the national archives

Requirements for Electronic Records Management Systems

Functional Requirements for the Sustainability of Electronic Records

Version 1; March 2006

The National Archives
Ruskin Avenue
Kew
Surrey TW9 4DU
United Kingdom

Website: www.nationalarchives.gov.uk/

Author:
Malcolm Todd, Records Management Department

Major contributors:
Adrian Brown, Digital Preservation Department
Richard Blake, National Advisory Services

Expert consultants:
Kevin Ashley, University of London Computer Centre
Andrew Wilson, Arts and Humanities Data Service

Contents

Introduction

These requirements have been developed in response to the need of Departments and Agencies in central government and authorities in the wider UK public sector to retain access to electronic records for extended periods. They follow on from the four volumes of *Generic requirements for the sustainability of electronic records* published by The National Archives in 2003 at:

http://www.nationalarchives.gov.uk/electronicrecords/generic.htm

It is hoped that business requirements of government organisations will prompt the offering of such solutions by suppliers of proprietary ERM solutions, although the National Archives has no plans to evaluate software against these requirements. It is also hoped that this publication will stimulate the development of alternative approaches to the same set of problems: a range of solutions is essential to the long-term information management needs of transformational government. ***No intention on the part of The National Archives to conduct formal software evaluations of the functionality described should be inferred by their publication***.

The purposes of long-term retention may be archival (especially where early archival transfer to a dedicated digital archive is for some reason not possible) or for purely business purposes. There are significant quantities of records now being created digitally that will be required beyond several generations of technology, but not required permanently for historical purposes. For example: some human resource records are required for superannuation purposes well into an individual's retirement.

The main issue, though, remains the same: the preservation of the records (or more precisely, *copies of records enabling the reproduction* of the record in an authentic form) through time and across technological changes. This guidance therefore follows and references principles from authoritative archival and records management texts, especially ISO 15489 *Information and documentation: records management*.

Note that this guidance is not intended to define an Open Archival Information System (OAIS) as set out in ISO 14721:2003: *Space data and information transfer systems – Open archival information system – Reference model*.[1] Accordingly, the terminology used in this document is generally consistent with that in *Requirements for electronic records management systems [Volume 1, frequently referred to as TNA 2002]*, with a few modifications. For specific meanings of various terms used in this guidance, see the glossary at Annex 1.

1 See http://nost.gsfc.nasa.gov/isoas/ref_model.html [last accessed 21/02/2006]

Audience and approach

The main audience for these requirements is information services managers, IT managers, senior Departmental Record Officers with responsibility for digital records. These requirements form part of the deliverables of a sub-Project of the National Archives "Seamless Flow" programme: *Management of semi-current records in Government Departments*[2]. The brief of that project is to ensure as far as is possible that digital records of government business activities can be maintained for as long as they are required by prevailing retention policies[3].

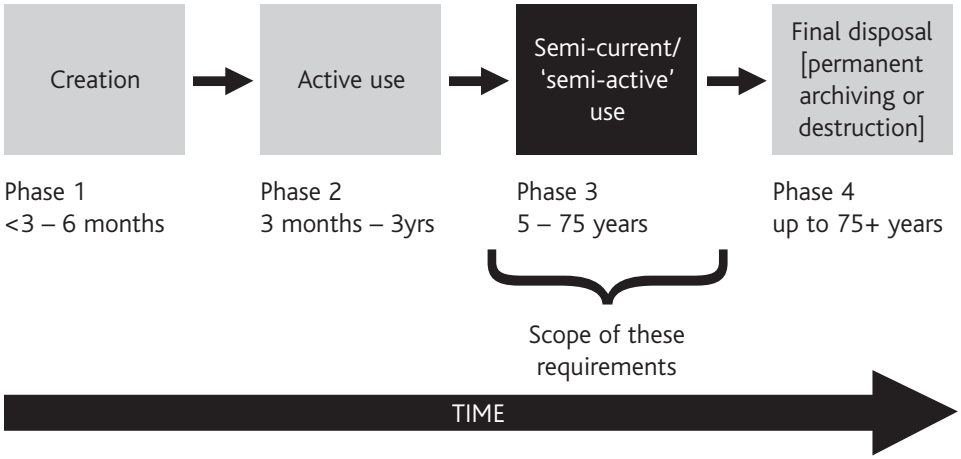
As with previous published requirements, they attempt to stress outcomes rather than specific solutions and *do not form part of any UK Government procurement scheme*. They have the status of guidance: individual organisations must examine their needs for further articulation of some requirements and possibly the reduction of some others in some circumstances. *Reduction of mandatory requirements should not be undertaken without careful consideration of the consequences for satisfying the accompanying rationale*.

2 They are also produced to meet a Key Performance Indicator for the current business plan year 2005-06. KPI 6[ii] reads: ***To produce, in collaboration with other government departments, requirements for the sustainability of electronic records.***

3 Previous deliverables include a governance regime comprising the National Archives' *Custodial Policy for digital records* and transfer agreements clarifying roles and responsibilities to comply with the 1958 Public Records Act in the digital environment.

The records lifecycle

Records are commonly understood to have a lifecycle from creation to final disposal. It can be represented simply as follows:



In the first phase, a document produced by a business activity is judged to be a corporate record. That is, it is recognised as being required for business, accountability or an historical purpose and associated with a logical position in the business classification scheme of the organisation, relative to its other records.

In the second phase, the business unit responsible for the creation of the record – and perhaps also others in the organisation – consults the record frequently. The information it contains, and its importance to understanding recent and current events is required frequently. It may be used to produce further documents and records: such as later drafts and final record versions of the same business process.

The third phase – which concerns us most here – applies to records required for periods *longer than about one generation of technology*. In the third phase consultation and active reuse of the record is no longer so frequent. The aggregation it is contained in should have been closed to prevent further additions. The unit of business it records is considered finished and its records complete and to be managed together for the remainder of their life, until the fourth stage is reached, that of final disposal. *These requirement concern themselves solely with the third phase and its linkage [interfaces] to the second and fourth*. For periods past 5-7 years from creation, there is an enhanced risk that current technology will not provide access to the records if they remain in the form in which they were created.

In the analogue world, paper records followed a very clear physical progression through these phases. The first two took place in the business unit carrying out the activity that created the record. A record written or printed on paper establishes a near-permanent physical bond between the content and the medium. Managing the physical entity managed the content for the rest of the lifecycle. The semi-current phase was when use of the records had declined sufficiently to allow them to be moved to off-site storage, which was generally cheaper than office space.

Whilst some aspects of the lifecycle are still helpful, a few modifications to the traditional understanding of this model are essential to retain its usefulness in the digital environment and they are vital to understanding the driver behind these requirements. ***Digital records will not survive if the approach taken is "out of sight, out of mind"***:

1. Firstly, digital records cannot be maintained simply by placing their carrying or storage media in secure environmentally controlled conditions. The media they are encoded on will decay, threatening the integrity of the data and its accessibility. As a result, the continuing needs of the business for access to them will be in jeopardy. Writing to removable media has a place in back-up routines but is not, of itself, a credible "archiving" solution.

2. Secondly, there is a separation between the physical carrier of the record, the content and the context. The format[s] the content and context are encoded in will become obsolete, requiring them to be migrated to more current technology if they are to remain readable. Such required changes must be done in accordance with proper procedures. Because digital records are far more prone to undetectable alteration, steps must be taken to ensure that unauthorised changes to the record content are prevented. The context of the creation and use of the records must be kept in close association with the content through the management of metadata.

3. Thirdly, public authorities have to comply with the provisions of the Freedom of Information Act 2000 and, in central government, the Public Records Act 1958. Requests from members of the public for access to official information have to be answered promptly, within 20 working days. Public authority records have to be managed in accordance with the Lord Chancellor's *Code of Practice* on records management issued under Section 46 of the FOIA which includes properly managed disposal. Public records required for permanent preservation must be preserved in accordance with the guidance and supervision of the National Archives.

Links to other guidance

In 2003 The National Archives published a set of *Generic requirements to sustain electronic information over time* that define generic functionality requirements for compliance with BS ISO 15489. There are four documents, available from http://www.nationalarchives.gov.uk/electronicrecords/, which should be read together for complete comprehension, as there are important interdependencies:

- Defining the characteristics for authentic records
- Sustaining authentic and reliable records: management requirements
- Sustaining authentic and reliable records: technical requirements
- Guidance for categorising records to identify sustainable requirements

This document extends the generic requirements by providing specific functionality requirements for ERMS that allow for records in those systems to be maintained over time in a sustainable manner.

The approach taken in this document is to provide functional requirements additional to those already contained in the *Requirements for electronic records management systems* [Volume 1], published by The National Archives in 2002 and available from: http://www.nationalarchives.gov.uk/electronicrecords/function.htm This specification refers to those requirements generally as "TNA 2002" and to the specific requirements occasionally as "2002 Functional requirements". A concordance table to those requirements is contained in Annex 2. The National Archives endorses and is participating in the EU-DLM initiative to revise the European Union Model Requirements for records management systems ["MoREQ"] which may incorporate and harmonise all or part of national standards from 2007-08.

Other relevant guidance relating to current records management is:

The remaining volumes of *Functional requirements for electronic records management systems* [accessible from the URL above]:

Vol. 2 *Metadata standard*
Vol. 3 *Reference Document*
Vol. 4 *Implementation guidance*

Rationale for requirements for electronic records management systems published at: http://www.nationalarchives.gov.uk/electronicrecords/rat2002/pdf/erms_section.pdf

Volumes 1-3 of the 2002 TNA requirements have for some time been associated with a software evaluation scheme. In December 2004, The National Archives announced that it was completing the software evaluation scheme for ERMS it began in 2000 and revamped in 2002. Whilst these Requirements are an extension of TNA 2002, there is no proposal for software testing implied by their production.

Instead, we envisage achieving the same objectives for these requirements by a two-pronged approach. Firstly, their implementation through the ITTs and related business requirement specifications produced by UK public authorities in association with their procurement activities. Secondly, exposure of the requirements to the wider records management and archival communities in suitable *fora*, such as EU-DLM, professional bodies, research initiatives, etc. ought to increase awareness of the issues implied by long term retention of digital records.

The National Archives will consider other products to support the implementation of this guidance. At the time of writing (March 2006), these are envisaged to be:

- The provision of tools and guidance for preservation planning, including DROID: http://www.nationalarchives.gov.uk/aboutapps/pronom/tools.htm
- The provision of a technology watch service to assist in the assessment of risks inherent in technical dependencies in digital objects comprising parts of the record[s] and supporting preservation decisions such as migration. See: http://www.nationalarchives.gov.uk/pronom/
- Further iterations of National Archives guidance on records management and archival metadata, as far as possible integrated within the wider framework of the eGovernment Metadata Standard and developed into machine-readable representation[s] such as XML schema or RDF;

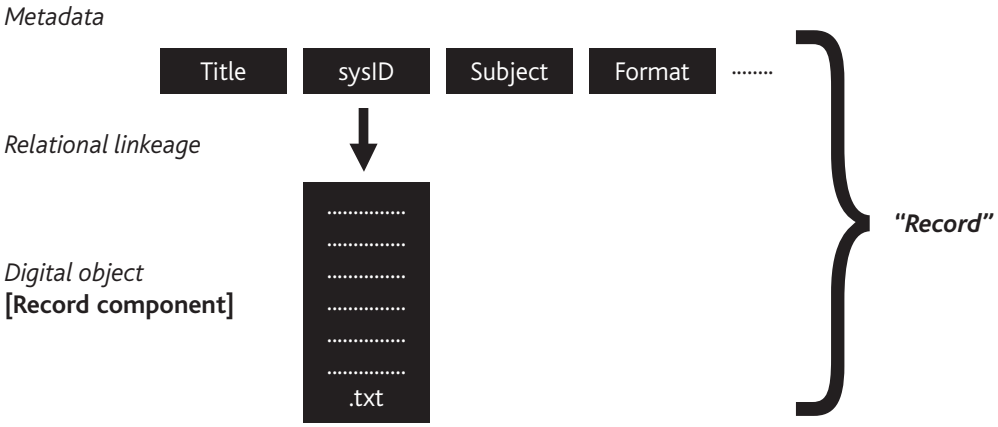
Note on terminology of record, digital object and metadata

BS-ISO 15489 defines a "record" [Section 3.15] as:

"Information created, received and maintained as evidence and information by an organisation or person, in pursuance of a legal obligation or in the transaction of business".

It goes on to discuss the characteristics of a record in section 7.2. Handling the interventionist processes of digital preservation similarly requires treating the "record" as a *conceptual* rather than a *physical* entity and this is consistent with the ISO definition. Pragmatically, it is helpful with these definitions to see the record as *a combination of one or more digital objects* ***and their metadata***. This logical abstraction facilitates migration of the objects from their original format to a more current one without destroying the conceptual record (the objects are linked to the metadata and to one another using relational identifiers).

This is best represented diagramatically in the following way



A *manifestation* of a record is a representation of its content for meeting various different accessibility requirements. These can include:

- *Migration* for preservation across format obsolescence (a *migrated manifestation* where the previous format may be retained);
- Making non-sensitive material available whilst keeping sensitive material secure by producing a *redacted manifestation* for viewing (where the full content version will be held in secure association until it can be released).

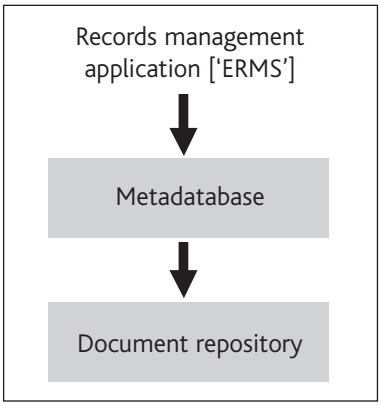
In these *Requirements* it has frequently been necessary to make separate stipulations for the behaviour and management of metadata and digital objects.

Metadata is vital to contextualise the information contained in the record in terms of the business processes that created and have been used to manage (including preserve) it. This is a very important part of meeting the requirement of Section 7.2 of BS-ISO 15489 *Characteristics of a record*. It should be possible to confirm that these processes are permissible and compatible with the record still remaining what it purports to be [authentic], by comparing the object's metadata with the organisation's records management policies, strategies and procedures.

See Terminology Annex [# 1] for definitions of these and other terms.

Logical architecture

A traditional view of an ERMS is that it comprises a metadatabase and a document repository, both under the *sole* control of a database management system or records management "application":



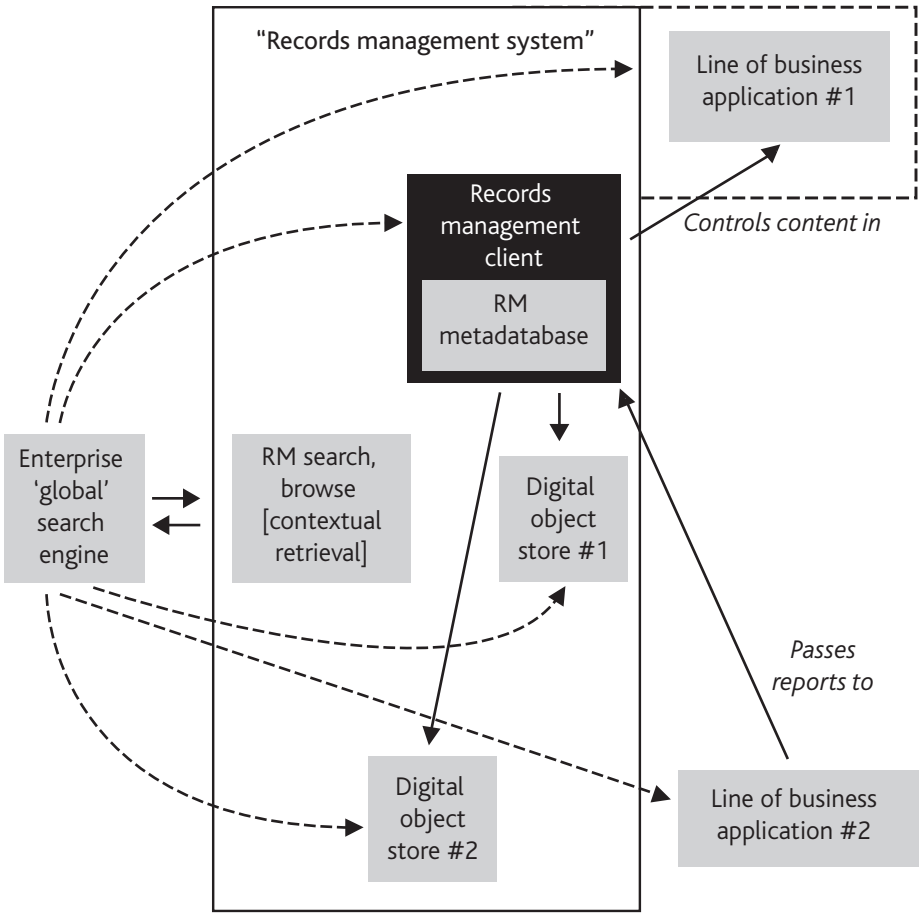
'Traditional' view

This is the simplest model. When a record is created, control passes from the creating application to the ERMS. This control is exercised by rules in the ERMS functionality and instruments such as disposal schedules, access control schemes, business classification schemes implemented through it.

The main technical scenario addressed in these requirements is the use of an application level infrastructure to provide a controlled environment for sustaining the records. This does not necessarily assume the same monolithic viewpoint as in the previous diagram: the issue is one of control not system architecture. In current records management, provided the necessary rules can be configured, there is no reason why the document repository has to be only used for storing records, nor why digital objects in other stores cannot be controlled by the records management application. This is an essential mental leap required to engage with enterprise content management.

It is often more helpful to view the "records management system" as a combination of rules, procedures and technologies and so it will be here. These requirements tackle significantly more detailed system functionality than needed in the TNA 2002 *Functional requirements*, but do not assume that this will all be delivered by the same piece of software. In particular, utilities for technology watch and active preservation activity are not expected to be provided by a single product. Functionality such as scalability may be better facilitated in a more distributed environment and the paradigm maps more closely to enterprise content management than the more monolithic model above. *Getting multiple products to work together may raise complex configuration issues, but the concentration must be on the robustness of the outcome in terms of records management rather than the actual system design / architecture*.

A more realistic logical architecture



Notes on diagram:
1) There may be documents in the stores not under the control of the records management client (undeclared documents, some website content, etc).
2) The two types of search may be provided by the same search engine if it has the integration capability

This is by no means the only approach to tackling this problem. Alternatives might be to use a carefully specified, managed service or even to build a dedicated digital archive. Whilst these solutions are out of the current scope, it is hoped that there may be some cross-fertilisation between these different approaches and contribution to other National Archives' workstreams.

Positioning sustainability within ERMS

Current ERMS functionality typically does little in the way of manipulating digital objects: it is aimed mainly at fixing the content of a declared record, typically preventing *any* subsequent changes apart from the accrual of descriptive metadata. Some software providers have responded encouragingly to certain non-mandatory National Archives requirements published in 2002, designed to lay the ground for longer-term retention (e.g. the rendition requirements). It is hoped that this document will increase the visibility, understanding, demand and hence the support for those 2002 requirements.

Supporting sustainability in this environment effectively means promoting and stressing a number of requirements that *may* be present in the current records management environment, but giving them new emphasis. This is the approach taken in these requirements. The authoring process has involved considering broad areas of ERMS functionality and teasing out what needs to be augmented to support these activities being carried out over extended periods of time. This requires considerably more attention to technical detail.

Some of these areas have then required further development and articulation and a concordance table with the TNA 2002 requirement is in Annex 2. Solutions aimed at meeting the previous published requirement[s] *may* meet this new requirement but should be evaluated carefully to ensure that they do before this can be assumed. The main additions are fuller articulation of digital preservation requirements. There is an important distinction to be made between 'active' and 'passive' preservation processes.

Preservation

'Passive preservation'

In these requirements, "passive preservation" is defined as the provision of secure storage and integrity of each record manifestation. This maps mainly to the ***Management requirements*** [Vol. 2] of the *Generic requirements for sustainability of electronic records* cited above.

- Storage management: including monitoring and refreshing of media to offset the risks of degradation
- Scalability, including particularly issues arising from:
  – Design and performance
- Security, comprising:
  – Functional access control at export
  – Environment outside the ERMS functionality: key references
- Certain management reporting specific to supporting these processes, such as:
  – Media monitoring
  – Integrity checking
  – System performance

'Active preservation'

In these requirements, "active preservation" is defined as such intervention(s) in the bitstream used to encode the content of the records as may be required over time to preserve access to the content of the records and their value as evidence and as cultural artefacts. This maps mainly to Volume 3 and 4 of the *Generic requirements: Technical requirements and Record categorisation*.

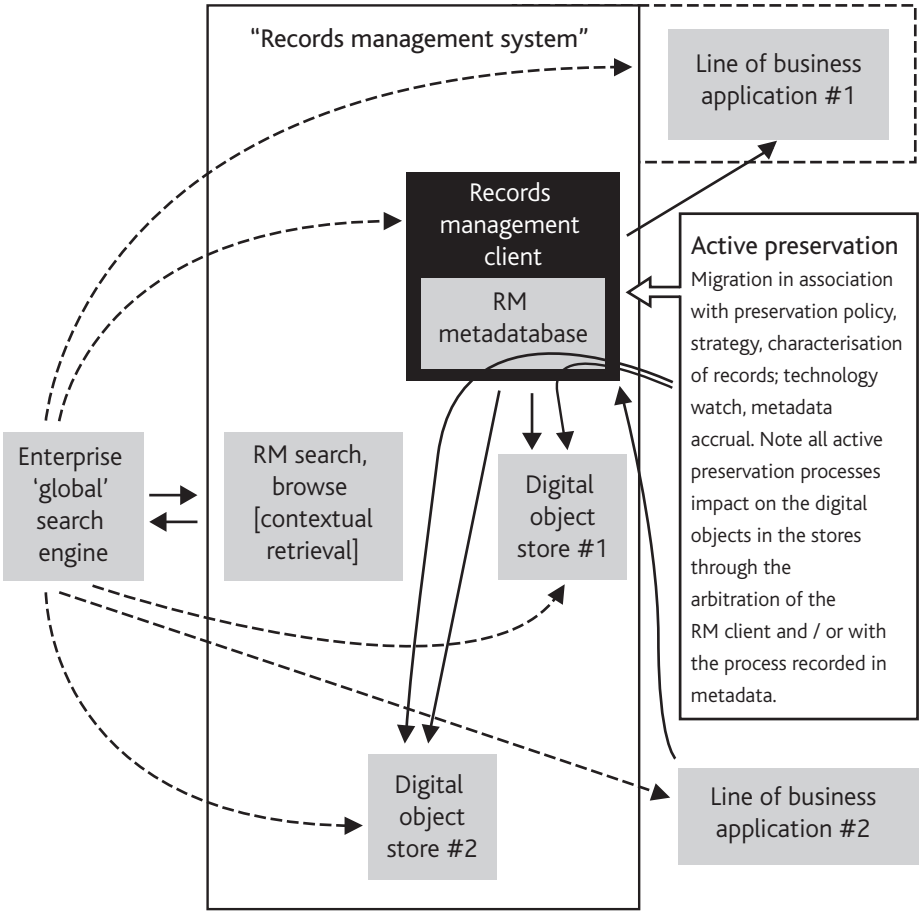
- Identification of current logical object formats and characterisation of how the current content of the records is manifested;
- Preservation planning, a technology watch programme to identify digital object formats and track their technical dependencies and ensure active preservation action is taken in good time;
- Migration of the content objects to preserve accessibility of content and maintain access to the content using more current technology;
- Maintaining interoperability / openness *across migration* through standardisation of metadata interfaces. This ensures the minimum interference in the documentary form of the records and the maximum of automation;
- Certain specific management reporting specific to supporting these processes.

Other sustainability requirements

- Interoperability / openness at import and export and more generally
- Retrieval requirements (particularly to support the servicing of FOIA, DPA and EIR requests according to the design approach taken to meet the above requirements[4])
- Custodial issues and archival mappings
- Other Reporting capabilities not covered in the sections above

4 These requirements occur in the sections on scalability and performance because they are dependent on significant design decisions likely to be made for these reasons.

Effects of active preservation processes on candidate logical architecture



Presentation of these requirements

Each requirement is now examined in turn in the remainder of this section. Their significance is justified in introductory text and then distilled into a high level requirement for the subsequent, more detailed articulation by functional requirements.

A standard format is used to present this information:

x.x Issue

x.x.x Rationale
What the issue is about and why it matters.

Functional requirements
What system functionality needs to be present to achieve the outcome desired in terms of sustainable records management[5].

Non-functional requirements
Other requirements included for completeness, but probably requiring to be implemented by business rule / procedure and not necessarily automatable. A significant number of requirements present problems of assessing compliance that can only be addressed in a 'real life' situation by the provision and assessment of information provided by a potential supplier, and understanding of the implications of particular design features by both parties.

Reporting requirements
Specific reporting requirements specific to supporting the requirements just articulated.

Note on obligation levels in these requirements

The rubric followed is the same as TNA 2002 apart from some small amendments because the present requirements are ***not*** *linked to any software evaluation or procurement scheme*. In this document:

- mandatory requirements are indicated by the phrase "The ERMS *must*..." *It is difficult to imagine a credible solution that does not fulfil these requirements*
- highly desirable requirements are indicated by the phrase "The ERMS *should*..."
- desirable requirements are indicated by the phrase "The ERMS *may* ...".

Each numbered requirement is labelled as:
(M) = Mandatory = "The ERMS *must* ..."
(HD) = Highly Desirable = "The ERMS *should* ..."
(D) = Desirable = "The ERMS *may* .."

MUST This word means that the numbered requirement as defined is an absolute requirement without which the solution is unlikely to be viable.

MUST NOT This phrase means that the numbered requirement as defined is an absolute prohibition.

5 This needs to be present and effective *somewhere in the solution*. It does **not** necessarily have to be provided by a single application; the acronym "ERMS" in these *Requirements* should be taken to mean a combination of software, infrastructure and procedures around them.

SHOULD This word means that there may exist valid reasons in particular circumstances to ignore a particular item, but the full implications must be understood and carefully weighed before choosing a different course. In particular, attention is drawn to the high level requirements at the start of the same section: there a clear idea of how the high level requirement[s] is [are] to be met if the decision is taken to ignore a detailed requirement. An implementation which does not include such an item MUST normally be prepared to interoperate with another implementation which does include the item, though perhaps with reduced functionality. An implementation which does include a particular item MUST be prepared to interoperate with another implementation which does not include the item (except of course for the feature which the item provides).

MAY This word means that the numbered requirement is optional. One vendor may choose to include the requirement because a particular marketplace requires it or because the vendor feels that it enhances the product, while another vendor may omit the requirement. An implementation which does not include such an option MUST be prepared to interoperate with another implementation which does include the option, though perhaps with reduced functionality. An implementation which does include a particular option MUST be prepared to interoperate with another implementation which does not include the option (except of course for the feature which the option provides).

These definitions are drawn from the meaning of terms as set out in RFC 2119 (http://www.ietf.org/rfc/rfc2119.txt).

WHERE / IF [*Conditional requirements labelled "Cond."*] Where a SHOULD option is taken, subsequent requirements may use the phrase "Where < a feature is provided>, the ERMS must ...". Here, MUST means that if the highly desirable or desirable option is offered, the mandatory rider is an absolute requirement. If the highly desirable or desirable option is not offered, the rider does not apply.

- '*Advisory'* requirements from TNA 2002 were a factor of evaluation scheme: indicative information was collected in the testing process. This sort of thing is generally expressed as non-functional requirements in this but must be present for the solution to achieve the outcomes specified.

- Must / should *facilitate* or *support* [see previous text on technical architectures: required in the environment and likely to be an integration but functionality is required if the requirement is mandatory]. These are generally non-functional requirements where some procedure is probably required to achieve a particular outcome.

- A number of requirements considered highly desirable or desirable in TNA 2002 are labelled as *Mandatory* in these requirements as they are seen to be more important for long term retention of records. A mapping table appears in Annex 2.

As with all statements of this kind, Departments and agencies must examine their own business needs and tailor the requirement accordingly, in this case bearing in mind the detailed rationales at the beginning of each section.

The requirements

S.1 Storage media management:[6] Scalability, monitoring and refreshing, and integrity

Scalability rationale

Maintaining older records alongside those being created currently can imply a very significant scalability requirement of the solution. Against that, it should be remembered that disposal management can be implemented more robustly, efficiently and precisely using an ERMS than in any other environment.

Depending on the pace of creation and disposal, though, there may be a strong case for using the same infrastructure provided it can be properly managed. The main advantages of this are:

- A single interface can give access to a substantial proportion of corporate information;
- Organisations with a limited quantity of records that require sustaining for extended periods can have this capability without more substantial investment in digital archiving technologies;
- Off-line storage solutions can be avoided as these can too easily become a case of '*out of sight, out of mind*'. Such an approach may have been more suitable in the paper environment, but are high risk in the digital (see next section on media monitoring);
- The same classification scheme, with its support for contextualised retrieval, metadata inheritance, disposal and security management can be used as for the more current records.

Depending on the architecture of the solution, additional retrieval requirements may be required to support business search / retrieval and the servicing of FOIA / EIR requests in particular. One such scenario is covered in these requirements as an indicative example: where some of the records are held off- or near-line.

Functional requirements

S.1.1 (M) The ERMS *must* support the management of the storage of digital objects and metadata, potentially across a range of removable and attached media[7].

S.1.2 (M) The ERMS *must* support system or administrator replication of digital objects and metadata between different storage media.

S.1.3 (M / HD[8]) The ERMS *must / should* maintain at least *two* copies of each digital object and its metadata.

S.1.4 (M) The ERMS *must* routinely back-up the system indices, digital object *and* record metadata.

6 See also section on security requirements.
7 This generic high level requirement is developed further by a number of other requirements throughout the first half of this specification, e.g. scalability, back-up / recovery requirements
8 This is a generic requirement about *physical* rather than logical storage. If a system is to preserve, physical replication is one of the principle means of supporting this. In many cases, it will need to be treated as mandatory. It may be possible to count live copies on attached media and near or offline copies on removable media towards meeting this requirement but the risk of such an approach must be thoroughly assessed and kept under review. See Requirements: S.1.4, S.1.6 – S.1.10, S.1.21 – S.1.33 and S.2.9 – S.2.14

S.1.5 (M) The ERMS *must* have sufficient capacity in terms of the digital object store and the metadatabase to accommodate the creation rate of new records (and possibly also documents if an integrated EDRMS), use of the records and the long-term retention of new ones until their disposal in accordance with the disposal schedules applied.

S.1.6 (D) The ERMS *may* have the capability to manage digital objects and metadata in multiple stores as a single integrated logical solution to address scaling issues.

S.1.7 (D) The ERMS *may* employ, for its digital object and metadata store[s], many or a mixed combination of storage devices [e.g. arrays of disk storage or tape libraries], possibly across different sites. These should be capable of seamless operation as an integrated solution, ideally with the same user interface.

S.1.8 (M, *Cond.*) If requirement S.1.7 is met, the ERMS *must* address metadatabase scaling issues in similar ways as in the preceding three requirements.

S.1.9 (D, *Cond.*) If requirement S.1.6 is met, the ERMS *may* support distributed digital object and metadata stores across different sites on a secure Wide Area Network.

S.1.10 (HD) If requirements S.1.7, S.1.8 & S.1.9 above are met, the ERMS *should* support the intelligent management of storage[9], e.g. as to performance, availability and robustness.

S.1.11 (D) The ERMS *may* support the use of the internal metadatabase to hold pointers to records held elsewhere ["elsewhere" in this requirement means places where record content is held on storage not permanently connected to the main digital object store of the ERMS; nearline or offline] and the management of their lifecycle[10].

S.1.12 (M, *Cond.*) Where requirement S.1.11 is met by the deployment of off- or near-line storage, the ERMS *must* support the definition and exchange of standard protocol messaging between the distributed solutions and implementation of the messages as system commands. These protocol messages *must* include search protocols ["federating search"] and the return of search results in the form of abstracted metadata profiles.

S.1.13 (HD) The ERMS *should* provide facilities to select and browse from abstracted metadata profiles of individual records, folders and classes returned in search results to the content and metadata of related records, folders and classes in the fileplan or classification scheme (e.g. either by invoking the ERMS fileplan view or navigating / viewing through stylesheets and identifiers).

S.1.14 (HD, *Cond.*) Where requirement S.1.13 is met, the protocol messages *should* include inherited and specific management metadata such as access control and disposal metadata.

S.1.15 (HD) The ERMS *should* support the retrieval of record *content* from the offline or nearline solution returning an abstracted metadata profile in a single integrated process.

9 Such routines *may* include, but need not, caching of frequently or recently used digital objects, automated or semi-automated redistribution of digital objects and metadata to reduce performance loading, etc.

10 This requirement and the following requirement[s] that are dependent on it, imply implementing an enhanced version of the *Optional module* B.3 on hybrid records published by TNA for the current records management environment, but for born-electronic records held elsewhere. See: http://www.nationalarchives.gov.uk/electronicrecords/reqs2002/pdf/requirementsfinal.pdf, pp. 58-60. The use of metadata profiles within the classification scheme to identify and manage physical objects ought to be considered as a means of managing removable media such as back-ups.

Non functional requirements

S.1.16 (M) The ERMS *must* be accompanied by information from the supplier[s] on the limitations of the system imposed by its design, hardware, software or other constraints. "Limitations" may relate to numbers of digital objects, size of metadatabase, numbers of records, etc.

S.1.17 (M, Cond.) Where requirement S.1.10 is met, the ERMS *should* be accompanied by information on the criteria use by the intelligent storage management, and a capability for the administrator to override them.

Reporting requirements

S.1.18 (M) The ERMS *must* support the definition of standard reports to provide scalability, and other storage management. Examples of such reports should include:

- Creation rates, i.e. expressed in terms of cumulative file sizes, metadata creation rate
- Disposal rates, i.e. of digital objects and metadata profiles and stubs
- Unavailable storage space within the solution
- Current capacity threshold for the digital object store[s] and metadatabase
- Predictive reporting of likely future storage requirements (e.g. xx GB) to assist media and hardware planning (i.e. algorithmically based on the above management information).

S.1.19 (M) The ERMS *must* support the proactive production of standard reports [i.e. optionally without requiring the report[s] to be initiated by an administrator] as above as a configuration option.

S.1.20 (M) The ERMS *must* alert that capacity thresholds are being approached.

Media monitoring and refreshing rationale

It is generally preferable for the monitoring of both content and the media it is stored on for the storage to be "nearline"[11]. Sometimes, there may be compelling reasons [e.g. scalability, performance] for storing content offline, where protocols and standards exist for loading the content swiftly but involving specific human or automated intervention not necessarily involved in the everyday functioning of the system.

There may be other reasons, such as security, where *airgapping* is employed to avoid the risks imposed by a permanent physical connection between system. In this scenario, there must be some way of continuing to monitor the state of the media and the records stored on it (as well as to retrieve them: see previous section). Examples of how this might be done are included in the final group of functional requirements but are scenario-specific.

11 Online storage would also support the same requirements, including background routines periodically checking integrity, etc., but would impose a type of solution.

Functional requirements

S.1.21 (M) The ERMS *must* monitor the performance and degradation of all attached media. In this requirement, "attached" means permanently *or temporarily* attached to the ERMS.

S.1.22 (M) The ERMS *must* facilitate the refreshing of storage media for digital objects and metadata.

S.1.23 (M) (M) The ERMS *must* not by its design or performance imperil the longevity of the storage media[12].

S.1.24 The ERMS *must* use storage media with an approved, explicit and proven lifespan[13]. ["Proven lifespan" in this context refers to industry standards and / or manufacturer's recommendation[14]].

S.1.25 (M) The ERMS *must* permit the entry of media longevity parameters based on industry standards and / or manufacturer's recommendation.

S.1.26 (M) The ERMS *must* alert the Administrator when the industry or manufacturer's estimated media lifespan is approaching.

S.1.27 (M) The ERMS *must* integrate with external software applications or checking procedures for verifying the physical integrity of storage media.[15]

S.1.28 (M) The ERMS *must* provide facilities for the replacement of obsolescent or failing media. Mandatory "facilities" in this requirement include:

- The ability to copy the digital objects and/or metadata to alternative media;
- The ability to verify the copying process through a bit-level or checksum comparison between the source and target versions.

S.1.29 (D) The ERMS *may* provide for the automatic commissioning of new storage media (e.g. adding a disk or tape to those already operating in the solution) and the implementation of the previous requirement.

Integrity Rationale

Integrity of the digital objects is critical to their continued use over time. In this context, integrity refers to the absence of errors in the data, ie. every digital bit that makes up the object has the correct value. Bit corruption can occur through errors in copying or disk read actions, through electromagnetic affects, or through the transition of stray gamma particles through the storage medium (highly unlikely in practice). The commonest and most practical method of checking integrity is through the use of checksums. A number of checksum algorithms exist, of varying degrees of complexity and likelihood of 'collisions' (the possibility of two different data objects producing the same checksum). When choosing a checksum algorithm, it is recommended that one of the SHA (secure hash algorithm) checksums be used as these are less susceptible to collisions. Examples of such SHA checksums are MD5, SHA-1, SHA-256 etc.

12 e.g. by imposing environmental conditions inimical to its performance or longevity through its other operations.
13 The requirement is worded as though the media is part of the ERMS. Easily removable media may be deemed not to be part of the ERMS core system, but will still need to meet this requirement.
14 Hard disk storage cannot reasonably be expected to comply with this.
15 In the case of optical media it is not appropriate to rely on CD or DVD drives, since the error checking function built-into these technologies will result in notification of physical media errors only when they have become degraded to such an extent that data recovery is impossible. Drive manufacturers supply disc checking and verification utilities for carrying out such checks.

S.1.30 (M) The ERMS *must* regularly check the integrity of digital objects and metadata.

[(a) In this requirement, "regularly" should be interpreted as:

- For live online storage, an ongoing system background routine which may be performed at times of low user demand and/or other regular intervals;
- For off- or nearline storage at an interval shorter than the mean time before failure.

(b) In this requirement, "integrity" should be interpreted as absence of corruption of the digital object[s] by any means including unauthorised tampering.

(c) In this requirement, the sort of "checks" envisaged include against inventory databases, checksums of individual items and aggregates, schemas and comparison with other copies.

(d) In this requirement, near- or offline storage monitoring should be instituted as a more manual routine by a report from the ERMS prompting the connection and checking of the removable media on a rotational basis].

S.1.31 (M) If the integrity check detects an error, the ERMS *must* support:

- Isolation of the storage volume[s] affected and reservation for restoration;
- Resolution of detected errors discovered by comparison of distributed replicating storage;
- If the uncorrupted copies are available to it, restoration to the appropriate location[s]; and / or
- Alerting the administrator to the replacement of the copies or the need for the replacement of the copies [the administrator must be given the opportunity to correct an automatically triggered replacement].
- Updating the metadata to record the recovery action with the action taken and the administrator's authority.

S1.32 (M, *Cond.*) If a new copy of an object is produced in association with requirement S.1.31, in response to an error detection, the system must update the new object's metadata to reflect its creation; and must update any objects which referred to the earlier manifestation.

S.1.33 (M, *Cond.*) Where a new copy of a digital object is produced by the functionality described in requirement S.1.32, the ERMS *must* update the metadata including history elements and relational metadata [linking different manifestations of the objects].

Reporting requirements

S.1.34 (M) The ERMS *must* allow the capture of defined refreshing or [media] migration timescales based on manufacturer's and industry recommendations [see non-functional requirements below].

S.1.35 (M) The ERMS *must* allow the setting of a default alert timescale for the timing of media refreshing alerts [e.g. 6 months or a year prior to the action required].

S.1.36 (M) The ERMS *must* provide facilities for the definition of standard reports on media obsolescence. Such standard reports may include:

- Particular media affected by obsolescence owing to their reaching the end of their working life (see non-functional requirements below);

- Particular media monitored by the ERMS and performance is found to be deteriorating;
- Overdue media refreshing (i.e. where the reports have either not been run or not executed);
- Consolidation of any of the above types of report.

S.1.37 (M) The ERMS *must* support the proactive production of standard reports as above as a configuration option.

Non-functional requirements

S.1.38 (M) Public authorities *must* take note of media manufacturers' recommendations on the longevity of storage media[16].

16 Further advice and Public Records Sector guidance will be made available on The National Archives' website.

S.2 Security

Rationale

One of the principal drivers for the production of these requirements is the need in central government for digital records selected for archival preservation to be maintained by departments until their security markings have been downgraded sufficiently to permit their transfer to The National Archives or other repository. In some cases this will be periods of up to 30 years, or even longer. Records required for long periods for purely business purposes may also have security markings attached and / or contain sensitive personal data that requires additional protection.

The requirements here break into two distinct categories: functionality likely to be required of the ERMS itself and security requirements in its surrounding environment. The former introduces certain innovations in functional access control at export. In the latter case, only a brief outline of the likely needs can be covered here. Departments should seek the guidance of their ICT department, their own security advisers, the *Manual of Protective Security* and appropriate security agencies such as CESG. In addition, there are the Security environment, back-up and disaster recovery requirements in TNA 2002, cited in Annex 2.

Some information security procedures are provided in the media management parts of section 1.

Access to the records can only be assured by the ERMS whilst they are within its controlled environment. Additional requirements addressing the behaviour at export is required to meet the main security scenarios likely to be encountered.

Functional requirements

S.2.1 (M) The production code of the ERMS solution *must* meet the security requirements of any relevant security authority.

S.2.2 (M) The ERMS *must* support the definition of a discrete record type to act as a placeholder in export.

S.2.3 (M) The exported placeholder record-type *should* be capable of having no digital object[s] attached and *should* be configurable to contain only a subset of the available metadata fields and an additional identifier to be used for ultimate reconciliation of the full metadata profile with [all] the relevant digital object[s].

S.2.4 (M) The ERMS *must* permit the setting of a maximum export security marking, to be matched with the clearance level of the export destination. For example, if the latter is infrastructure cleared to *Restricted*, nothing bearing the security markings *Confidential*, *Secret* or *Top Secret* can be exported from the ERMS.

S.2.5 (M) The ERMS *must* withhold any records (digital objects and their metadata) or superior aggregations (i.e. folder parts, folders or even classes) contained in larger aggregations of records scheduled to be exported (i.e. classes, folders or folder parts) and substitute [a] placeholder record type instance[s] for them at export[17].

17 i.e. to avoid the export of sensitive material to infrastructure without the requisite clearance, see Paras. 9.2.3 and 9.2.3.

S.2.6 (HD) On alerting the administrator or records manager role to the existence of records unexportable under the previous requirement, the ERMS *should* prompt the consideration of the production of a redacted instance of the digital object[s] and metadata involved that is capable of being exported within the aggregation [e.g. by the presentation of a report on withheld records].

S.2.7 (M) The ERMS *must* enforce the retention of metadata stubs for all exported records to retain structural integrity and context for the retained records until such time as they can be finally disposed.

S.2.8 (M) On declassification, the ERMS *must* support the separate export of the full metadata and withheld digital objects and the return of the disposal routines for the other metadata stubs to those normally implemented in the ERMS.

[Note: See also subsection on data entity models in interoperability section below]

Non-functional requirements[18]

[Note: These requirements are not necessarily concerned with the system functionality of the ERMS itself but may imply system capabilities more widely].

S.2.9 (M) The ERMS and all its integrations *must* form part of a comprehensive assessment of information security management, such as contained in BS-ISO7799, other continuity and disaster recovery planning (i.e. 'around' back-up), procedures, HR issues, audit, etc.

S.2.10 (M) The ERMS *must* facilitate the maintenance of at least two further copies[19] of each digital object in secure off-site storage.

S.2.11 (M) The ERMS *must* facilitate the maintenance of at least two back-ups of the index data (i.e. metadata *and* audit trails) in secure off-site storage.

S.2.12 (HD) The back-up procedures *should* establish a regime to rotate off-site copies to ensure consistency in checking to the level described in requirements S.1.30 and S.1.31 above. The interval between such checks should not be less than that advocated in requirements S.1.24 and S.1.25 above.

S.2.13 (M) The backup copies *must* be stored, and moved between the system and the separate site, so as to protect them to the highest security classification applied to the records stored on the back-up copies and comply with data protection legislation[20].

S.2.14 (D) The backup copies *may* be kept on a site at least ***x*** kilometres from the main system. In this requirement, "***x***" shall be a distance identified by a full assessment of the risk [impact and likelihood] that back up media might be affected by the same disaster.

18 Back-up is handled as a "non-functional" issue in the remainder of this section as the functionality will almost certainly lie outside the ERMS. Some requirements relevant to back up routines but more generic are in sections 8.1 and the previous group of functional requirements in this section: these functional aspects relate to the creation of copies and their reloading, metadata requirements and *not preventing* the other operations outside the ERMS.

19 i.e. further to the replication / back-up requirements in section S.1, but see *fn* to Requirement S.1.3.

20 The intended meaning of this requirement is as a *physical* protective security measure.

S.3 Interoperability / openness

Rationale: promoting independence of records from the current platform

Records can very often be required for longer periods than the system currently being used to create and manage them. Their value as corporate assets may be very many times in excess of the value of the information technology they are maintained in.

In these *Requirements*, there is no presumption that the records (digital objects and metadata) should have any significant dependency on the system used to create and [initially] to store them. As a result, most of the requirements are articulated in terms of the *outcomes* necessary to support the sustaining of the records rather than actual system *design* criteria. Typically, any good ERMS ought to be able to store any digital object in its repository and capture metadata about it. More demanding is the requirement to export objects and metadata out of the system in close association and import them coherently on the receiving platform with the relevant metadata mapped between equivalent fields..

Semantic and syntactic interoperability through metadata standards and schemas

National Archives' guidance in the current ERM environment has done a great deal, if followed, to reduce the likely level of dependency between records and the infrastructure used to create them. This is mainly a question of standardising descriptive metadata and must be taken to a new level to meet the present needs. Ultimately, interoperability at both the semantic [*meaning*] and syntactic [*textual encoding*] levels will be required to achieve this. The National Archives will continue to promote this level of interoperability to the Cabinet Office as an essential consequence of the eGovernment Interoperability Framework requirement for records required for extended periods. It is essential to support archival transfer, long-term sustainability, data sharing and platform migrations undertaken in response to changes in the machinery of government.

The requirements of ISO 15489 and 23081 in this respect are comprehensive and any solution aiming at best practice requires careful design and configuration. To maximise the evidential value of the records, all significant processes carried out to them should be evidenced in some way[21]. Metadata about them is also required to support their preservation. The construction of a plan or *schema* gives some structure to the multiplicity of metadata fields required in records management. This is an essential metadata management device.

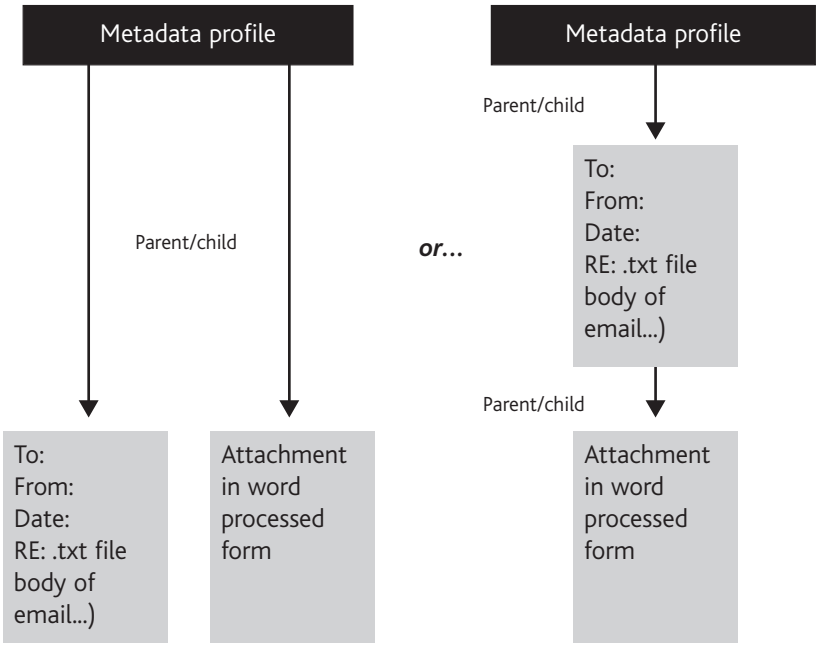
Role of standardised data entity models

In this respect, it is particularly important that standardised data entity models can be observed (or computed) at system export and import interfaces. These requirements publish the results of important work on this done under the auspices of The National Archives' digital preservation programme to promote this standardisation. For example, a record comprising a single digital component and its metadata is a simple matter; the linkage specified in TNA 2002 requirement A.2.49 (reproduced below) is clearcut: an identifier in the metadata points to the digital object currently manifesting the record's

21 Digital records being prone to alteration unless in a highly controlled environment, meaning that a full audit of operations carried out on them needs to be maintained and ideally in a form independent of the management system[s] and capable of accompanying them throughout their life, i.e. the record metadata. This is a business / evidentiary rather than an archival requirement.

content. Even an email with an attachment or a word processing file with an embedded spreadsheet chart presents a more complex scenario and questions about whether the "record" is more accurately modelled as a metadata profile with several 'child' objects or a prime object and its metadata with a secondary object (possibly with its own metadata). The email scenario is modelled very roughly in the following diagram:

Alternative representations of data entity relationships: same email with word processed attachment



Even in this very basic, everyday example it is clear that achieving maximum interoperability in how systems handle these issues requires standardisation, preferably from the beginning. Very similar issues arise with managing multiple manifestations of the same (or similar) record content, e.g. where a rendition or redaction is present. The principles behind this will also apply to far more complex scenarios where the records are best understood as comprising multiple digital objects or *record components*. In many computing environments, the relationships between them may be highly proprietary and difficult to preserve.

In addition, it is necessary to settle what approach is to be taken to compound file types, which nest more than one digital object inside another so that there is apparently one digital object present, with [an]other[s] possibly hidden until it is opened. The most common types of these are complex compound email files such as *.msg* and *.vmbx* and web page formats such as *.mht*. These present the additional challenges in that the specification is often a proprietary one[22]: the danger is that once the specific format is not supported by either the creating software or the operating system, it will not be usable. (There are further interoperability issues to be considered with file formats; these are discussed in the migration section).

22 *.msg* and *.mht* are formats owned by Microsoft Corporation. Many ERMS solution providers are apparently aware of this issue and some proprietary solutions render emails to alternative formats on capture to hedge this risk at migration.

The National Archives will represent the implications of this work to the eGovernment Unit of the Cabinet Office for later inclusion in subsequent editions of the Government Data Standards Catalogue ["GDSC"].

Functional requirements

S.3.1 (M) The ERMS *must* support descriptive standards mandated by eGovernment initiatives such as eGIF and eGMS including the specialised applications of these developed by The National Archives in collaboration with the eGovernment Unit to standardise structured ingest [import] and export protocols[23].

S.3.2 (M, *Cond.*) Where the support of requirement S.3.1 is impracticable without a degree of transformation (e.g. of XML metadata using XSLT), the infrastructure which supports the ERMS *must* facilitate the maintenance of this technical documentation. This may be done either:

- by their declaration as records and subsequent maintenance within the ERMS;
- download of essential identification information of DTDs, schemas and / or stylesheet perpetually preserved elsewhere[24]; or
- by real-time linkage to a maintained schema and / or stylesheet registry according to any protocol published by a relevant authority such as the eGovernment Unit of the Cabinet Office or the Local Government eStandards Body.

S.3.3 (M) The ERMS *must* present data entity models implemented in the relational metadata of digital entities in the forms specified for common record types in the diagrams in Annex 3 and other such requirements as included in future National Archives / Cabinet Office metadata standards.

S.3.4 (HD) The ERMS *should* provide facilities to capture or subsequently render digital objects to formats advised as suitable for long term retention / sustainability by either The National Archives direct or in collaboration with the Cabinet Office eGovernment Unit in the GDSC.

S.3.5 (D) The ERMS *may* support the import of digital objects and their metadata where there are multiple manifestations (renditions, redactions etc.) of the same record content[25].

23 This facilitates the handling of the records or their metadata by other platforms. There are two main scenarios for this: those systems needing to receive a cogent export of the entire content [such as successor systems and archival systems] or those requiring a defined subset of the metadata or content only [such as portals or eGovernment line-of-business systems].

24 Preservation of the *DTDs* / document schemas and stylesheets may need to be supplemented by the logical schemas of the business workflow if logical and technical characterization to support active preservation is to be achieved. See Active preservation.

25 See section S.3 and Annex 3: Generic data entity models for common record types to be observed or computed for this requirement to be viable.

S.4 Active preservation

Rationale

'Active preservation' involves taking steps to ensure access to digital objects and metadata can be maintained across long periods of time by planning and taking action to offset technical obsolescence as it might affect the objects and metadata to be preserved. How this is approached in the preservation strategy is likely to be affected by the characteristics of the collection being preserved and the time periods involved.

Migration is currently the preferred means of preserving access to digital records over extended periods. Other options are the use of emulation, or the adoption of open / persistent formats strategies. Emulation involves supporting the historic behaviour of computer hardware and operating systems. Obviously, this can be enormously expensive although, in theory, it minimises the amount of intervention that needs to be made in the bitstreams of the objects themselves. Open format strategies rely on the availability of open[26] and either long-term or perpetually-supported formats. Such strategies are variants of the migration approach. Generic multi-format viewers may also be of use as a 'cut-down' emulation strategy, particularly where the affected records are not required permanently and/or the existing formats are relatively static.

In some circumstances, the most active degree of preservation may be the best approach to take for records required for long periods. This involves taking decisive action at [or before] the creation stage to maximise the preservability of the records: e.g. by creating, capturing or promptly rendering record content to open and / or standard formats[27]. There will be a cost to this approach in terms of flexibility and overhead at creation stage so it will not be suitable in all circumstances. [The above interoperability requirements will, if followed, promote the preservation of descriptive metadata accompanying the digital objects by observing an open and readily transformable XML standard as published by the World Wide Web Consortium (W3C)].

The remainder of the approach outlined here is concerned with migration of digital objects in *pre-existing* formats. The strategy to be adopted also rather depends on exactly *what* is to be sustained, for how long and for what purpose and not solely the current formats the record content is encoded in:

For example, the essential characteristics of a spreadsheet may be the historic values it calculated at the time of record capture. According to whether the 'point' of the record is to explain the complete reasoning of a decision or to show the statistical data underlying it, the full functionality of the formulae and the creating application may or may not be required. Some records may be required for a time period equating to a few generations of software but neither for many decades nor permanently. Such records, if static and relatively simple, may be best preserved by bit-level preservation if a risk assessment shows generic viewers can be used to access the content in cogent form.

Records management and archival science both accept that migration of the underlying digital object[s] by a trusted controlled process does not *of itself* destroy the authenticity of the record (see Volume 1 of the *Generic requirements for sustainability: Defining the characteristics for authentic records* and the *Custodial issues and archival mappings* section of these requirements).

The requirements in this section are designed to ensure that active preservation decisions can be taken based on the most robust criteria and at the most appropriate

26 See Annex 1: Terminology entries on open source, open standards and open formats.
27 Parts of such an approach are already contained in the TNA 2002 *Functional requirements* designed for the current records environment, see particularly the rendition requirements: A.2.12 and A4.57.

stage in line with a consistent approach. They are organised in five main sections:

Characterisation rationale

The first set are concerned with characterisation of the significant technical and logical properties of the digital objects. This involves the evaluation of the objects'[28] method of conveying the trace of the business activities they record. This is partly a question of the actual format used and partly how encoding the content in that way [i.e. in its existing format] forms part of the record.

This needs to be done at a level appropriate to the expenditure of resources and value of the records: broad categorising of the records as outlined in **Volume 4** of the *Generic requirements* and comparing the distribution of objects amongst these categories along with broad mapping to the business classification scheme and its disposal criteria. This implies a developed means of obtaining management information on the digital objects and the records they form part of (see Reporting requirements). It begins with validating any information or metadata that may already exist on the format of the digital objects.

Some properties not present in the record metadata may be capable of automated extraction at this point. The National Archives will develop, make available and support a range of tools to assist with the identification and extraction operations.

Preservation planning rationale

Preservation planning involves developing the practical means to implement higher-level preservation policies and strategies. It involves assembling the picture gained from the Characterisation information and preparing to take active preservation decisions based on an evaluation of risk and the impact of the course of action indicated by the various indicators. This has four main stages: assessing risk based on the format of the objects and their known properties, a technology watch function[29] to track the obsolescence of file formats and other technical components in terms of compatibility and support and observe any impact on the risk criteria, the assessment of the overall impact of the migration plan on the collection and finally the generation of a preservation plan. The resulting Preservation Plan is likely to consist of a plausible migration pathway applying to many [or all] of the instances of a particular file format held in the digital object repository[-ies] and it requires testing and validation as a means of meeting the requirements and expectations of the strategy before it is finally executed.

Migration execution rationale

Migration of the digital objects contained in records may need to be done to preserve access to the records but needs to be done in accordance with preservation policies and strategies and under appropriate levels of security and control of the ERMS. Central to these latter is the maintenance of the binding between the record metadata and the digital object[s]: if this is broken there is no record present. Meanwhile, the approach to preservation planning ought to mean that considerations such as 'look and feel' and other non-archival views on authenticity have also been accommodated. Particular attention should be paid to relevant sector-specific enactments.

[See also generic data entity models and metadata requirements below].

28 This section assumes that the demands of the previous sections on accrual and openness of metadata have been met and can therefore be kept linked with the objects.

29 The National Archives is developing a technology watch called 'PRONOM' to provide such a service to UK public authorities over the internet. See http://www.nationalarchives.gov.uk/pronom/ It has at its heart a file format registry.

Migration validation rationale

Following migration of the digital objects, some validation is required to check that it has been successful. This obviously cannot involve a 100% visual check of all the objects affected, but must be effective, especially if the obsolescence risk is imminent and/or the previous formats of the digital objects are not being retained[30]. This is likely to involve some selective manual checking [especially where migration has been problematic or complex] and some automated using the characterisation information gathered earlier. The latter automated checking should ascertain that each migrated object expected in fact exists and is a valid instance of its expected format. It is also important to ensure that unmigrated objects have not been changed or corrupted by the migration process, by carrying out an integrity check using previously calculated checksums.

Reporting rationale

Many types of management information are implied by the above requirements.

Note: A collection of material with a very restricted range of formats and use of formats may permit the stages of technology watch, characterisation, preservation planning, execution and validation to be greatly simplified. For example, a collection consisting of text and image records with little additional functionality such as embedding, macros employed from complex digital object types could credibly be managed using standard formats[31]. The requirement first to assess the collection and the risk of the proposed approach remains.

Functional requirements

S.4.1 (M) The ERMS *must* support the migration of digital objects comprising the content of records *prior* to the technology they are encoded in becoming completely obsolete and access to the records becoming possible only through expensive techniques of digital archaeology.

S.4.2 (M) The ERMS *must* allow the recording of a pointer to an external escrow service (to which a subscription had been made) to link particular records or record component formats to technical documentation not currently disclosed to the organisation operating the ERMS[32] or the Technology watch service.

Characterisation, Validation

S.4.3 (M) The ERMS *must* support the identification of the file formats of digital objects it manages.

S.4.4 (M) The ERMS *must* support the identification of encrypted digital objects in its repository and alert the administrator to the need to determine appropriate action.

S.4.5 (HD) Where it identifies encrypted information, the ERMS *should* support the identification of the relevant encryption algorhythm.

30 One advantage of the approach taken to preservation validation is that it provides the characterisation of the target format as part of the process so that this information can be exploited in the next migration cycle.
31 e.g. **ASCII**, **.tiff**, **PDF-A** [ISO19005/1].
32 e.g. proprietary file format specifications may be held in escrow by a trusted third party whilst the IPR owner is still asserting its moral rights for commercial purposes.

S.4.6 (D) If requirement S.4.5 is supported, the ERMS *may* offer options for the creation of an unencrypted manifestation of the digital object(s) affected[33].

S.4.7 (M) The ERMS *must* support the identification of the specific version of the file formats of digital objects it manages[34].

S.4.8 (M) The ERMS *must* be capable of capturing preservation metadata about file formats it identifies to the relevant field indicated in the current version of the digital records management specialisation of the eGovernment Metadata Standard.

S.4.9 (M) The ERMS *must* be capable, where required, of querying an authoritative file format registry [e.g. PRONOM[35]] and assign the relevant PRONOM unique identifier to the relevant metadata field for each digital object.

S.4.10 (M) The ERMS *must* be capable of extracting additional metadata from the relevant file property fields of digital objects it manages.

S.4.11 (M) The ERMS *must* extract or integrate with extraction tools to compute additional preservation metadata about the instance of the digital object[s] and map it to the relevant field indicated in the current version of the digital records management specialisation of the eGovernment Metadata Standard.

Technology watch

S.4.12 (M) The ERMS *must* support the entry and storage of technical dependency data about file formats and making it available to the reporting functionality described below to support preservation planning for the digital objects it manages[36] in its repository.

S.4.13 (D) The ERMS *may* integrate with an IT systems portfolio management application to compare current creation formats with ERMS repository formats and compile additional statistical reports.

S.4.14 (M) The ERMS *must* respond to any change in the status of stored technical dependencies to issue an alert to the administrator of any enhanced risk of obsolescence applying to digital objects under its management.

Impact assessment, Migration plan generation

S.4.15 (M) The ERMS *must* provide facilities for the compilation of statistical reports to support preservation decision-making based on appropriate criteria. In this requirement, "appropriate criteria" include:

33 the handling of such alternative manifestations produced in response to this requirement must follow the generic data entity modelling and requirements for other forms of migration that follow

34 e.g. the generic file format might be WORD and the specific WORD 6. It is recommended Departments and agencies use the identification tool DROID available from: http://www.nationalarchives.gov.uk/aboutapps/pronom/tools.htm . The PRONOM unique identifier system referenced in requirement S.4.9 expresses both the *generic* and *specific* file format.

35 See: http://www.nationalarchives.gov.uk/pronom/ PRONOM is recognized as an encoding scheme for the Format element of the eGMS.

36 In particular, PRONOM is a major part of a comprehensive technology watch programme being implemented at The National Archives to support its digital preservation activities. PRONOM is also referenced more specifically in the functional requirements section below. In this context, its main contribution is the provision of information to track the technical dependencies of current file formats and support the taking of migration decisions *before* formats are completely obsolete. Refer to Volume 3 of the *Generic requirements for sustaining electronic information over time: technical requirements* and http://www.pronom.gov.uk . Ideally, a download copy of extracts from the relevant PRONOM database table might be queried by the ERMS.

- Number of instances of file format before and after proposed migration execution;
- Number of instances of specific format version before and after proposed migration execution;
- Number of records possessing the above characteristics:
  - Forming part of compound records;
  - In particular areas of the classification scheme;
  - With particular disposal rules applied.
- Both *before* and *after* proposed migration execution.

Migration

S.4.16 (M) The ERMS *must* provide a report of migration activity to be carried out in order of urgency based on obsolescence alerts it has received [or computed].

S.4.17 (M) The ERMS *must* generate checksums for all digital objects due for migration according to the migration plan and associate them *temporarily* with the corresponding digital objects.

S.4.18 (M) The ERMS *must* allow the selection of a random sample of digital objects not due for migration, and compute checksums for the objects in the sample.

S.4.19 (M) The ERMS *must* integrate with a range of software tools to migrate digital objects from their source to their destination format. Such tools will include:

- Standard software applications with backward [or forward] compatibility;
- File format conversion tools;
- Any bespoke utilities [transformation utilities] for the same purpose.

S.4.20 (M) The ERMS *must* permit the staged execution of the highlighted migration actions according to decisions taken by the administrator. In this requirement, "staged" may mean:

- According to the area of the classification scheme the records are attached to;
- According to the digital object main or sub-format type;
- According to the urgency of migration indicated by the report.

S.4.21 (M) The ERMS *must* require the administrator to confirm all single or batched migration actions before executing them.

S.4.22 (M) On execution, the ERMS *must* link the newly migrated manifestation of the digital object[s] to the relevant record metadata using a new instance of relational metadata in accordance with the digital records management specialisation of the eGovernment Metadata Standard and Annex 3 of these requirements.

S.4.23 (M) The ERMS *must* provide a report on the execution of the migration action[s], or such information must be obtained from any separate migration tool prior to the migration being finalised [e.g. by disposal of the previous manifestations]. The report must specify:

- Number of objects affected
- Characteristics of object affected
- Number of conversion errors and / or corruptions identified during the migration process[37]

37 Identified by integrity check failures; reported copy failures, etc.

S.4.24 (D) The ERMS *may* prompt the viewing of selected portions of the system audit trail relating to the executed migration.

S.4.25 (HD) In the case of migration actions reported as having been unsuccessful, the ERMS *should* present options for the resolution of the errors.

S.4.26 (M) The ERMS *must* support the retention of previous manifestations of digital objects forming migrated record components.

S.4.27 (HD) The ERMS *should* provide for the global retention or disposal of the previous manifestation objects, as a configuration option.

S.4.28 (M, *Cond.*) Where the configuration option in requirement S.4.27 is set to dispose of the previous manifestations, the ERMS *must* prompt for a further two-staged confirmation before disposal is finally executed[38].

S.4.29 (HD) The ERMS *should* present options for the disposal of previous manifestations that have been the subject of previous migration[39], e,g, dispose immediately, after 3 months, or 5 years.

Validation

S.4.30 (M) The ERMS *must* validate the migrated objects as valid instances of the target format (i.e. *technical validation* using file format identification tools previously mentioned).

S.4.31 (M) The ERMS *must* compare the temporary checksums made under requirement S.4.17 above for the unmigrated instances of the digital objects that have undergone migration, with, newly generated checksums for the unmigrated instances using the same algorithm.

S.4.32 M) The ERMS *must* apply the checksums generated in requirement S.4.18 above to compare the sample set of digital objects not undergoing migration[40].

Non-functional Requirements

S.4.33 (HD) Migrated versions of digital objects *should* be confirmed as valid objects by manual checking procedures. Such procedures *may* include:

- Opening and using a random sample of the migrated objects
- File format checking of a random sample using external tools such as JHove and DROID.

Reporting

S.4.34 (M) The ERMS *must* prompt for the running of a full migration validation report [or automatically produce such a report] to facilitate the carrying out of additional manual spot checks by viewing migrated digital objects.

S.4.35 (M) The migration validation report *must* offer refinement parameters such as:

38 i.e. to allow for the evaluation of the success of the migration.
39 Whether immediately or some period before.
40 A suitable sample is to be identified according to the preservation risk assessment.

- A certain proportion of migrated digital objects at random;
- Migrated digital objects weighted according to the area of the classification scheme, [see risk assessment criteria previously applied at requirements: S.4.15, S.4.16 and S.4.23].

S.4.36 (M) The ERMS *must* support navigation from the migration validation report to the relevant digital objects in a single integrated process [i.e. without having to repeat the report].

S.4.37 (M) The ERMS *must* record the retrieval of the migrated digital objects itemised in the migration validation report and save the record of these actions as technical documentation.

S.4.38 (M) The ERMS *must* support the definition of reports to check the statistical outcome of migration[41].

Preservation *process* metadata requirements [at migration]

S.4.39 (M) The ERMS *must* maintain a navigable relational link between different manifestations of the same record content.

S.4.40 (M, *Cond.*) Where requirement S.4.27 is met, the ERMS *must* record the pre-existence of disposed record manifestations by retaining a minimum set of metadata. In this requirement, this "minimum set of metadata" must specify:

- The system identifier and record reference of the disposed manifestation;
- The creation and disposal date of the disposed manifestation;
- The format of the disposed manifestation;
- The exact time of the disposal; and
- The identity [user details] of the administrator authorising the disposal.

S.4.41 (M) The ERMS *must* record the details of the migration actions executed on record components in accordance with the electronic records management specialisation of the eGovernment Metadata Standard[42]. (Identical 'batched operations' carried out on large numbers of digital objects may be recorded in technical documentation subject to the meeting of requirements S.4.39 and S.4.40 above and other record integrity requirements from TNA 2002).

41 e.g. the separation of complex digital objects in the source format into more than one in the target format.
42 These sub-elements will be developed in association with The National Archives' digital preservation programme.

S.5 Ingest[43]

Rationale

It may sometimes be necessary to import records that have *not* been managed within a controlled environment capable of associating significant quantities of descriptive metadata to record their provenance, relationships, technical characteristics, subject etc. Such records are sometimes [and despite their compromised nature] relied upon by the business as the records of their activities. This is often as a last resort because nothing better is available. Whilst it is not possible to recreate much of the metadata that might have evidenced these things from an earlier stage, some metadata can usefully be associated with them to assist in managing and using them for the remainder of their life and evidencing their management regime from ingest onwards.

Functional requirements

S.5.1 (M) The ERMS *must* be capable of accepting copies of digital records in "as received" form, that is unprocessed save for the allocation of a simple unique temporary component identifier (e.g. a sequential number).

S.5.2 (M) The ERMS *must* notify of any encrypted material in an ingest and alert the administrator prior to ingesting it, pausing the process to allow resolution of the position.

S.5.3 (M) The ERMS *must* be capable of enhancing fragmentary or compromised metadata accompanying ingested digital objects by integrating with a variety of extraction tools, placing the newly-extracted metadata in the relevant metadata field in the ERMS metadatabase.

S.5.4 (M) The ERMS *must* be capable of integration with tools provided for file format recognition, such as DROID, to identify ingested formats.

S.5.5 (M) The ERMS *must* have the capability of applying a PRONOM unique identifier to the relevant metadata field for the recording of file format of each ingested digital object. ["Relevant metadata field" in this context means as identified in the current records management specialisation of the eGMS].

Non-functional requirements

S.5.6 (M) Processes *must* be in place to ensure that all digital objects scheduled for ingest have been imported by the ERMS.

43 This section deals with a specific scenario where records are taken into the system from an unmanaged environment and with fragmentary metadata. See Annex 1: Terminology definitions and explanatory rationale of this and the concept of bulk Structured Ingest (mapping to Import in TNA 2002) to clarify the difference between these concepts.

S.6 Reporting

Rationale

As stated in the Scalability requirements above, many of the present requirements imply a greater degree of management of the ERMS solution than might be necessary if the records are retained for relatively short periods.

Functional requirements

S.6.1 (M) The ERMS *must* have the capability of generating all required reports in a standard format (e.g. XML, *.csv*, HTML).

S.6.2 (M) The ERMS *must* have the capability of allowing the Administrator to save each report generated under any of the preceding requirements as either a discrete record or as technical documentation.

S.7 Audit

Rationale

To maintain trust in the authenticity of the records, it is essential for the sustainability solution to maintain an audit of actions carried out to them. As with the record and other metadata, this evidences that the actions carried out to the digital objects and metadata were those authorised by the policies and procedures used to govern the system and were carried out by duly authorised personnel.

Note: *This section assumes that all of the audit requirements contained in the 2002 Requirements need to be met, including the two articulated there as highly desirable [A.6.6 and A.6.12). Additional operations requiring full audit in this environment and not enumerated in 2002 comprise an additional series of bulleted events within requirements A.6.2 & A.6.3: see "augmented requirements" below. These are also repeated in Annex 2].*

Functional requirements

S.7.1 (M) The ERMS *must* be able to record in the audit trail all unauthorised attempts to view digital objects and metadata.

S.7.2 (HD) The ERMS *should* be able to record in the audit trail all viewings of the record content, specifying the identifier of the object[s] viewed, the time and the identity of the relevant user.

Augmented TNA 2002 functional requirements [changes tracked]:

A.6.2 (M) The ERMS must be able to record in the audit trail all changes made to:

- groups of electronic folders
- individual electronic folders
- electronic parts
- ~~electronic records~~ digital objects
- extracts of digital objects or other manifestations of record content
- metadata associated with any of the above.

A.6.3 (M) In particular, the ERMS must be capable of recording information in the audit trail about the following events:

- *the ingest and / or import of digital objects and metadata*
- the date and time of declaration of all electronic records
- re-location of an electronic record to another electronic folder, identifying both source and destination folders
- re-location of an electronic folder to a different class, identifying both source and destination classes
- re-allocation of a disposal schedule to an object, identifying both previous and re-allocated schedules
- placing of a disposal hold on a folder
- the date and time of a change made to any metadata associated with electronic folders or electronic records
- changes made to the allocation of access control markings to an electronic folder, electronic record or user
- the creation of additional manifestations of record content
- export actions carried out on an electronic folder
- separately, deletion or destruction actions carried out on an electronic folder or electronic record, by all users including an Administrator.

[Operations carried out to manage the storage, performance of the solution and affecting the physical but not the *logical* storage of the records need not be present in the record metadata but must be present in the audit trail]

S.8 Custodial issues and archival mappings

Rationale

The Generic requirements for sustaining electronic information published in 2003 referenced archival scientific research but the main intellectual framework was that of BS-ISO 15489 (*Information and documentation: records management*). Although this is an International Standard and an authoritative publication owing to its successful completion of the standardisation process, it claims not to apply to records held in an archival institution.

Given that the positioning of this set of requirements also has an important archival purpose (see previous page on the Seamless Flow programme), the opportunity has been taken in these requirements to articulate further the relationship between sustaining digital records for extended periods by government departments and archival scientific thinking on digital preservation.

Archival science has demonstrated that:

- The concept of an original is meaningless in the digital environment; instead the issue is whether the record [comprising digital one or more objects and metadata] is what it purports to be[44];

- It is not necessary to preserve the bitstream originally used to encode the digital objects [and metadata] to have an authentic [representation of the] record. This permits active preservation activity, such as migrating the digital object[s], if it is subject to appropriate controls.

The requirements to be observed vary considerably according to the perspective being applied: especially marked are the potential differences between the perspective of the creator, the custodian and the preserver where these are separate entities. The requirements set out in this document focus on the needs of the business maintaining its own records – or appointing an external custodian to do it as its agent – for long periods, irrespective of whether the records are ultimately destined for archival custody.

For example: from the purely archival perspective if a record is asserted and certified in some way by a trusted representative of the creator to be the record created and relied on by the creating organisation for reference, accountability and evidence of its business activities it is a valid record. Early transfer of digital archival records irons out many of the problems likely to be encountered here: the link between the creation of the records and the business function that led to it[45] is closer and may not require the creation of an additional controlled environment where trusted intervention can be taken, unauthorised intervention is prevented, and records will be maintained by persons who have no stake in what the records [digital object and metadata] actually *say*. Technical documentation about the records will also be easier to come by.

44 The first phase of the InterPARES project took this a stage further and declared that it was not possible to preserve the record at all, but only the ability to reproduce the record in authentic form http://www.interpares.org This is a logical extension of the same argument.

45 The abstraction of the business function in the classification scheme and the logical location of the record within it constitutes the "archival bond".

This is not to say that an archival authority is likely to be totally uninterested in the effectiveness of the management of the record throughout its lifecycle: it may also have responsibility for records management standards, for example. With digital records, the presumption of authenticity relies on a minimum meeting of rudimentary characteristics and some of these have to be present from the record creation stage.

As a result of such custodial considerations, from both a business and an archival perspective these requirements lean more towards the records continuum viewpoint in accepting the need for comprehensive metadata and a consciously high level of accountability for preservation processes. This requires comprehensive management of the digital objects throughout their life and the accrual of metadata to record every significant event.

It is likely in an increasingly distributed digital environment that records and their component parts will often be encountered in varying custodial situations. Accordingly, it must in principle be possible for much of the evidence of the authenticity of the record to travel with it. For example, a custodian must have custody of the records and their accompanying metadata but also the technical metadata, the custodial contract [if any] or other mandate under which it operates.

'Trusted custodian' concept

The digital preservation community has articulated a concept of a "Trusted custodian"[46]. A Trusted custodian has a number of important characteristics:

- Technical, financial and organisational ability to manage and preserve a variety of digital resource types for extended periods;
- Provision of a contractually-stated level of security for and access to the content [or other appropriate arrangement, e.g. statutory access regimes like FOIA];
- Little or no interest in changing what the content being cared for actually *says*[47];
- Transparent and auditable compliance with policies, practices and performance measures;
- Ability to meet stakeholder expectations of trustworthiness.

This idea resembles what a traditional archive might have done with analogue records, reinterpreted for the digital environment given that digital records are so much more vulnerable to alteration. Although it is normally assumed that a trusted custodian is a wholly *external* agent, this concept does map to a number of scenarios that may apply within government, such as:

- Contracted out provision of IT infrastructure, applications or other services;
- Centrally provided common services;
- Services contracted out by a consortium of Departments;
- Management of semi-current records by information services or departmental records functions within the same organisation;
- Management of archival material by places of deposit appointed under S.4 of the Public Records Act 1958.

46 Or "trusted digital repository", see http://www.rlg.org Only a small amount of summary information about the concept can be provided here.

47 This can only be achieved inside the same organisation by a rigorous separation of functional roles, documented in policies and technical documentation.

Functional requirements

S.8.1 (M) The ERMS *must* provide evidence in its technical documentation and record metadata that trusted and permissible processes and procedures have been carried out on the objects within its control and unauthorised processes and procedures have *not* been carried out[48].

S.8.2 (M) The ERMS *must* maintain the archival bond between the digital objects, their accompanying metadata and the arrangement of the records based on the business classification scheme used in the creating environment[49].

S.8.3 (M) The ERMS *must* support the documentation of comprehensive information on what controls and intervention the digital objects have been subject to throughout their life within the system. This will exist at a number of different levels and must support the mandatory *Requirements* in this specification.

- At a high level, this will comprise policies and [if the provision of all or parts of the technical environment are supplied under contract] contracts;
- Other documentation will consist of audit / inspection reports of the public authority or third party auditors, quality management processes;
- Other documentation at system /ERMS level will consist of audit trails / technical documentation held to show how the solution and parts of its content have been managed and batch level metadata about digital objects and their metadata [best stored once and linked to all the entities it applies to for performance and scalability reasons];
- Linkage of record and metadata;
- Some of the documentation will be applicable to individual objects and should be recorded in the relevant metadata field. As a general principle, the record metadata should record the full life history of the object and its description. This will include:
  – Authenticity of digital objects in the ERMS;
  – Metadata and mapping it to archival description.

Non-functional requirements

S.8.4 (M) Operating and governance procedures for the ERMS *must* ensure that records are formally held in the custody of persons with no direct interest in manipulating the content of the digital objects and description comprising the records in the interests of changing their interpretation by others[50].

48 There remains a point of weakness where access to the underlying document repository is feasible without the arbitration of the ERMS. See [non-functional] *Security* requirements above. System documentation should give a categoric list of activities feasible without going through the interface of the ERMS client itself; such a list would in any case be required to comply with requirements S.2 and A.10.1 from TNA 2002 [cited in Annex 2].

49 i.e. through persistence in the relational and classification metadata. This requirement is closely related to TNA 2002 Requirement A.1.68 cited in the next sub-section.

50 Many agents can potentially operate as trusted custodians: for example a Departmental Record Officer in a government department, an ICT department, a trusted contractor where the terms of the contract can be relied on to enforce the requisite level of control.

S.9 Authentication and certification[51]

Rationale

The transmission of records [comprising digital objects and metadata] from the sustainability solution to other infrastructure for any purpose sometimes needs to be accompanied by some authoritative statement about their status. This can be a very formal requirement sanctioned or mandated by legislation or regulations, or an administrative activity such as at the export of some or all of the contents of the digital object repository and their metadata to another system.

In some cases this transmission must be accompanied by a formal type of certificate, such as a signature of an authorised agent, or merely another record attesting to the authenticity of the records.

Juridical types of authentication

Many public authorities have powers and duties to issue extracts from their records to other parties. For example, for presentation as legal evidence or attestation of an entry in a public register. Longstanding legislation sometimes requires these extracts to be accompanied by a signature and, once it is applied, various juridical implications may kick in. For example, the copy can [and in some cases must] be treated in evidence or some other process as having the same status as the "original"[52]. Occasionally, a statutory provision requires facts attested in such a certified extract to be treated as fact. This is unusual and is the highest level of standing that can be accorded.

All of these types of authentication depend for their repute on an unusual level of trust placed in the organisation involved by the Supreme Court *Civil Procedure Rules* or Parliament (in statute)[53]. Obviously, continued trust in them relies on those who have these power vested in them to use them responsibly. The relationship with the trusted custodian issues outlined in the preceding section is that the credibility of the certification depends on the trust in the policies and procedures that have been applied in the care of the records.

As a consequence, procedures must be put into place to ensure that a formal certification is only applied where it is justified by the authenticity of the records and thus the extent to which they have been managed according to trusted processes [i.e. the policies, procedures of the solution clearly designed to support the authenticity of the records in the first place]. This will require an assessment of the metadata attached to the digital objects, other technical documentation such as metadata schemas used to manage the records, the system audit trail[s] etc. Other types of certification may exist within the same organisation, e.g. to support eGovernment transactional processes and these must be procedurally distinct from formal juridical certification. A presumption of authenticity is dependent on certain minimum metadata and trust in its veracity based on the technical and policy documentation of the solution.

51 These requirements do not recommend the use of asymmetrical cryptographic techniques to "assure" the "authenticity" of digital records, as such techniques are neither capable of supporting preservation, nor have any role within the sustainability solution and do not apply an archival science view of authenticity to the problem. There are many authorities who make the same judgement.

52 This is often the wording of an ageing statutory provision or the concept behind it, but is meaningless in the digital environment except where it is interpreted as "a copy that is what it purports to be". [See above].

53 This is something of a legal anomaly in English law. The highest level of evidence in a common law system such as ours is typically accorded to sworn testimony subject to cross-examination in a court of law. The Roman [Civil] Law systems of other jurisdictions including Scotland and many of the EU member states accords the highest evidentiary value to certain types of document.

Effect of the Electronic Communications Act 2000 ["ECA 2000"]

The ECA 2000 is a primary enactment that allows any legislation that calls for a manual signature or seal of an authorised officer representing the Chief Executive [or other person] to be amended by secondary legislation [typically regulations contained in a Statutory Instrument] to permit this to be done by electronic means.

The records management system of a public authority may be used to manage, preserve and transmit such extracts in the future. This is a type of the more generic issue of certifying copies or extracts from information resources held by a public authority in the following section [certification of records not held in the environment specified in these requirement is, obviously, not covered here].

Policy lead for this and other aspects of the implementation of the EU Directives on electronic commerce and digital signatures rests with the Department of Trade and Industry. Other EU jurisdictions have been far more radical in their domestic enactment than the UK where the above Act is only augmented with regulations to regulate the third party provision of digital signatures[54].

Certification of extracted copies

Certifying digital objects and metadata from a controlled environment such as an ERMS used to sustain digital records will normally involve accompanying extracts from the digital object store and metadatabase with an asymmetrical cryptographic digital signature, watermark or other technical device.

Signatures have their place for this purpose, but probably *cannot be preserved*. This is for a range of reasons:

- They cannot be preserved except where the approach is one of bit-level preservation. If the objects are migrated to preserve them or additions made to the metadata, the digital signature will indicate that the object[s] has [have] changed[55].
- They cannot be preserved owing to the likelihood of the signature provider being a commercial company subject to take over, bankruptcy, etc.
- Preserving metadata about the fact that a signature was validated at some time *in the past* is about as much as can be reasonably achieved in the long term.

[This ephemeral nature could be turned to advantage if providing the certified copy is benefited by the signature being dependent on no changes being made to the extract after it has left the controlled environment of the ERMS, as a rudimentary rights management feature. Digital watermarks have the disadvantage that they change materially the objects themselves].

54 The Irish and Italian jurisdictions have, for example, legislated for what the EU eSignature directive calls standard and 'advanced' digital signatures. The latter have non-repudiation provisions that may be more suitable for civil than common law jurisdictions.

55 For the same reasons the use of such signatures within a sustainability of preservation solution itself – i.e. to "ensure" authenticity – is not recommended by any significant number of authorities.

Use of signatures to control administrator workflows within the ERMS and other transactions

A more low-level type of certification without a formal juridical mandate may be established as an internal procedure within the ERMS or a procedure in identifying citizens or public servants in transactions. These may use similar technologies as more formal juridical authentication – but need not. It is important that the significance of the digital signature is clearly understood for its deployment within records management systems to be appropriate. For example:

- Checking the identity of the administrator carrying out certain key processes in the ERMS is just an elaboration of the functional access control of the ERMS;
- A risk assessment of any business process may suggest that the additional security provided by a digital signature would be beneficial.

Functional requirements

S.9.1 (M) The ERMS *must* support the issue of individual or aggregations of records [digital objects and their metadata] in a form that is signed with a digital signature [in this requirement, "support" will normally mean integrate with another utility actually providing the signature].

Non-functional requirements

S.9.2 (M) The ERMS *must* be accompanied by procedures for the evaluation of whether the requirements for the presumption of authenticity are met[56]:

- The condition of the records on import or ingest [evidenced by metadata, technical environment of creating or other previous environment[s][57];
- The procedures applied to the records whilst under the control of the ERMS [ditto].

If they are not met a qualified or no certificate must be issued and procedures for this should be instituted.

Reporting requirements

S.9.3 (M) The ERMS *must* support the running of reports to compile the documentation to asses the extent to which the presumption of authenticity is supported [see requirement S.9.2].

56 There may be, by extension, additional requirements imposed by any conditions particularly on juridical authentication powers which also need to be met.
57 As evidenced in metadata, technical documentation, etc.

Annex 1: Terminology

Term	Definition
Active preservation	See *Preservation*.
Audit trail	System *index data* normally only under the control of a system administrator and capturing a full history of system events.
Authenticity	An authentic record is one that can be proven a) to be what it purports to be, *b) to have been created or sent by the person purported to have sent or created it, and* *c) to have been created or sent at the time purported.* *To ensure the authenticity of records, organisations should implement and document policies and procedures which control the creation, receipt, transmission, maintenance and disposition of records to ensure that records creators are authorised and identified and that records are protected against unauthorised addition, deletion, alteration, use and concealment.* [BS-ISO15489] See also the sub-attributes that follow in the Standard: *Reliability, Integrity and Useability*.
Bit-level preservation	*As a Preservation* approach this involves maintenance of the original bitstream of *digital objects*. Normally requires either the maintenance of obsolete hardware or *emulation for access and use. Bit-level preservation is also used as an 'insurance policy' in conjunction with other preservation approaches to ensure that the original object can be made available for access or further transformations if necessary.*
Capacity threshold	The total size of the storage solution that can be used without compromising the *performance metrics*.
Characterisation	Process of defining what attributes of a *record* constitute its essential parts and therefore need to be preserved across *migration. This activity is derived from the CEDARS project concept of the 'significant properties' of digital objects*.
Component	See *Record component*.
Compound object	Scenario where a *digital object* is nested inside another, either facilitated by the software specification or the behaviour of the operating system. Problematic for records management and preservation because of the risk that the contained object will be 'invisible' and not supported across technological change.
Digital object	Discrete software data capable of being handled by the operating system as an entity and interpreted by relevant software. A digital object comprising part of a record is also referred to as a *record component* where important for the meaning. [A digital object *repository* is a storage device for digital components]
Document	Informational by product of business activity in any format. May either be kept as a corporate *record* for reference, accountability or historical purposes or treated as ephemeral. [Note this is *not* the same definition as used in TNA 2002]
EDRM [Electronic Document and Records Management]	Acronym conveniently abbreviating recent convergence between technologies for collaborative [drafting] work and those supporting records management.

Term	Definition
ERMS [Electronic records management system]	Records management system where the content is managed in electronic format throughout its life.
Emulation	*Preservation* approach involving writing programmes to mimic the behaviour of obsolete hardware systems to ensure continuing access to digital objects in their original formats.
Encoding language	Convention of formatting text meaningful to computer programmes. E.g. eXtensible Mark-up Language (*XML*), ASCII.
Extract [use is deprecated – see redaction]	*Manifestation* of a *record component* where part of the content has been masked or removed for security, privacy or sensitivity reasons.
File format	Generally used to refer to the encoding structure(s) of discrete identifiable packets of digital data (files). *[Encoding of representation information of data objects [OAIS]* *Organisation of digital information according to preset specifications [PREMIS]]* See *migration*.
Fixity	In this document, refers to the validity of the data (the bits) that makes up a digital object. For file / record validity see Integrity. *[Property that a digital object has not been changed between two points in time [PREMIS]]*
Import	See *Ingest*.
Index data	*[Indexing: Process of establishing access points to facilitate retrieval of records and/or information [BS-ISO 15489]]* In these requirements, data generated for system management purposes, e.g. audit trail. May overlap with metadata – of which it forms a subset – especially where system identifiers are used to maintain record integrity.
Ingest	Function of the Open Archive in the OAIS model concerned with all aspects of taking in content representation information. Here the term is used to cover both of the following *narrower* senses of the term: (a) taking into the ERMS of digital objects with no or very limited records management metadata owing to the creating environment they have been created and/or held in. (b) taking in digital objects and metadata with the latter structured according to some sort of logical schema that may or may not map precisely to current eGovernment standards but comes from a document and / or records management environment. Scenario (b) is called "Import" in TNA 2002.
Integrity	*The integrity of a record refers to its being complete and unaltered. It is necessary that a record be protected against unauthorised alternation. Records management policies and procedures should specify what additions or annotations may be made to a record after it is created, under what circumstances additions or annotations may be authorised, and who may be authorised to make them. Any authorised annotation, addition or deletion to a record should be explicitly indicated and traceable.* [BS-ISO 15489] A concept from archival science that relates to issues of records authenticity and reliability. See section 8 above for a fuller discussion of this topic.

Term	Definition
Interoperability	*Coherent exchange of information and services between systems..... [If this is achieved, then the system can be regarded as truly interoperable. Furthermore, it must be possible to replace any component or product used within an interface with another of a similar specification while maintaining the functionality of the system.* eGIF, v. 6.1 accessed from http//:www.govtalk.gov.uk February 2006.
Manifestation	Instance or example of digital object[s], rendering the content of a record accessible in some way, possibly for a defined period of time.
Metadata	*Data describing context, content and structure of records and their management across time*. [BS-ISO 15489/1]
Metadatabase	Database within the ERMS architecture used for the storage of structured *metadata* and its linkage to *record components*.
Metadata profile	Set of metadata associated with one or more digital components together comprising a record.
Metadata stub	Cut down or complete metadata remaining after the disposal of either a record or record components.
Metadata schema	Plan or 'map' of relationships imposing structure and organisation on metadata fields to support their management. See also XML schema which is a machine readable binding of a metadata schema, but not a replacement for working out these relationships logically.
Migration	Changing the format encoding of digital objects to maintain accessibility on more current technology. OAIS uses migration to denote several scenarios falling under its definition (*the act of transferring a digital information object to new storage media or to new forms*): (a) Data migration between storage media (b) Metadata migration; and (c) Format migration (also referred to as transmutation), where the format used to encode the content is altered but the content is intended to stay as close as possible to the original. [OAIS] This last sense is the one mapped to here.
OAIS	*Open Archival Information System* reference model originally developed to support the maintenance of the [digital] spatial data collected by NASA space exploration missions and widely referenced in the digital preservation community. The OAIS reference model sets out a conceptual framework for an archival system dedicated to preserving and maintaining access to digital information over the long term. Now an international standard: ISO 14721:2003. See http://www.ccsds.org
Offline storage	See *storage*.
Open format specification	Format where the specification [rules governing its encoding] is public domain [not to be confused with *open source*].
Open source software	Software where the code used to create digital objects is public domain [not to be confused with *open format*].

Term	Definition
Open standard	An *open format* specification that has attained wide adoption or recognition through an authoritative formal standardisation body such as ISO. E.g. PDF-A [Archival].
Passive preservation	See *Preservation*.
Performance metrics	Measures of the performance of an IT system stipulating acceptable tolerances.
Personal data	*Data which relates to a living individual who can be identified -* *(a) from those data, or* *(b) from those data and other information which is in the possession of, or is likely to come into the possession of, the data controller* *- and includes any expression of opinion about the individual and any indication of the intentions of the data controller or any other person in respect of the individual*. [DPA 1998, *S*. 1]
Preservation	*Processes and operations involved in ensuring the technical and intellectual survival of authentic records through time*. [BS-ISO 15489] In these requirements this is broken down into: (a) *Active preservation*: intervention in the encoding of record components to maintain access to their contents across technological change (e.g. technology watch, *migration*); and (b) *Passive preservation*: measures taken to preserve the current encoding of record components (e.g. media monitoring and refreshing).
Record	*Information created, received and maintained as evidence and information by an organisation or person, in pursuance of a legal obligation or in the transaction of business*. [BS-ISO15489]
Record component	A constituent part of a *record*, e.g. [normally] *digital objects*, *metadata*.
Records management system	Procedures, rules, tools and infrastructure used to manage the lifecycle of records. See *ERMS*.
Redaction	A *manifestation* of record content where some material has been removed or securely masked. Typically made when all the *record* content cannot be made available, but part can. [TNA 2002]
Relationships	Logical, technical or semantic linkage between entities including *metadata* and *digital objects*.
Refreshing	Term frequently used in digital preservation community to denote processes called *migration* in OAIS model but here, and most commonly, the rewriting of digital objects and / or metadata to new media.
Reliability	*A reliable record is one which contents can be trusted as a full and accurate representation of the transactions, activities or facts to which they attest and can be depended upon in the course of subsequent transactions or activities. Records should be created at the time of the transaction or incident to which they relate, or soon afterwards, by individuals who have direct knowledge of the facts or instruments routinely used within the business to conduct the transaction*. [BS-ISO 15490]

Term	Definition
Rendition	A *manifestation* of a r*ecord component* where the file format has been converted.
Replication	Creation of [a] bit-identical digital copy[ies]. [OAIS]
Repository	Database within the ERMS solution used for the storage of digital objects forming record components.
Scalability	Ability to increase the capacity of the ERMS solution in terms of total *repository* and metadata storage whilst remaining within the *capacity threshold*.
Sensitive personal data	*Personal data consisting of information as to:* *(a)the racial or ethnic origin of the data subject,* *(b)his political opinions,* *(c) his religious or other beliefs of a similar nature,* *(d)whether he is a member of a trade union (within the meaning of the Trade Union and Labour Relation [Consolidation] Act 1992),* *(e) his physical or mental health or condition,* *(f) his sexual life,* *(g) the commission or the alleged commission by him of any offence, or* *(h) any proceedings for any offence committed or alleged to have been committed by him, the disposal of such proceedings or the sentence of any court in such proceedings*. [DPA 1998, S. 2]
Standard format	Format in sufficiently wide use to be deemed an industry standard. May be an *open standard* or produced by *open source* software but need not be.
Sustainability [of digital records]	Concept developed by The National Archives in 2000-03 to distinguish preservation for the business purposes of government from archival *preservation* (e.g. by TNA). A type of *preservation*.
Storage	3 main digital storage scenarios are referred to in this document: *Online* – permanently connected to access or management interfaces; *Nearline* – not routinely or permanently connected but possible to load relatively quickly as a standard and possibly automated procedure; *Offline* – without physical connection to access or management devices, e.g. on removable storage remote from this infrastructure.
Technical documentation	Any information related to the workings of a system or its contents of general pertinence. Unlikely only to refer to small portions of its content and impracticable to record comprehensively as part of the metadata of every *record* so stored as technical documentation in association.
Technology watch	Programme monitoring changes in the technical environment likely to affect the preservation of digital objects to enable preservation planning to occur.
Trusted custodian	Party trusted to preserve [digital] records owing to capabilities, policies and impartiality as to the content. See discussion in *Custodial Issues and Archival Mappings*. See also the attributes of a *trusted digital repository* published by the Research Libraries Group: http://www.rlg.org

Term	Definition
Useability	A useable record is one that can be located, retrieved, presented and interpreted [BS-ISO15489]. It should be capable of subsequent presentation as directly connected to the business activity or transaction that produced it. The contextual linkages of records should carry the information needed for an understanding of the transactions that created and used them. It should be possible to identify a record within the context of broader business activities and functions. The links between records that document a sequence of activities should be maintained. [BS-ISO 15489]
Volume	Storage partition used by a solution in combination with an identifier system largely invisible to the end user to manage the physical storage of the system's content. May have no relation to the logical storage of the content.
Viewer	Technology providing ability to view some or all aspects of an *instance* of a software format without a full software application being present.
XML schema	Validation method for examining *XML* text against a set of logical rules. A validation tool, e.g. XML Spy, will determine whether an instance of XML code is [a] well formed according to the XML specification and [b] complies with the rules contained in the specific schema. This makes it valuable for supporting some aspects of *interoperability*.
XML	eXtensible Mark-Up Language readable by machine and – to an extent – humans. An encoding language whose specification is owned and maintained by the W3C. Widely touted as an essential tool for web infrastructure owing to its transformability, flexibility and compatibility with other web technologies (e.g. HTML, XSLT). See also *XML schema*.

Annex 2: Mappings to 2002 National Archives Requirements

TNA 2002 reference	Previous wording	Sustainability requirements ref	Comment
A.2.20 (M)	The ERMS *must not* impose, by its architecture or design, any practical limit on the number of records that can be captured and declared into a folder; *or on the number of records that can be captured and declared into the ERMS as a whole*.	S.1.1 - S.1.15	The second half of this requirement is emphasised here: all It systems ultimately have scalability issues; the point is that it should not be imposed by a conscious design constraint.
A.9.1 (M, A)	The ERMS *must* provide a robust and flexible architecture that can evolve to meet the needs of a changing organisational environment, appropriate to the types of implementation for which the ERMS is intended.	S.1.1 – S.1.15	This requirement has had to be articulated in more definite terms for this environment and subdivided into the groups of new requirements cited.
A.9.18 (HD)	*The ERMS should support a distributed repository with multi-site service*.	S.1.1 - S.1.15	There is a far greater likelihood that this requirement [and others cited below in the same sequence] will need to be promoted to **mandatory** requirements in this environment [the marking "A" (*Advisory* – and therefore untested in previous TNA software evaluation activity) is not relevant in this context*]. They are substantively rearticulated in the new requirements set out above to fit the new context*.
A.9.19 (HD)	*The ERMS should support caching of frequently and recently used repository content*.	S.1.10	There is a far greater likelihood that this requirement [and others cited below in the same sequence] will need to be promoted to **mandatory.** Text runs: *If the previous 3 requirements are met, the ERMS should support the intelligent management of storage, e.g. as to performance, availability and robustness*. This is a more generic statement of the same aims. It could be met in the ways envisaged in 2002, but need not.
A.9.20 (HD)	When querying a remote repository, the ERMS *should* minimise the amount of data exchange required.	S.1.10 – S. 1.15	There is a far greater likelihood that this requirement [and others cited below in the same sequence] will need to be promoted to **mandatory**. [See above]
A.9.21 (M)	The ERMS *must* provide facilities for monitoring storage facilities, and automatically alert an Administrator when a capacity threshold is reached, or when an error condition requiring attention occurs.	S.1.15 – S.1.20	The previous requirement has been split to unite the first half with more general reporting concerns.

TNA 2002 reference	Previous wording	Sustainability requirements ref	Comment
A.9.22 (M)	The ERMS *must* provide evidence of adequate performance and response times for commonly performed functions under the normal operating conditions for which it is intended. A benchmark for normal operating conditions is: • 75% of the user population actively using the system • *total record volume to be expected after 5*[58] *years use stored* • multiple, concurrent and representative active use of system functionality Benchmark metrics for performance are: • time taken to display a graphical view of the class and folder structure • time to store a set of standard documents at capture and/or declaration • time to return a search response for a simple search • time to return a search response for a complex (Boolean) search • time to display a recently captured record • *time to display an 'inactive' record.*	S.1.16 – S.1.20	A sustainability solution will need its own set of [system] performance criteria according to the business environment. The logical and technical architecture may impose important constraints on availability and response times. [The remainder of this previous requirement is concerned with the management of current records]
A.9.23 (M)	The ERMS *must* provide evidence of the degree of scalability that it can support over time, as organisational needs change and develop. Benchmark metrics for scalability are: • number of geographical locations at which users can be supported, while maintaining the performance metrics demonstrated • total size of the record repository which can be supported, in Gigabytes or Terabytes, while maintaining the performance metrics demonstrated • number of total users which can be supported while maintaining the performance metrics demonstrated • systems management overhead in maintaining growth rate *for the number of records and users anticipated in the first five years of operation.*[59]	S.1.1 – S.1.20	The sustainability requirements assume a more distributed logical architecture to support preservation so the entire requirement is reworked; see also Logical architecture. Timescales specified are too short to be relevant here.

58 The 5 year restriction in this and the following requirement for the current records environment is, obviously, to be set aside in these functional requirements.
59 To become mandatory, in accordance with the schedule for implementation of the finalised XML schema when published.

TNA 2002 reference	Previous wording	Sustainability requirements ref	Comment
A.9.23 (M) Continued	• amount of re-configuration and downtime required to maintain a growth rate for the number of records and users anticipated *for the first five years of operation* • amount of re-configuration and downtime required to make bulk changes to organisational structures, class and folder structures, and user roles with the number of folders, records and users anticipated *after five years of operation*.		
A.2.56 (HD).	The ERMS *should* be able to allow the creation of an extract directly from an originating record, where portions of original content have been masked in the extract, while ensuring the original record remains intact.	S.2.1 – S.2.8	This requirements needs to be *mandatory* in this environment. [This guidance denigrates the use of the terminology "extract" in favour of *redaction*].
A.9.11 (M) – A.9.17 (M) Disaster recovery requirements	–	S.1.1 – S.1.4, S.1.18 – S.1.20, S.1.21 – S.1.38	A fuller set of requirements for the management of and replication to removable media are articulated in 1, 2, 5, 7, 8, 9, 10 and elsewhere. The disaster recovery elements of A.9.11 – A.9.17 remain relevant, as does A.10.1. [Compliance with Standards including BS-ISO 17799]
A.10.1 (M)	Wherever relevant, the ERMS *must* comply with, or support compliance with, the following standards: • UK Government eGIF (*eGovernment Interoperability Framework*) • UK GDSC (*Government Data Standards Catalogue*) • *eGovernment Strategy Policy Framework and guidelines: Security* • BSi BIP PD0008 2004 *Code of practice for legal admissibility and evidential weight of information stored electronically* • BSi BIP PD0018 2001 *Code of practice: Information management systems: building systems fit for audit* • ISO17799 / BS7799 *Information security management*.	Interoperability / openness, also Annex 3 Generic data entity models. [See also migration requirements]	This section articulates these requirements to a far greater degree than previously and will be extended with revisions of the records management specialisation of the *eGovernment Metadata Standard* [including its machine-readable bindings].

TNA 2002 reference	Previous wording	Sustainability requirements ref	Comment
A.1.68(M)	The ERMS *must* maintain full structural integrity of the class, folder, part and record structure at all times, regardless of maintenance activities, user actions or component failures.	Interoperability	Section 3 covers the main interoperability rationale and the particular requirements of the sustainability solution. Many other requirements articulated in TNA 2002 for the current records environment impact on this area: only the most significant are reproduced here. The same record integrity, metadata editing rules, import / export requirements and particularly the persistent linkage of metadata profiles to record components criteria apply as for current RM.
A.2.31 (M)	The ERMS *must* support the capture and presentation of metadata for electronic records as set out in the records management specialisation of the eGovernment Metadata Standard ["eGMS"] published by The National Archives in collaboration with the Cabinet Office and other public authorities.	*See above*	*See above*
A.2.32 (M)	The ERMS *must* ensure the capture of all required metadata elements specified at systems configuration, and retain them with the electronic record in a tightly bound relationship at all times.	*See above*	*See above*
A.2.39 (M)	The ERMS *must* prevent any amendment of selected elements of metadata of the electronic record which have been acquired directly from the application package, the operating systems of the ERMS itself (for example, certain dates) as defined by the records management specialisation of the eGMS.	*See above*	*See above*
A.2.41 (M)	The ERMS *must* ensure that the content of selected items of metadata (a subset of those that may be changed) of the electronic record can only be changed by an authorised user, as defined in the records management specialisation of the eGMS.	*See above*	*See above*

TNA 2002 reference	Previous wording	Sustainability requirements ref	Comment
A.2.49 (M)	The ERMS *must* be able to allocate an identifier, unique within the system, to each electronic record on declaration, that serves to identify the record from the point of declaration throughout the remainder of its life within the ERMS.	*See above*	*See above*
A.2.63 (M)	The ERMS *must* be able to capture in bulk records exported from other records management and document management systems, including capture of: • electronic records in their existing format, without degradation of content or structure, retaining the relationship between the components of any individual record • electronic records and all associated metadata, retaining the correct relationship between individual records and their metadata attributes • the folder structure to which the records are assigned, and all associated metadata, retaining the correct relationship between records and folders.	*See above*	*See above*
A.2.64 (HD)	The ERMS *should* be to import any directly associated audit information with the record and/or folder, retaining this securely within the imported structure.	*See above*	*See above*
A.2.65 (M)	Within the schedule for implementation, the ERMS *must* be able to directly import, in bulk, electronic records in their existing format with associated metadata that is presented according to a pre-defined XML schema[60] (schema to be defined based on the accompanying records management metadata standard), mapping this to the receiving ERMS folder and metadata element structures.	*See above*	*See above*
A.4.55 (M)	The ERMS *must* be able to export metadata for folders, parts and records in an XML format.	*See above*	*See above*

60 This brief list of metadata attributes are taken from the records management specialization of the eGMS. It is not a substitute for the comprehensive evidencing of all processes the records have been subjected to advocated elsewhere in these Requirements; they are the absolute minimum

TNA 2002 reference	Previous wording	Sustainability requirements ref	Comment
A.4.56 (M)	The ERMS *must* be able to support the export of metadata as defined by the electronic records management standard schema, as versions become available through the GovTalk site (www.govtalk.gov.uk), and in accordance with the schedule for compliance.	*See above*	*See above*
A.4.57 (M)	The ERMS *must* also be able to export records: • in their native format, or a current format to which they have been migrated And in order of preference: • in an XML format which falls within the UK eGIF framework, where possible • in a rendition which is consistent with the range of formats currently specified in the eGIF set, where an XML format is not available. Such renditions may be achieved by: • capturing an appropriate rendition as part of the record capture process • rendering the record as part of the export process • exporting directly to another package which is capable of rendering the record within a controlled environment.	Active preservation	The core active preservation functionality outlined as part of a sustainability solution means that restating this "rendition at export" scenario may well be redundant.
A 2.12 (HD)	When capturing a document in its native format, the ERMS *should* also be capable of also capturing a rendition of that document in a standard format, and of storing native format and rendition in a close association. Standard rendition formats include: XML, PDF and Postscript.	S.3 interoperability / openness requirements, S.4 Active preservation requirements, Annex 3: Generic data entitiy models for common record types.	[Rendition at export – TNA 2002 requirement A. 4.57 and import is similarly handled here].

TNA 2002 reference	Previous wording	Sustainability requirements ref	Comment
A.2.31 (M)	The ERMS *must* support the capture and presentation of metadata for electronic records as set out in the accompanying metadata standard for electronic records management (and numerous other metadata requirements).	S.4.3 - S.4.10 Active preservation requirements S.4.39 – S.4.41 Preservation process metadata	Invoked by extension from A.2.31 and development of the records management specialisation of the eGovernment Metadata Standard.
A.2.66 (M)	The ERMS *must* be able to indirectly import, in bulk, electronic records in their existing format with associated metadata that is presented in a non-standard format, mapping this to the receiving ERMS folder and metadata element structures.	S. 5 Ingest	
A.2.67 (HD)	The ERMS *should* be able to import, in bulk, existing electronic documents, in any and all supported formats, that have no associated metadata presented separately from the document, by: • placing documents in queues for further processing • automatically extracting metadata from the document properties where possible • providing facilities for the addition of missing metadata, and the assignment of documents to folders • supporting the declaration of documents from these processing queues.	S.5.1 – S.5.6 Ingest	
A.4.57 (M)	The ERMS *must* also be able to export records: • in their native format, or a current format to which they have been migrated And in order of preference: • in an XML format which falls within the UK eGIF framework, where possible • in a rendition which is consistent with the range of formats currently specified in the eGIF set, where an XML format is not available. Such renditions may be achieved by: • capturing an appropriate rendition as part of the record capture process • rendering the record as part of the export process • exporting directly to another package which is capable of rendering the record within a controlled environment.	S.3 interoperability / openness requirements, S.4 Active preservation requirements, Annex 3: Generic data entity models for common record types.	[Rendition at *capture* – TNA 2002 requirement A. 2.12 and *ingest* is similarly handled here]

TNA 2002 reference	Previous wording	Sustainability requirements ref	Comment
A.6.2 (M)	The ERMS *must* be able to record in the audit trail all changes made to: • groups of electronic folders • individual electronic folders • electronic parts • electronic records • extracts • metadata associated with any of the above.	S.7 Audit	Additional operations require to be audited in this environment. The full rearticulated requirement is as follows [tracked changes shown]: The ERMS *must* be able to record in the audit trail all changes made to: • groups of electronic folders • individual electronic folders • electronic parts • ~~electronic records~~ digital objects • extracts of digital objects *or other manifestations of record content* • metadata associated with any of the above.
A.6.3	In particular, the ERMS *must* be capable of recording information in the audit trail about the following events: • the date and time of declaration of all electronic records • re-location of an electronic record to another electronic folder, identifying both source and destination folders • re-location of an electronic folder to a different class, identifying both source and destination classes • re-allocation of a disposal schedule to an object, identifying both previous and re-allocated schedules • placing of a disposal hold on a folder • the date and time of a change made to any metadata associated with electronic folders or electronic records • changes made to the allocation of access control markings to an electronic folder, electronic record or user • the creation of additional manifestations of record content • export actions carried out on an electronic folder • separately, deletion or destruction actions carried out on an electronic folder or electronic record, by all users including an Administrator.	S.7 Audit	Additional operations require to be audited in this environment. The full rearticulated requirement is as follows [track changes shown]: In particular, the ERMS *must* be capable of recording information in the audit trail about the following events: • *the ingest and / or import of digital objects and metadata* • the date and time of declaration of all electronic records • re-location of an electronic record to another electronic folder, identifying both source and destination folders • re-location of an electronic folder to a different class, identifying both source and destination classes • re-allocation of a disposal schedule to an object, identifying both previous and re-allocated schedules • placing of a disposal hold on a folder • the date and time of a change made to any metadata associated with electronic folders or electronic records • changes made to the allocation of access control markings to an electronic folder, electronic record or user • the creation of additional manifestations of record content • export actions carried out on an electronic folder • separately, deletion or destruction actions carried out on an electronic folder ~~or~~ electronic record, *or record manifestation* by all users including an Administrator.

TNA 2002 reference	Previous wording	Sustainability requirements ref	Comment
A.6.6 (HD) & A.6.12 (HD)		S. 7 Audit	These requirements need to be mandatory in this environment.

Annex 3: Generic data entity models for common record types

[This section for completion when further work emerges from The National Archives digital preservation programme.]

Annex 4: Minimum directly-linked metadata[61]

This reduced metadata set is the bare minimum adequate to take digital objects asserted to be the records of a business activity into a more controlled environment, such as a sustainability solution. As such they are minimum requirements for ingest if the subsequent management of the digital objects in a controlled environment is to have any possibility of redeeming limitations in the earlier management environment.

The metadata must be persistently linked to the content, normally by being present in the *record profile* which has some relational linkage to the record component[s].

If this minimum set is not available [i.e. neither present with the records in the creating environment nor possible to reconstruct with complete confidence at the point of import from other associated documentation or reliable first-hand knowledge[62]], these *Requirements* propose that ***the records ought not to be certified as authentic without a suitable qualifier***. See section on Authentication and certification.

- Identifier System ID
- Title
- Creator
- Date Created
- Date Acquired (mandatory for e-mail)
- Date Declared
- Addressee (mandatory for e-mail)
- Type Record type (mandatory where applicable)
- Relation Copy (pointer) (mandatory where applicable)
- Relation Parent object
- Relation Redaction/Extract (mandatory where applicable)
- Relation Reason for redaction/extract (mandatory where applicable)
- Relation Rendition (mandatory where applicable)
- Aggregation
- Rights Protective marking

In addition, there is extra contextual information that is required to attest to the integrity of the records should they no longer exist in the original creating environment. This may, but need not, be linked to every record or its components unless it is specific to them. If it applies across aggregations, it may be stored in the *technical documentation* but it must be clear to which records it applies and be available for export should they be sent to another platform.

To support a presumption of integrity of the records, the custodian must possess or obtain evidence that the following attributes are supported:

- name of the creating organisation that regards the record as part of its official corporate record
- name of the organisation which has custody of the record (if different from the creating organisation)
- indication of types of annotations added to the record
- indication of technical modifications.

61 This brief list of metadata attributes are taken from the records management specialization of the eGMS. It is not a substitute for the comprehensive evidencing of all processes the records have been subjected to advocated elsewhere in these Requirements; they are the absolute minimum for a viable record to be present at all.

62 Such attestations are required to be preserved with the records.

Tel: 020 8876 3444
Fax: 020 8392 5286
email: records.management@nationalarchives.gov.uk
The National Archives Kew Surrey TW9 4DU