon a required experiment sequence, hardware/software requirements, etc.

Annex 2.7 Resource Specification template

Stage 5. Resource Specification
Experiment process [number]

Author :
Date :
Version :

This document is intended to provide a detailed estimate of the time and resources needed to build, test and perform the Experiment. Estimates are to include possible options for building, testing and performing the experiments and the implications of each. Along with Stage 4's Experiment Design document, it forms the basis for the Go/No Go Decision required in Stage 6 by the Project Manager.

---

1. Requirements

2. Resources

2.1 Mandatory Items

2.2 Desirable Items

2.3 Enhancements

Annex 2.8 Declaration of Go/No Go Decision template

Stage 6. Go/No Go Decision
Experiment Process [number]

Author :
Date :
Version :

Before the Experiment Process and Experiment can be taken any further, a Go/No Go Decision must be made by the Project Manager, taking into account the Experiment Design and the Resource Definition and/or any of the other documents produced for this Experiment Process. This document charts the decision made and the reasons why.

---

1. Decision Maker

This section allows for the naming of the decision maker. This will generally be the Project Manager but may be made by someone else when the Project manager is absent

2. Decision

One of the following options must be checked:

i. Go - Proceed as planned
ii. Provisional Go - Revisions to the Design or Statement are needed before the Experiment can proceed
iii. Deferred Go - The Experiment should be delayed for a specific period or until a specific event occurs
iv. No Go - The Experiment should be terminated at this stage

3. Reason for Decision

This section allows for any explanatory notes considered necessary about the decision.

4. Action Needed

If the Decision is either Provisional Go or Deferred Go, action is required. This section should detail or identify that action or event.

Annex 2.9 Test Plan template

Stage 7. Develop Experiment: Test Plan
Experiment Process [number]
Experiment [number]

Author :
Date :
Version :

This document will provide an explicit plan for testing experiments and will define mechanisms for testing or capturing, as appropriate.

Annex 2.10 Development Plan Template

Stage 7. Develop Experiment: Development Plan
Experiment Process [number]
Experiment [number]

Author :
Date :
Version :

The Experiment Development Plan should include all of the steps and/or activities needed to implement the Experiment design produced in Stage 4 of the Experiment Process.
This plan is to include a i) an indication of the steps needed to build and test the specified software, ii) a definition of the procedures required to run the experiment, and iii)the provisions required to prepare the Testbed environment for the experiment.

Annex 2.11 Experiment Plan template

Stage 7. Develop Experiment: Experiment Plan
Experiment Process [number]
Experiment [number]

Author :
Date :
Version :

This document serves to define the sequence of events and/or activities that are required to run an experiment. It may also serve as a checklist for running a specific experiment. The content of this plan has already been specified in Stage 4 - Experiment Design, and will be elaborated upon during the experiment development. This plan should also include details of any activities which are required to capture the results of an experiment.
There will be one Experiment Plan per Experiment, but there may also be more than one Experiment Plan for each Experiment Design (the Design belongs the Process and can therefore be used in different experiments.)

Annex 2.12 Final Evaluation Checklist template

Stage 8. Test Experiment: Final Evaluation Checklist
Experiment Process [number]
Experiment [number]

Author :
Date :
Version :

This checklist is a refined version of the Basic Evaluation checklist that is produced in Stage 3 of the Experiment process. By this stage it should be in a form that allows for values to be assigned to each of the elements and attributes previously specified. It is intended to perform as a means of evaluating the Experiment, to track any changes produced in the objects via the preservation process

---

1. Background

2. Checklist

Introductory paragraph explaining procedure and why.
This is left blank for the time being to accommodate any changes which may yet be made to the basic Evaluation checklist.

Annex 2.13 Test Results template

Stage 8. Test Experiment: Test Results
Experiment Process [number]
Experiment [number]

Author :
Date :
Version :

These Test results stem from the Testing specified in Stage 7, Test Plan. When the Testing has been carried out, the results are to be captured in this document. It is anticipated that the tests may provide aspects of an experiment that need to be corrected or places where additional development may be required.

Annex 2.14 Experiment Results template

Stage 10. Experiment Results
Experiment Process [number]
Experiment [number]

Author :
Date :
Version :

This document captures the essential information about the changes produced (or not, as may be the case) during the preservation process. Some results may be captured automatically, others manually, but all need to be entered here. These results will form the basis for the Evaluation of the experiment, and will also provide useful information for the focus of future experiments.
Experiment results may include any of the following: Converted Files; Captured and Generated Metadata; Experiment Log Files of the sequence and completion of activities performed in an Experiment, and the successful completion of the required tasks.

Annex 2.15 Evaluate Experiment template

Stage 11. Experiment Evaluation
Experiment Process [number]
Experiment [number]

Author :
Date :
Version :

The Stage 8 Final Evaluation Checklist and the Stage 10 document, Experiment Results, are the basis for this document. It is intended that this serve as a comparative focus point between the respective SDO's and TDO's.
The evaluation should consider results of any form, either automatic or manual, and also any additional observations not accounted for in the Checklist. It should focus on specific experiments.

Annex 2.16 Consider Results template

Stage 12. Consider Results - Testbed and Experiment Recommendations
Experiment Process [number]
Experiment [number]

Author :
Date :
Version :

This final stage in the Experiment Process allows for the entire Testbed Team to consider not only the results of the Experiment but also the various stages of the Experiment itself. This will allow for recommendations for refinements and enhancements for future experiments to be made, to propose further experiments, and also to provide input into the evaluation of the Testbed.
This review considers the results in a broader view from that of Stage 11 in that it takes into account the development of the Testbed itself.
