JISC

JISC Project Plan

InSPECT Project

Overview of Project

1. Background

Significant properties are those aspects of the digital object which **must** be preserved over time in order for the digital object to remain accessible and meaningful. The various properties of a digital object may be categorised as content, context (metadata), appearance (eg. layout, colour), behaviour (eg. interaction, functionality), and structure (eg. pagination, sections). Deciding which aspects of each of these categories must be preserved over time is essential to proper preservation planning but is an essential investigation which has not yet been undertaken. There is a critical need to initiate this work in order to establish best practice approaches to preserving digital objects.

The current consensus about digital preservation holds that approaches that are data-centric, ie. concerned about keeping the data object useable over time, offer better prospects for success than those which are process-centric, ie. concerned to keep original software and/or hardware environments operational over time. For a data-centric approach, understanding the significant properties of digital objects is crucial to successfully preserving those objects over time. Process-centric approaches, however, make an underlying assumption that all properties of digital objects are significant and all must be preserved. This assumption needs to be tested and either justified, modified or abandoned.

The fundamental challenge of digital preservation is to preserve the accessibility and authenticity of digital objects over time and domains, and across changing technical environments. This requires acceptance both of the inevitability of change,¹ and of the inherent separation of the logical information object from its physical environment. Any successful preservation strategy must reconcile the requirement to maintain the fixity/integrity of that logical information object, with the inevitable transformation of the technical environment in which the object resides. A useful conceptual model for understanding this is provided by the 'performance model and the concept of essence' developed by the National Archives of Australia (NAA) in 2002.²

The essence of a performance can be defined in terms of a set of measurable properties, which must be preserved. These properties apply to the logical information object. In addition, the planning and implementation of successful preservation strategies requires a detailed understanding of the properties of the physical object. The most fundamental of these is the representation format in which the information object is encoded, although other explicit and implicit properties also need to be understood. Without an understanding of these 'representation' properties, it is impossible to assess the need for preservation action (in terms of risk), or to identify appropriate preservation strategies.

An organisation with curatorial responsibility for digital objects cannot assert or demonstrate the continued authenticity of those objects over time, or across transformation processes, unless it can identify, measure, and declare the specific properties on which that authenticity depends. Nor can it

¹ It has been observed by Su-Shing Chen (at the time Professor of Computer Engineering and Computer Science at the University of Missouri-Columbia) that "If we hold onto digital information without modifications, accessing the information will become increasingly more difficult, if not impossible." See "The Paradox of Digital Preservation", *Computer* March 2001, 2-6; available at: http://www.gseis.ucla.edu/us-interpares/pdf/ParadoxOfDigitalPreservation.pdf [last checked 14 June 2006].

² I I I lealer at al. An Americant to the Preservation of Digital Decords National Archives

² H Heslop et al, An Approach to the Preservation of Digital Records, National Archives of Australia, 2002.

undertake the preservation actions required to maintain access to those objects, unless it can characterise their current technical representations with sufficient detail.

2. Aims and Objectives

The project will examine the whole concept of significant properties, determine which properties are significant for a range of object types and assess the importance of each of these for future representation of the object, and finally propose a generalised methodology that will enable resource curators to determine the significant properties of classes of digital objects that must be preserved over time. The project will not attempt to define significant properties for high level objects types, such as document like objects. Instead the project will work at a lower level of granularity, focussing on 4 object types initially:

- raster images
- emails
- structured text
- digital audio

These list may be altered or refined during the early phases of the project. The focus of the project will be on the methodologies to be developed for quantifying, comparing and assessing the significant properties of digital objects generally, rather than just on the specific properties of specific object types.

In this way, the project will take forward suggestions regarding the need for further research to support preservation of digital images, relating especially to risk assessment and data loss through migration, made in the recent Digital Image Preservation study undertaken by AHDS for JISC.

In summary the project will:

- expand and articulate the concept of 'significant properties'
- determine sets of significant properties for a specified group of digital object types
- evaluate methods for measuring these properties for a representative sample of representation formats
- investigate and test the mapping and comparison of these properties between different representation formats
- identify issues which will require further research

3. Overall Approach

The identification of those properties which are significant to the continued preservation and accessibility of authentic digital objects, across changing technical environments, is a fundamental task for successful digital preservation. Unless such properties can be defined in a rigorous and measurable manner, cultural memory institutions have no objective framework for identifying, implementing, and validating appropriate preservation strategies, nor for asserting the continued authenticity of their digital collections.

In recent years, the need to identify such properties has been highlighted within a number of notable digital preservation programmes. These include the National Archives of Australia, the Electronic Record Archives programme at the National Archives and Records Administration, The National Archives' Seamless Flow programme in the UK, and the EU-funded DELOS project. Some conceptual work on authenticity and object properties has been undertake as part of the InterPARES 2 project at the University of British Columbia.³ However, to date, little research has been undertaken on the practical application of the concept and approach. It is therefore widely recognised that there is a pressing need for practical research in this area, to develop a methodology, and begin identifying quantifiable sets of significant properties for specific classes of digital object.

³ InterPARES 2 Project website: http://www.interpares.org/ip2/ip2_index.cfm.

This project will address this need, firstly by locating and investigating all references to 'significant properties', 'significant characteristics', or 'essence' in order to clearly articulate a complete and appropriate working definition of the concept. If necessary, the concept will be developed and expanded to reflect current thinking in digital preservation and to provide a baseline definition for the purposes of the project. Once the concept is properly understood and articulated, the project will examine a range of digital object classes to analyse, assess and specify the significant properties of each class of object.⁴ This work will be used as a basis for developing and testing a generic methodology for identifying and measuring significant properties of digital object types. The project will demonstrate the practical application of this methodology to the preservation process, and the effect maintenance of significant properties has on the usability of migrated objects. The project will also examine the application of this methodology to preservation planning, and the validation of preservation actions. It will also identify issues for future research.

4. Project Outputs

- Detailed Project Plan
- Project Website
- Progress Reports
- Completion Report
- Significant Properties Concept Report
- Significant Properties Template
- Significant Properties for Specific Representation Formats (list)
- Significant Properties Assessment Methodology (report)
- Significant Properties Measurement Methodology (report)
- Significant Properties Comparison Methodology (discussion paper)
- PRONOM Enhancement Requirements (report)
- Object Migration Scenarios (paper)
- Source Representation Formats for Migration (list)
- Target Representation Formats for Migration (list may be combined with previous)
- Migration and Significant Properties Testing Analysis (report)
- Enhanced PRONOM Registry (database)

5. Project Outcomes

Several items of value to the wider community will emerge from this project:

- literature review of work on the concept of essence/significant properties
- developed and fully articulated statement of the concept
- a set of documents setting out significant properties for specific digital object types
- tested generalised methodology for determining the significant properties for other digital object types
- increased understanding of the outcomes of various preservation techniques applied to digital resources
- understanding of how well various representation formats preserve the significant properties of digital object types
- contributions to pilot implementations and models for choosing preservation approaches
- an enhanced version of the PRONOM registry to implement the project outcomes

⁴ In this context we use class to mean a specific category of objects, eg. images.

6. Stakeholder Analysis

| Stakeholder | Interest / stake | Importance |
|--------------------------|---|------------|
| InSPECT project partners | Direct and explicit interest in a successful outcome for the project. | High |
| Expert Advisor | Quality of project outputs. | High |
| UK Government agencies | Enhanced reliability of digital preservation methods. | Medium |
| DPC and DCC | Availability of preservation methodology and usable templates and tools. | High |
| Preservation community | Will have an interest in outcomes and ramifications for digital preservation approaches. | High |
| Research Community | Interested in the outcomes of the project so they can be assured that authentic versions of data are preserved. | Low |

7. Risk Analysis

| Risk | Probability (1-5) | Severity (1-5) | Score (P x S) | Action to Prevent/Manage Risk |
|---|----------------------|-------------------|------------------|--|
| Staffing - inability to attract and retain staff with appropriate skills and experience. | 2 | 4 | 8 | Spread expertise throughout the project, create a clear project plan and document current work to ensure knowledge is not lost. Timely advertising of vacancies. |
| Organisational – project management and work activities not carried out effectively and efficiently. | 1 | 5 | 5 | Employ appropriate staff; provide training as necessary. Ensure appropriate evaluation of outputs. |
| Key stakeholders do not buy in to/support the project. | 1 | 4 | 4 | Ensure regular information flow to all stakeholders, and seek feedback on direction and progress at every opportunity. |
| Technical hardware and software issues. | 2 | 5 | 10 | Thoroughly evaluate hardware and software to be deployed in the project and ensure adherence to standards and best practice. |
| Unsuitable software to perform the representation information extraction and file format conversion. | 2 | 4 | 8 | Thorough desk research to identify software available, and discussion with experts to assess suitability of each for project use. |

8. Standards

| Name of standard or specification | Version | Notes |
|--------------------------------------|---------|-------|
| | | |
| | | |

9. Technical Development

The project will not be undertaking specific software development. Testing of existing tools and interaction with the PRONOM database will take place within the existing technical environments of the project partners.

10. Intellectual Property Rights

IPR in all working papers and reports etc, and any software tools or enhancements produced by the partners will be retained by the authors and host institutions but made freely available on a non-exclusive licence as required by JISC. IPR and copyright in any enhancements made to TNA's PRONOM remain with TNA, and PRONOM code will not be available outside TNA. All other results and outputs will be freely disseminated and available for use by the HE and FE communities. Access to file format information held in TNA's PRONOM is freely available via the Internet.

All project outputs and working data will be archived permanently and made accessible in line with JISC requirements.

Project Resources

11. Project Partners

Project Lead Arts & Humanities Data Service (King's College London) Contact: Alastair Dunning, Communications manager Role: project management and strategic direction; contribution to concept development in work package 2; significant contribution to technical aspects of work packages 2-5.

Project Partner The National Archives (TNA) Contact: Adrian Brown, Head of Digital Preservation Role: contribution to project management; major technical contribution to work packages 2-5.

A consortium agreement will be negotiated with TNA during the first 6 months of the project.

12. Project Management

Project Directors Alastair Dunning AHDS <u>alastair.dunning@ahds.ac.uk</u> Adrian Brown TNA Adrian.Brown@nationalarchives.gov.uk

Dunning and Brown will be joint Project Directors, with project management support provided by AHDS. The project directors will oversee the work of the project and act as advocate for the work in the wider community. They will contribute to the investigation and development of the concept of significant properties under work package 2.

Consultant Andrew Wilson, previously of the AHDS Wilson initiated the project whilst working at the AHDS, and will work on Workpackage 1.

Expert Advisor Paul Wheatley Paul.Wheatley@bl.uk

Expert advice on conceptual and methodological issues will be provided by Paul Wheatley of the British Library who participated in the earlier JISC funded CEDARS and CAMILEON projects.

Technical Officer, 0.2 FTE Gareth Knight Gareth.knight@ahds.ac.uk Technical and research activities associated with work packages 3-5.

Research Officer(s), 0.4 FTE To be appointed 1 – 2 staff provided by TNA who will carry out technical components of work activities 3-5.

13. Programme Support

JISC could provide support in the following areas:

- Managing stakeholder expectations including communication of project scope and limits;
- Understanding and fulfilling JISC reporting requirements;
- Managing the interface between the project and other related work within the repositories and preservation programme.

14. Budget

See Appendix A. The AHDS host institution, King's College London, is making a significant contribution to the project in the form of indirect costs. TNA is making a significant contribution to project costs in the form of indirect costs, and the direct costs of PRONOM enhancement.

Detailed Project Planning

15. Workpackages

See Appendix B.

16. Evaluation Plan

Evaluation will be an on-going process throughout the lifetime of the project. The Project directors will evaluate the deliverables and project outputs as a regular part of their project management role. A final evaluation report will be produced by the Project directors at the end of the project, following completion of all project activities. The role of the Expert Advisor, Paul Wheatley, is to read and evaluate major outputs of the project.

| Timing | Factor to Evaluate | Questions to Address | Method(s) | Measure of Success |
|----------|--------------------------|--|---|---|
| Apr 07 - | State of the art report. | Does the report provide an exhaustive and comprehensive compilation and assessment of the concept of 'significant properties'? | Circulate to Expert Advisor, DCC and DPC for feedback. | Reviewers consider report is comprehensive and exhaustive and the proposed definitive statement of the concept is useful. |

| Sep 07 - | Significant properties methodology. | Does the canonical form adequately describe the properties? Does the methodology properly assess, define and express the significant properties of digital objects? | Testing on a wide range of object types. Feedback from DPC and DCC. Feedback from Expert Advisor. | Reproducible results; agreement from DCC and DPC that methodology is useful and appropriate. Positive feedback from Expert Advisor. |
|----------|---|---|--|---|
| Mar 08 - | Significant properties representation format measurements. | Does the methodology measure the way in which representation formats contain the significant properties of the object type? | Testing across a wide range of representation formats. Circulation to DPC and DCC for testing and comments. Feedback from Expert Advisor. | Reproducible results; agreement from DCC and DPC. Positive feedback from Expert Advisor. |
| Sep 08 | Retention of significant properties across representation formats. | Do the migrated objects retain the significant properties of the object types? Are the tools useful? | Testing across repositories. DPC and DCC assessment using methodology. Outside testing of tools. Feedback from Expert Advisor. | Reproducible results; successful retention of properties following migration; successful use of tools to migrate objects. Positive feedback from Expert Advisor. |
| Sep 08 - | Value of PRONOM enhancements. | Are the enhancements to PRONOM useful to stakeholders. | Testing by stakeholders; request for feedback from users and the wider community. | Positive reaction from user community; increased usage statistics. |

17. Quality Plan

| Timing | Compliance With | QA Method(s) | Evidence of Compliance |
|------------|---------------------------|--|--|
| Throughout | Fitness for purpose | See evaluation plan | See evaluation plan |
| Throughout | Accessibility legislation | Project webpage: the webpage will adhere to JISC-standards | Webpage easily accessible by a variety of hardware and software platforms |

18. Dissemination Plan

A key element of the project will be to share the experience, evaluations and results with the wider community. Dissemination will be an on-going activity throughout the project and will take place through a variety of mechanisms. In particular the project will use the following dissemination pathways:

- a project website to announce and disseminate results and outcomes of this project;
- Digital Curation Centre: the project will develop a relationship with the DCC and seek to disseminate its outcomes to a wider audience through the DCC website;
- email announcements: important milestones and deliverables will be announced on appropriate discussion lists;

- •
- publications: articles will be written for publication in appropriate journals; conference presentations: various key conferences will be targeted for dissemination • opportunities.

| Timing | Dissemination Activity | Audience | Purpose | Key Message |
|------------------|--------------------------------|---|---|---|
| Ongoing | Project Website | Wider community. | Detailed information about the project and related activities. | Importance of significant properties concept. |
| Ongoing | JISC Meetings | JISC Community; Stakeholders. | Detailed information about the project and related activities. | How outputs from this project can be used by others. |
| Ongoing | JISC mailing lists | HE and FE digital preservation community. | To inform of project outputs and outcomes. | Raise awareness of key outputs of the project. |
| October 2008 | Final Report | HE and FE digital preservation community. | Evaluation of the InsPECT Project. | How the aims and objectives of the project have been met. |
| October 2008 | Completion Report | JISC Programme Manager and JISC | Evaluation of project's development | Internal and external issue raised during project |
| Post Sep 2008 | Conference / Journal Papers | Digital preservation community. | Critical evaluation of significant properties methodologies. | Relevance of concept and methodologies to digital object preservation |

Dissemination will take place at an international level in addition to the UK, as appropriate.

19. Exit and Sustainability Plans

| Project Outputs | Action for Take-up & Embedding | Action for Exit |
|--|---|--|
| Project Website (WP 1.3) | | Maintained by AHDS for at least 3 years following end of project. Will be archived by UKWAC and permanently available through UKWAC. |
| Concept Report (WP 2.2) | Circulation to the digital preservation community through DCC, DPC, and JISC email lists. | Maintained on project website and with UKWAC. Circulation to community via DPC and DCC. |
| Migration results analysis (WP 5.7) | | Maintained on project website and with UKWAC. Circulation to community via DCC and DPC. |

| Project Outputs | Why Sustainable | Scenarios for Taking Forward | Issues to Address |
|---------------------------------------|---|---|--|
| Assessment Methodology (WP 3.5) | Practical methods for determining which are the significant properties of any digital object type | Develop practical implementation template. Commitment to maintenance and circulation by AHDS and TNA | Who will develop template. Who will provide resources to maintain and update? |

| Measurement Methodology (WP 4.4) | Practical methods for measuring how representation formats contain significant properties of digital object types | Develop practical implementation template. Commitment to maintenance and circulation by AHDS and TNA | Who will develop template. Who will provide resources to maintain and update? |
|---|--|---|--|
| Comparison Methodology (WP 4.5) | Practical methods for comparing retention of significant properties across various migration target formats | Develop practical implementation template. Commitment to maintenance and circulation by AHDS and TNA | Who will develop template. Who will provide resources to maintain and update? |
| Enhanced PRONOM database (WP 5.8) | Unique global source of digital file format information | PRONOM will continue to be developed and maintained by TNA | Long-term resourcing commitment |

Appendix: Workpackages

| WORKPACKAGES | Month | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 |
|--------------|-------|---|---|---|---|---|---|---|---|---|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|
| | | | | | | | | | | | | | | | | | | | | | | | | | |
| 1: | | Х | Х | Х | Х | Х | Х | Х | Х | Х | Х | Х | Х | Х | Х | Х | Х | Х | Х | Х | Х | Х | Х | Х | Х |
| 2: | | | | Х | Х | Х | Х | | | | | | | | | | | | | | | | | | 1 |
| 3: | | | | | | | | Х | Х | Х | Х | Х | Х | | | | | | | | | | | | |
| 4: | | | | | | | | | | | | | | Х | Х | Х | Х | Х | Х | | | | | | |
| 5: | | | | | | | | | | | | | | | | | | | | Х | Х | Х | Х | Х | Х |
| | | | - | | | | | - | | | | | | | | | | | | | | | | | |

Project start date: 12-10-2006

Project completion date: 30-09-2008

Duration: 24 months

| Workpackage and activity | Earliest start date | Latest completion date | Outputs (clearly indicate deliverables & reports in bold) | Milestone | Responsibility |
|---|---------------------------|------------------------------|---|-----------|-------------------|
| | | | | | |
| WORKPACKAGE 1: Project Management | October 2006 | Sep 2008 | | | |
| Objective: This package will develop a detailed project plan, manage and coordinate the work of the partners, to prepare and report as required, and to assess risks and opportunities as the project progresses. This work package also covers dissemination and sharing of outcomes. | | | | | |
| 1.1 Project coordination | 12 Oct 2006 | 20 Sep 2008 | | | Project Directors |

| 1.2 | Project plan | 12 Oct 2006 | 8 Dec 2006 | Detailed Project Plan | Project Directors |
|------|--|----------------|-------------|---|--------------------------|
| 1.3 | Establish and maintain InSPECT website | 1 Dec 2006 | 30 Sep 2008 | Website available from January 2008, containing description of project and links to project documents. Ongoing updating and maintenance | Technical Officer |
| 1.4 | Prepare JISC biannual progress report | 12 Oct 2006 | 31 Jan 2007 | Biannual report for Oct 2006 to Jan 2007 with financial statement | AHDS project director |
| 1.5 | Prepare JISC biannual progress report | 1 Jul 2007 | 30 Jul 2007 | Biannual report for Feb to Jul 2007 with financial statement | AHDS project director |
| 1.6 | Prepare JISC biannual progress report | 1 Jan 2008 | 31 Jan 2008 | Biannual report for Aug 2007 to Jan 2008 with financial statement | AHDS project director |
| 1.7 | Prepare JISC biannual progress report | 1 Jul 2008 | 31 Jul 2008 | Biannual report for Jan 2008 to Jul 2008 with financial statement | AHDS project director |
| 1.8 | Prepare JISC Final report | 1 Apr 2008 | 11 Oct 2008 | Final biannual report for Feb to Sep 2008 | AHDS project director |
| 1.9 | Prepare JISC Final evaluation report | 1 Apr 2008 | 11 Oct 2008 | Final evaluation report | Project Directors |
| 1.10 | Prepare JISC Completion Report | 1 Oct 2008 | 11 Oct 2008 | Completion report for duration of project with full financial statement | Project Directors |

| WORKPACKAGE 2: Investigate and expand the concept of 'essence' / 'significant properties' <u>Objective</u> : This package will explore the current versions, views, and visions of the concept at the core of this proposal, the idea of 'significant properties', also referred to as 'significant characteristics' and 'essence by collecting all discussions of the concept, analysing them and developing a clear, consistent articulation of the | Dec 2006 | Mar 2007 | | |
|--|----------------|-------------|---|--|
| concept. The project will review work being undertaken within the DELOS project, and now continuing within the EU-funded PLANETS project (in which TNA is participating), within NARA's Electronic Records Archive project in the US, and at the National Archives of Australia. | | | | |
| 2.1 Data gathering and analysis | 11 Dec 2006 | 16 Mar 2007 | Preliminary data gathering for investigation of the concepts involved in the project | Project Directors; Technical Officer, Consultant |
| 2.2 Significant Properties Report | 4 Mar 2007 | 30 Mar 2007 | Comprehensive state of the art report evaluating and assessing existing discussions of the concept of 'significant properties', and clearly stating an agreed definition for the project | Consultant |
| | | | | |

| WORKPACKAGE 3: Analysing the significant properties of digital object types | Apr 2007 | Sep 2007 | | |
|--|----------------|-------------|--|---|
| Objective : This package will review and amend if necessary the list of object types to be examined. This will be followed by identification of a set of significant properties for each object type, which will be expressed in the form of a template of properties. The package will also identify an appropriate canonical form for expressing these properties. In some circumstances, the demands of authenticity may best be served by defining a significant property in terms of a range of permitted values, rather than as an absolute. This work package will define permitted tolerances for each significant property. The package will articulate a generalised methodology for assessing, defining and expressing the significant properties of digital objects, capable of broader application. | | | | |
| 3.1 Finalise list of object types to be investigated | 1 Apr 2007 | 14 Apr 2007 | Working paper reviewing and assessing object types to be examined | Technical Officer + Research Officer(s) |
| 3.2 Develop first iteration of properties template | 15 Apr 2007 | 31 May 2007 | Draft template for each object type, setting out significant properties of the type | Technical Officer + Research Officer(s) |
| 3.3 Develop canonical property expression methodology | 15 May 2007 | 31 Aug 2007 | Detailed method for a canonical way of expressing the significant properties of digital objects (may be expressed as an XML schema) | Technical Officer + Research Officer(s); Project Directors |
| 3.4 Develop final version of properties template | 1 Jul 2007 | 31 Aug 2007 | Detailed template for each object type setting out significant properties of the type, the canonical expression of the property, and the tolerances for each property | Technical Officer + Research Officer(s) |
| 3.5 Properties report | 1 Sep 2007 | 30 Sep 2007 | Report setting out a generalised methodology for assessing | Technical Officer + Research |

| | | | significant properties of digital objects | Officer(s) |
|---|------------|-------------|--|---|
| | | | | |
| WORKPACKAGE 4: Assessing representation formats | Oct 2007 | Apr 2008 | | |
| Objective: This work package will determine methods for describing how the significant properties defined in work package 3 can be identified within a range of common representation formats, and for measuring those properties. The package will identify a representative sample of 3-5 formats for each object type | | | | |
| 4.1 Identify sample representation formats | 1 Oct 2007 | 31 Oct 2007 | Working paper setting out representation formats for each object type | Project Directors; Technical Officer + Research Officer(s) |
| 4.2 Update PRONOM database | 1 Oct 2007 | 31 Dec 2007 | Updated content of PRONOM registry to reflect collated technical documentation | Research Officer(s) |
| 4.3 Analyse property representations | 1 Nov 2007 | 28 Feb 2008 | Analysis Of representation formats to see how well they represent the significant properties | Technical Officer + Research Officer(s) |
| 4.4 Develop measurement methodology | 1 Dec 2007 | 30 Mar 2008 | Report describing generalised methods for measuring significant properties in representation formats | Technical Officer + Research Officer(s) |
| 4.5 Develop comparison methodology | 1 Dec 2007 | 30 Mar 2008 | Discussion paper elaborating methods for comparing significant properties between representation formats | Technical Officer + Research Officer(s) |
| 4.6 Develop PRONOM requirements | 1 Jan 2008 | 30 Mar 2008 | Requirements document for PRONOM enhancements | Research Officer(s) |
| | | | | |

| WORKPACKAGE 5: Tools for assessing and testing significant properties | Apr 2008 | Sep 2008 | | |
|---|----------------|-------------|---|---|
| Objective: This work package will test the methodology developed in work package 4, using sample digital objects from the collections of TNA, AHDS and possibly the British Library. For each object type, this work package will identify source and target representation formats for migration. The package will identify tools to be used in the testing process and will consider the performance of the tools used, their limitations, and any current gaps in tool provision. This work package will measure the properties of the source sample objects, and convert them into their canonical form. The properties of the target sample objects will then be measured, converted to canonical form, and compared to the properties of the source sample | | | | |
| 5.1 Identify migration scenarios | 1 Apr 2008 | 30 Apr 2008 | Working paper setting out scenarios for migration of digital objects (NB may be combined with WPs 5.4 and 5.6) | Project Directors; Technical Officer + Research Officer(s) |
| 5.2 Create sample object collections | 15 Apr 2008 | 15 May 2008 | Locate and collect sample representation formats for each category of object types | Technical Officer + Research Officer(s) |

| 5.3 | Identify and evaluate tools | 15 Apr 2008 | 15 May 2008 | Identify, test and assess tools for assessing properties of digital objects and tools for migrating digital objects | Technical Officer + Research Officer(s) |
|------|------------------------------------|-----------------|-------------|--|---|
| 5.4 | Measure source properties | 16 May 2008 | 1 Jul 2008 | Working paper setting out source representation formats for migration of each object type (NB may be a combined document with WPs 5.1 and 5.6) | Technical Officer + Research Officer(s) |
| 5.5 | Perform migrations | 1 Jun 2008 | 15 Aug 2008 | Undertake format migrations on chosen objects | Technical Officer + Research Officer(s) |
| 5.6 | Measure target properties | 15 June 2008 | 1 Sep 2008 | Working paper setting out target representation formats for migration (NB may be a combined document with WPs 5.1 and 5.4) | Technical Officer + Research Officer(s) |
| 5.7 | Analysis and reporting | 1 Jul 2008 | 30 Sep 2008 | Report detailing testing process and outcomes for each object type, assessment of tools, recommendations for future tool development. The report will assess how well each representation format retains significant properties after migration. | Technical Officer + Research Officer(s); Project Directors |
| 5.8 | Develop enhanced version of PRONOM | 1 Apr 2008 | 31 Jul 2008 | Modified version of the PRONOM registry, to support the automated measurement of properties | Research Officer(s) |
| 5.9 | Release enhanced version of PRONOM | 1 Aug 2008 | 30 Sep 2008 | Public release of enhanced PRONOMN | TNA Project Director |
| 5.10 | Final report | 1 Sep 2008 | 30 Sep 2008 | Final project report , summarising conclusions from work packages 2 to 5 | Project Directors; Technical Officer + Research Officer(s) |
| | | | | | |

Members of Project Team:

Alastair Dunning (AHDS) Andrew Wilson (ex-AHDS, Acting as Consultant to Project – paid from costs to AHDS) Adrian Brown (TNA) Gareth Knight (AHDS) 1/2 Research Officer(s) (TNA) tba