

Because good research needs good data

About (/about)	▼	Home (/) » Guidance (/guidance) » Briefing Papers (/guidance/briefing-papers) » Technology Watch Papers (/guidance/briefing-papers/technology-watch-papers)
News (/news)	▼	
Events (/events)	▼	
Services (/services)	▼	
Guidance (/guidance)	^	
Briefing Papers (/guidance/briefing-papers)		
How-to Guides (/guidance/how-guides)		
Case Studies (/guidance/case-studies)		<i>By Martin Donnelly, DCC</i> <i>Published: September 2008</i> - Introduction - The Planets Testbed - Functionality - Selected Implementations - Additional Resources 1. Introduction Planets (Preservation and Long-term Access through Networked Services) is a four-year EU-funded project (IST-2006-033789) which aims to meet some of the key challenges of digital preservation. Its primary goal is the creation of practical services and tools which help ensure long-term access to cultural and scientific assets. The Planets Testbed comprises a software application within a research framework, and aims to bolster what we currently know about preservation tools with new knowledge derived from practical experimentation. In their own words, the Planets Testbed developers aim to ensure that "digital preservation research is tackled as an engineering problem rather than as an art and craft." The Testbed offers a controlled, measured, part-automated and reproducible environment for the testing and evaluation of third-party preservation tools, and a forum for the empirical comparison and analysis of experimental results across the sector. This should encourage the development of increasingly-automated digital
Policy Analysis (/guidance/policy)	▼	
Metadata (/guidance/standards)	▼	
Curation Lifecycle Model (/guidance/curation-lifecycle-model)		
Data Management Plans (/dmps)		
Research (/research)	▼	
Publications (/publications)	▼	
FAQ (/faq)	▼	

Information for
(/information)



preservation solutions, and concurrently raise awareness of preservation- and obsolescence-related issues within the software development community.



2. The Planets Testbed

Preservation tools can be split into three main categories:

Characterisation tools abstract the essential characteristics of a digital object from a file, and can be used to confirm that a file is of the type that it purports to be. Characterisation tools can also be used to extract and identify the significant properties of files, such as image bit depths, document paragraph spacing, and so on.

Migration tools convert digital objects from one file format to another in order to improve the prospects of long-term accessibility. (N.B. it is important to note that important information can easily be lost in this process.)

Emulation tools render digital objects in their original context on a different platform. (With emulation, the Planets Testbed web-service architecture faces the challenge of evaluating tools and services that do not change the object itself, but rather its *representation environment* while leaving the original untouched.)

By gaining an understanding of which tools best serve their needs, users can build preservation plans that indicate which of these preservation tools will be most effectively applied to a collection of digital objects in order to meet (operational) requirements.

The Planets Testbed is built atop a platform-independent, robust, scalable, service-oriented architecture. Tools are wrapped as web services, and accessed directly by the Testbed application. This approach enables experiments to be carried out on legacy tools, and tools which are not directly compatible with the platform on which the experiment is being run.

In order to be analysed as part of an experiment, each preservation tool must first be wrapped as a Web Services Interoperability Organisation (WS-I)-compliant web service, registered within the Planets Testbed system, and have a service template created for it. The service registration process consists of a five-step wizard-based interface which guides the administrator through the tasks required to make a service available. Experimenters can then

access these templates to *simulate* the specific usage of a tool under different circumstances, and users can build complex preservation workflows comprising a series of simpler preservation services. Workflows can therefore be built with confidence that the tools they involve are appropriate and reliable for their given purposes.



3. Functionality

The Testbed workflow is as follows:

Stage 1: Define basic properties

Stage 2: Design experiment

Stage 3: Specify outcomes

Stage 4: Approve experiment

Stage 5: Run experiment

Stage 6: Evaluate experiment

In the first three stages, the user designs the experiment, specifying its focus and intended goals, and selects the services/data on which the experiment will be carried out. An automated decision is then made as to whether or not the experiment should be carried out via an embedded experiment approval algorithm. The algorithm checks the intensity of the experiment against a number of criteria (e.g. the complexity of the selected services, the volume of experimental data, and the current server load). If multiple experiments are already running, a new experiment may not be automatically approved. In this case the Testbed administrator is notified, and can then decide whether to override the algorithm manually, or to reschedule it to execute once the experiments already running are complete.

Once the experiment receives the green light, it is executed at stage 5, and evaluated at stage 6 according to the user's requirements.

A typical Testbed workflow-based experiment will involve (all or a subset of) the following stages:

invoking a **characterisation** service on the input data to determine their significant properties and appropriate migration tools

subsequently invoking (one or many) **migration** service(s) for the execution of data migration, in sequence or in parallel

automatically invoking a **second characterisation** service to assess the results of the migration

All Testbed users can explore the database of prior experiments, comment on their design, methodology and outcomes, and browse or create knowledge trees based on experiment references. It is therefore possible to build up a fuller picture of a tool's performance over the course of multiple discrete experiments. These results can then be aggregated to give average information about the performance of tools on various types of digital objects; in short, as the number of experiments carried out increases, so too does the credibility of this experimentally-derived information.



4. Selected Implementations

The public release of the Planets Testbed is scheduled for Spring 2009. While a downloadable version will also be made available, Planets project partners and other users will initially be encouraged to use a central instance (hosted by HATII at the University of Glasgow), which will enable the Testbed developers to monitor and aggregate experiment results within a single, searchable database. This will provide the preservation and memory institution sector with a valuable resource for analysis and further experimentation.

The DCC will also utilise the Planets Testbed in order to test curation and preservation tools and strategies. To facilitate the experimentation process, we have developed the DCC Methodology for Designing and Evaluating Curation and Preservation Experiments ([/webfm_send/~documents~publications~TestBedMethodV1.1.pdf](#)) **[PDF, 600KB]** which will serve as a workflow framework for designing experiments to validate the effectiveness of curation and preservation strategies.

Another outcome of the Planets project is the development of new tools for digital preservation. These will in turn be deployed within the Testbed, and evaluated using data from the Testbed corpora themselves (comprising damaged files, edge cases, and unusual formats, which can be utilised in testing tools and services), or datasets supplied by the wider user community.



5. Additional Resources

Planets website (<http://www.planets-project.eu/>)

Brian Aitken, Petra Helwig, Andrew Jackson, Andrew Lindley, Eleonora Nicchiarelli, Seamus Ross (2008). "The Planets Testbed: Science for Digital Preservation (<http://journal.code4lib.org/articles/83>)" , The Code4Lib Journal, Issue 3, 2008-06-23



(<https://www.linkedin.com/company/digital-curation-centre/>)
(<https://www.digitalcurationcentre.org/collections>)
(<https://www.digitalcurationcentre/>)

(<https://www.youtube.com/user/TheDigitalCuration>)
(<https://www.digitalcurationcentre.org/about-site/accessibility>)

Accessibility Unless explicitly stated otherwise, all material is copyright ©

(<https://www.digitalcurationcentre.org/about-site/accessibility>) Digital Curation Centre 2004-2026

(<https://www.digitalcurationcentre.org/about-site/accessibility>)

Contact
(/contact-us)

Privacy
&
Cookies
(/about-us/about-site/website-terms-use/privacy-cookies)

Terms
&
Conditions
(/about-us/about-site/website-terms-use)