



**DIGITAL CONTINUITY FRAMEWORK
STAKEHOLDER CONSULTATION RESPONSE**

ORGANISATION NAME _____

CONTACT NAME _____

CONTACT EMAIL _____

**CONTACT ROLE AND
RESPONSIBILITIES** _____

CONTACT PHONE _____

OUTLINE

The Digital Continuity project is delivering a service for the public sector that will enable it to use its business-critical digital information for as long as it needs to. Part of this service will include procuring a Framework of tools and services.

The purpose of this document is to invite government and public sector stakeholders to comment on the outline requirements we have developed for this Framework.

Digital Continuity background

Government and the wider public sector rely on digital information for ongoing business, to drive efficiencies and support accountability, and to safeguard the record of government for posterity. The Digital Continuity project, managed by The National Archives, and funded by all central government departments, was set up in 2007 to deliver a shared service that enables government to take action to ensure that essential business information remains complete, available and usable for as long as it is needed (see page 4).

This service will comprise guidance and a Framework of tools and services to support public sector organisations to identify and resolve Digital Continuity issues and to understand and manage exposure to Digital Continuity risks.

Digital Continuity Framework

The Digital Continuity Framework will provide a comprehensive catalogue of **technology tools** and **services**. It will focus on addressing issues associated with managing Digital Continuity, including digital obsolescence, information management, re-use and potentially delivering data management efficiencies.

It is important to note the tools and services on the Digital Continuity Framework are to meet the requirements of public sector organisations in a business context. The Framework is not to provide solutions for use within The National Archives' digital archive.

The Framework will be made available to the wider public sector, including Central Government departments (26), local government, Executive Agencies (77), Non Departmental Public Bodies (~1000), Police Authorities, the NHS and the charitable sector.

Categorising tools and services in the outline requirements

This document provides an outline of the requirements for tools and services that public sector organisations may need to manage their digital continuity. The outline requirements for this Framework are presented as notional 'lots' according to the role they play in addressing Digital Continuity issues, and our understanding of technology segmentation within the market place. Some tools or services may, therefore, address the requirements of more than one lot.

Consultation

The purpose of this consultation is to ask for stakeholder feedback on the definition and segmentation of the requirements.

Your feedback is important. It will support the development of the requirements, assure the commercial viability of the proposed Framework procurement and fit to stakeholder needs. We are requesting feedback in the following two areas:

- **Definition comments**
 - Are the requirements clearly defined and structured?
 - Is there anything missing from the definition or scope of the requirement?
- **Your requirements**
 - Can you envisage using the tool or service, now or in the future? If so, in what context?
 - Do you know of any particular products or suppliers?

Please give any general feedback on the scope of requirements in the dedicated section at the end of this document. In particular, please let us know if you think there is anything missing that you would like to see on the framework.

All responses provided to this consultation will be used by The National Archives and Buying Solutions only to inform the procurement framework definition. They will not be shared with other departments or suppliers.

Deadlines

Once you have completed your response please return your document by the 30th of September to: DCSprocurement@nationalarchives.gsi.gov.uk

What we mean by 'Digital Continuity'

Digital Continuity is when digital information remains complete, available and useable for as long as it is needed for business purposes. This means ensuring that digital information has the necessary context, can be located and opened with available technology and therefore can be understood and used as intended by the business.

Maintaining Digital Continuity depends on the alignment of an organisation's information assets and the technical environment that supports them in order for the organisation to use its information as it needs to.

The diagram below shows this alignment between information assets, environment and utility providing continuity:

- *Assets* could be entire file stores, or databases, or it could refer to individual files.
- *Environment* refers to any technology required to mediate access to the assets – everything from hardware to operating systems, networks, platforms and applications.
- *Utility* refers to any uses a business wants to make of its information assets. This can be defined at a technical level, for example the access that particular technologies support, and at a higher business level, for example, the need to use information as evidence, requiring provenance and audit trails.

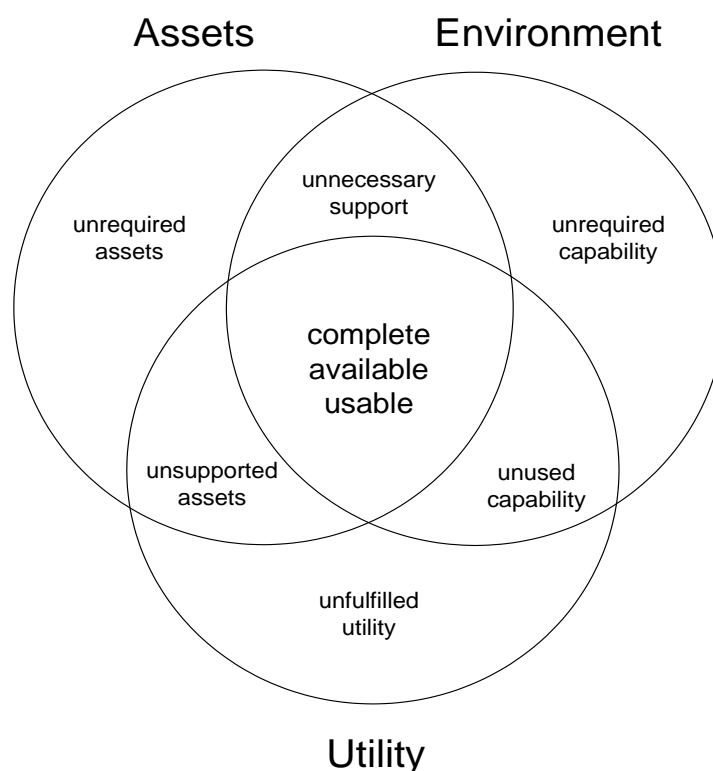


Figure 1 - Aligning assets, environment and utility

Over time the technology environment, the way the information assets are managed, and the business needs are likely to change - and it's during times of change that digital information is most vulnerable. Active intervention is required to keep digital information and its supporting technological environment aligned to meet business needs.

Improving the alignment of information and technology can also lead to efficiency gains through the elimination of unnecessary capability or assets. For example:

- Tools and services to help an organisation understand its technical environment can lead to efficiency savings through the identification of unneeded applications and licences.
- Tools and services to store information assets efficiently can lead to reduced storage requirements, potentially leading to savings in hardware and running costs.
- Tools and services to help identify information assets can improve efficiency in resource discovery, disposal and reuse.

For further information on the Digital Continuity Project please visit our website:
www.nationalarchives.gov.uk/digitalcontinuity

1 Tool Framework structure

At the present time we have defined four high-level use-cases to support Digital Continuity planning and risk mitigation:

- Understand Assets
- Understand Environment
- Change Assets
- Change Environment

Each of these high level use-cases is currently split into sub-lots. **These sub-lots are not organised by technology market sector, but by the continuity functionality they provide.** The eventual structure of the lots may or may not reflect this approach and your feedback will be a key driver in this respect.

We understand that any given product may fulfil functions across several sub-lots, and we are particularly interested in those cases. We may create specific sub-lots for multi-function technology if there is sufficient commonality between them. Please also let us know if you have technology which may be of use to organisations in managing Digital Continuity, but which does not appear to fit into an existing sub-lot.

Definition Comment, Section 1

1.1 Understand Assets

Understanding the nature of information assets is vital to assess the impact of change in the technological environment and usage requirements, and to change assets to align with them. It is also necessary to be able to assess the value of the assets to determine what type of support they will require across their lifecycle.

We have defined the following technology tool sub-lots to enable an organisation to understand their information assets:

File Characterisation

Database Characterisation

- Metadata Extraction
- Information Modelling

Content Analysis

- Asset Usage Analysis

Definition Comment, Section 1.1

1.1.1 File Characterisation

File characterisation tools give organisations the ability to discover important technical facts about files they hold so they can plan for the impact of technology changes. For example, knowledge about the precise file formats and versions an organisation holds will inform what applications may be needed to open and edit these files. Similarly, dates associated with a file help in understanding its age and activity, such as how often it is used or accessed. Organisations can use this information to help assess the risk of technical or economic obsolescence, decide what data objects / files to keep, what technology it needs, or in assessing the success of a file format conversion process.

In addition to characterising a single file, the ability to scan large file systems or other repositories in a batch processing mode, and to record the facts for use in later analysis is a useful additional capability. It is also useful if file characterisation can be automated using scripts or by integrating with third-party software through an API (Application Programming Interface).

Definition Comment, Section 1.1.1

Your Requirements, Section 1.1.1

1.1.2 Database Characterisation

Database characterisation refers to discovering and documenting the underlying structures of databases and the information held in them. Having this understanding is essential in planning and supporting development, standardisation, interoperability and data migration processes. The information collected by a such a tool may include, but is not limited to, the schema (e.g. tables, keys, indexes, views, data types), queries, stored procedures and relationships within or across databases.

A useful additional function is to analyse data values in a database for data quality. An example is checking that all dates adhere to the same format. This can be useful in ensuring successful migration of data into a new system.

Definition Comment, Section 1.1.2

Your Requirements, Section 1.1.2

1.1.3 Metadata Extraction

Metadata extraction refers to the process of pulling out information which is either explicitly embedded inside files, databases or other systems, or which can be derived from the content (e.g. number of words in a document). The extracted metadata may include, but is not limited to, creator information (e.g. author of a document), geographic information (e.g. a geo-tag in an image) and links to other assets (e.g. hyperlinks inside a web page). This can be useful to search and index assets by the metadata, to value them and to support their migration to other systems or formats.

Aside from scanning single assets at a time, the facility to scan large numbers of assets in a batch processing mode, and to record the extracted metadata for further analysis, is a useful additional capability. The ability to be controlled externally is also useful to integrate into existing systems (e.g. using scripting or via an Application Programming Interface).

Definition Comment, Section 1.1.3

Your Requirements, Section 1.1.3

1.1.4 Information Modelling

Information Modelling tools allow the organisation to define and document how information is managed in their organisation, or across organisations. Understanding the structure and relationships that exist in business information lets the organisation manage information continuity as systems are migrated to new systems or new systems are developed. This also facilitates interoperability between existing systems, and enhances information reuse in the organisation and beyond.

Information modelling technology may include, but are not limited to, tools which:

- Model information at a “business level”, showing the logical structure and relationships of information regardless of which system it is used in
- Show how the model is realised in underlying systems (e.g. logical to physical mapping)
- Allow the automatic discovery of information models by querying existing systems
- Validate that information models or system implementations conform to defined information standards.
- Facilitate the creation and management of controlled vocabularies, taxonomies, or ontologies.

Definition Comment, Section 1.1.4

Your Requirements, Section 1.1.4

1.1.5 Content Analysis

Content analysis tools enable an organisation to make sense of the information assets they possess, by analysing their content and providing ways to intelligently identify and locate related assets. This is useful to decide what assets have value and require ongoing continuity protection, and which can be disposed of.

Functionality provided by analysis tools includes but is not limited to producing groupings of related content, search for similar or duplicate content, and automatic classification or summarisation of content. Tools in this space may include, but are not limited to text analytics, data-loss prevention and e-discovery tools.

Definition Comment, Section 1.1.5

Your Requirements, Section 1.1.5

1.1.6 Asset Usage Analysis

Asset usage analysis tools allow the organisation to understand the usage patterns of their information assets. The information might include, but is not limited to, access dates and times, the types of access (e.g. edit or view), who accessed it, where the information was accessed from, and for what purpose.

Understanding usage can help an organisation determine the business value of its information assets, or identify suitable continuity and disposal strategies for them. A useful capability is the ability to report usage across assets held in many different systems.

Definition Comment, Section 1.1.6

Your Requirements, Section 1.1.6

1.2 Understand Environment

Access to information assets is mediated by their supporting technological environment. It is essential to understand your technological environment to assess the impact of change on your information assets, and to maintain alignment with business use requirements.

This understanding also facilitates efforts to mitigate the risks of unnecessary technological diversity and complexity, and to drive down costs, through standardising usage and by not continuing to license unused technology. It also indicates to the organisation where there may have been non-standard functionality or data formats in use which may pose a continuity risk.

We have defined the following technology tool sub-lots to enable an organisation to understand their technological environment:

- Technology Inventory
- Licence Management
- Software Usage Analysis
- Configuration Management

Definition Comment, Section 1.2

1.2.1 Technology Inventory

Technology inventory tools should help an organisation understand what information technology they have deployed. They may use this information to ensure that they have an appropriate technical environment to access and use their digital information in the ways they need. This information can also contribute to a continuity risk assessment, where an organisation determines their risk of being unable to access information appropriately, and it helps them understand the consequences of supporting or not supporting particular applications or hardware.

The technology inventoried may include, but is not limited to, hardware, operating systems, desktop applications, server applications, databases and application platforms. Determining the specific versions deployed is a useful additional capability.

Definition Comment, Section 1.2.1

Your Requirements, Section 1.2.1

1.2.2 Licence Management

Licence management technology should help an organisation understand what licences it has in use and the associated costs of this.

Licensing information can help an organisation understand what technology is deployed and used within the organisation, and the cost of using the technology. This can help it to decide whether supporting an information asset using that technology is justified, or whether there are more cost-effective ways of managing the information asset.

Definition Comment, Section 1.2.2

Your Requirements, Section 1.2.2

1.2.3 Software Usage Analysis

Software usage analysis tools help an organisation understand what software is actually being used, out of all the software that it may have. Knowledge of the software in use, and its capabilities, can help an organisation understand how its information assets are being used. Efficiencies can be realised through the identification of unnecessary or little-used technology.

Information collected by these tools may include, but is not limited to, how frequently the software is used, and by which users, divisions or departments.

Definition Comment, Section 1.2.3

Your Requirements, Section 1.2.3

1.2.4 Configuration Management

Configuration management tools should enable the organisation to understand what technologies are deployed across the organisation (similar to Technology Inventory), how those technologies are actually configured, and the technical dependencies between them.

As an example, a subset of the information collected might be the server hardware, what applications this hardware runs, and in what ways these applications depend on each other. Such tools allow the organisation to understand what impact a particular change may have over the whole technical environment.

Definition Comment, Section 1.2.4

Your Requirements, Section 1.2.4

1.3 Change Assets

It is necessary to change some digital assets to keep them aligned with changing technologies, to support new business usage requirements, and to manage them across their lifecycle.

Changes may involve migrating assets to new formats to keep them aligned with technology, moving them into new systems, repurposing the assets to enable re-use in new situations, enhancing the quality of the data in the information assets, and archiving the assets as they become semi-active or inactive.

We have defined several tool types which may assist an organisation in managing the continuity of their digital assets through active intervention:

File Format Conversion

- Content Transformation
- Database Migration
- Asset Deduplication
- Data Quality Enhancement
- Asset Archiving

Definition Comment, Section 1.3

1.3.1 File Format Conversion

File format conversion tools are used to convert a file into a different file format, e.g. changing a Word file into a PDF. Converting the format of a file can avoid lock-in to proprietary technologies, facilitate the re-use of information in different systems, and avoid the risk of holding information formatted for obsolescent technology.

It is not always necessary to preserve the exact information or functionality of the original file. In some cases, reducing the information or functionality of a file can make it easier and more cost-effective to manage. For example, a Computer Aided Design (CAD) file might be converted to a PDF if viewing the diagram is valuable but the diagram no longer needs to be edited.

In addition to converting a single file, the ability to handle many files in a batch processing mode is a desirable additional feature. It is also useful if file format conversion can be automated using scripts or by integrating with third-party software through an API (Application Programming Interface).

Definition Comment, Section 1.3.1

Your Requirements, Section 1.3.1

1.3.2 Content Transformation for Re-use

Content transformation allows an organisation to reuse an existing electronic file or data object for a new purpose either manually or in bulk. Examples include:

- compressing large files (such as picture/video/audio) to improve download and email transfer, and
- a new document (e.g. a PDF) is assembled automatically from structured source texts (e.g. XML) using a transformation script.

The information and file types might include, but are not limited to, images, video, audio, documents, and structured text. Of particular interest is the ability to handle bulk operations, where the same transformation is applied to many files at once. Additional desirable features are quality assurance processes, and the ability to be controlled by other programs or scripts, so that the transformation can be embedded within existing processes with minimal input from the user.

Definition Comment, Section 1.3.2

Your Requirements, Section 1.3.2

1.3.3 Database Migration

Database migration tools should allow the organisation to move information from one database to another. This is useful when standardising platforms in use across the organisation, to facilitate information re-use, and to mitigate against the risk of economic or technical obsolescence. They may also be used to export information from a database into a file which is archived, preserving the information, but avoiding the cost of keeping a live database system running.

Database migration tools may include, but are not limited to, functionality to:

- allow the structures of different databases to be related (e.g. map between schemas);
- produce diagrams of the database structures and their relationships;
- export or import those database structures to and from files;
- export or import the information held in the database to and from files;
- analyse the information found in a database;
- move information directly between databases.

Definition Comment, Section 1.3.3

Your Requirements, Section 1.3.3

1.3.4 Asset Deduplication

Asset deduplication tools should help an organisation locate duplicate or similar assets. This can lead to benefits such reduced storage costs, improved asset management and better success in search and discovery. Examples of duplicated or similar assets include multiple copies of the same attachment in an email system, or incremental drafts of a document in a file system.

Technologies in this space may include, but are not limited to, e-discovery and forensic tools. However, storage deduplication, which does not remove duplicate assets from the user perspective, but only reduces the space of storing them, is not in scope.

Definition Comment, Section 1.3.4

Your Requirements, Section 1.3.4

1.3.5 Data Quality Enhancement

Data quality enhancement tools should help an organisation identify and correct inconsistencies and errors in their information assets. Problems identified and/or corrected by these tools include, but are not limited to:

- Inconsistent formatting (e.g. figures recorded in different units);
- Incomplete data (e.g. missing information in a name and address);
- Inconsistent dates (e.g. an end date preceding a start date);
- Incorrect data types (e.g. text where only numbers are expected).

Organisations may perform data quality enhancement in order to help ensure their assets can continue to be located and used. It can also be used in a migration process, to help ensure that assets will be migrated successfully to their new environment. Data quality enhancement technology includes, but is not limited to, database schema and data value analysis, data cleansing and master data management.

Definition Comment, Section 1.3.5

Your Requirements, Section 1.3.5

1.3.6 Asset Archiving

Asset archiving technology is used to move information assets into dedicated archival systems. This may be done when the assets do not need to be changed further and when they are of less urgent need to the business. This ensures that older assets can be economically maintained without negatively impacting live systems.

When information assets are archived, it is worth considering what functionality and information is actually required once in an archival state. Information can often be converted at this point into less functional formats. For example, a database can be migrated into a file, or a word document can be converted into a PDF, lowering costs and enhancing the longevity of the asset.

Information assets may be, but are not limited to, web sites, intranets, wikis, emails, files and databases. Archival systems may allow the application of retention and disposal schedules to these assets, and facilitate the discovery and preservation of them.

Definition Comment, Section 1.3.6

Your Requirements, Section 1.3.6

1.4 Change Environment


The technological environment of digital assets changes rapidly, and those changes will be driven by individual business need and changes in the wider technological landscape. However, there are also specific technologies that can generally help to preserve access to digital assets and drive down the costs of maintaining access to them.

We have identified three types of technology that can be deployed to manage Digital Continuity:

Data Management Technologies

- Hardware Virtualisation
- Viewer Software

Definition Comment, Section 1.4



1.4.1 Data Management Technologies

We are interested to understand more about current technologies that will enable an organisation to reduce data volumes. The benefits we are seeking in reducing data volumes include:

- 1) A reduction in overall volumes will lead to easier identification and management of digital continuity issues.
- 2) A reduction in overall volumes will lead to more efficient use of physical storage resulting in potential technology cost savings (e.g. fewer physical devices) and potentially reduced running cost (less electricity) which in turn can contribute to an organisations Green targets.

For this lot we are interested to hear about technologies which can deliver one or more of the above outcomes.

Definition Comment, Section 1.4.1

Your Requirements, Section 1.4.1

1.4.2 Hardware Virtualisation

Hardware virtualisation is a technology widely used to run more services with less hardware infrastructure, and to enable more flexible management of those services. It allows services to be managed separately from any particular hardware. This can lead to efficiencies in terms of cost, maintenance, testing and energy usage.

Virtualisation can be used to maintain older operating systems and applications where access to them is still required, but the cost of migrating them to modern technology, or maintaining old hardware, is prohibitive. This can be achieved by running the old systems as a live virtual system, or by archiving a virtual “snapshot” of the old system off-line to file-storage.

Definition Comment, Section 1.4.2

Your Requirements, Section 1.4.2

1.4.3 Viewer Software

Viewer software allows an organisation to see the content of information assets, without providing the full functionality of the tools originally used to create them.

Such technology allows an organisation to provide required access to its information assets, without performing file migration or maintaining expensive software whose full functionality is no longer required. A further benefit is these tools may allow assets to be accessed on a wider range of platforms, such as mobile phones.

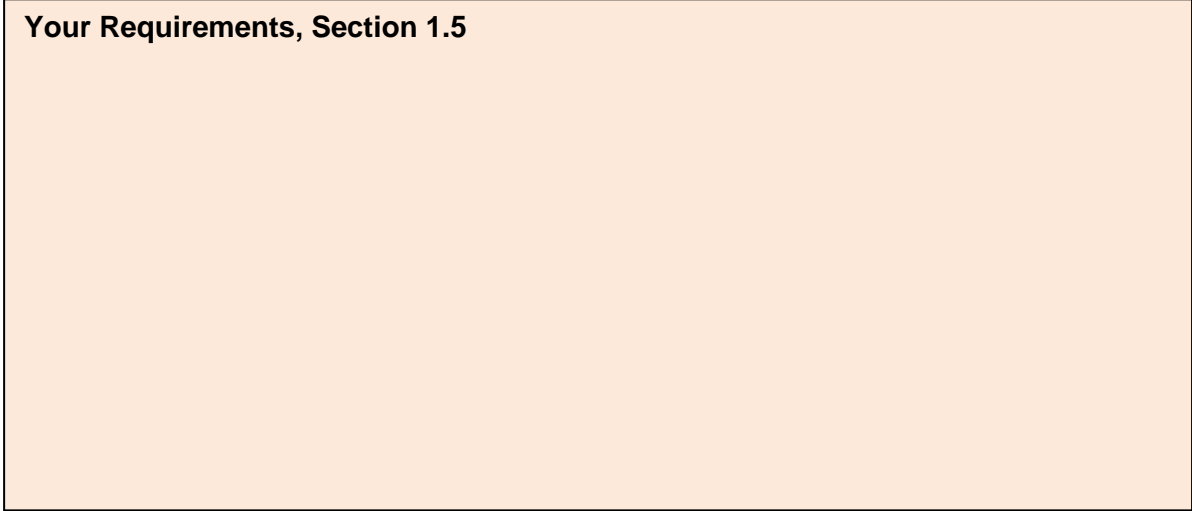
Definition Comment, Section 1.4.3

Your Requirements, Section 1.4.3

1.5 Licensing

We are asking suppliers to outline their licensing models. Please tell us if you have any particular issues or requirements for software licensing.

Your Requirements, Section 1.5



2 Services Framework Structure

Within this Framework, we are looking for suppliers of services to help clients follow Digital Continuity Guidance. Services include both consultancy and implementation of recommendations, perhaps using tools procured elsewhere in this Framework.

The services are divided into five lots according to their use in Digital Continuity. These are:

- Information Management Services
- Software Licence Management Services
- Efficiency Improvement/Technical Environment Change Services
- Format Conversion/Content Transformation Services
- File/Data Recovery Services

Even if some of these services are available on other frameworks, an important distinguishing factor is that the contract will oblige the supplier to deliver in accordance with Digital Continuity Best Practice Guidance.

Definition Comment, Section 2

2.1 Information Management Services

This lot is to help clients understand and manage their information assets.

Services may include consultancy in project management and business case justification; analysis of information access and sharing requirements; consultancy and delivery in performing a Digital Continuity risk assessment; consultancy in and management of information lifecycle processes; assessing the content and business value of information held; creating metadata schema.

Definition Comment, Section 2.1

Your Requirements, Section 2.1

2.2 Software Licence Management Services

This Lot is to help clients ensure they have the right applications to make appropriate use of their information assets.

Services may include consultancy; analysis of application usage; management of software licences.

Definition Comment, Section 2.2

Your Requirements, Section 2.2

2.3 Efficiency Improvement/Technical Environment Change Services

Services provided through this Lot should help clients understand and/or improve their technical environments through which information assets are held and accessed. This lot is divided into the following six specialisms:

Consultancy

- Virtualised hardware environment services

File archiving services

- Email archiving services

Data management services

- Database migration services

It is recognised that different suppliers may supply a different combination of services in this space. Furthermore, it also allows clients to solicit vendor-independent advice.

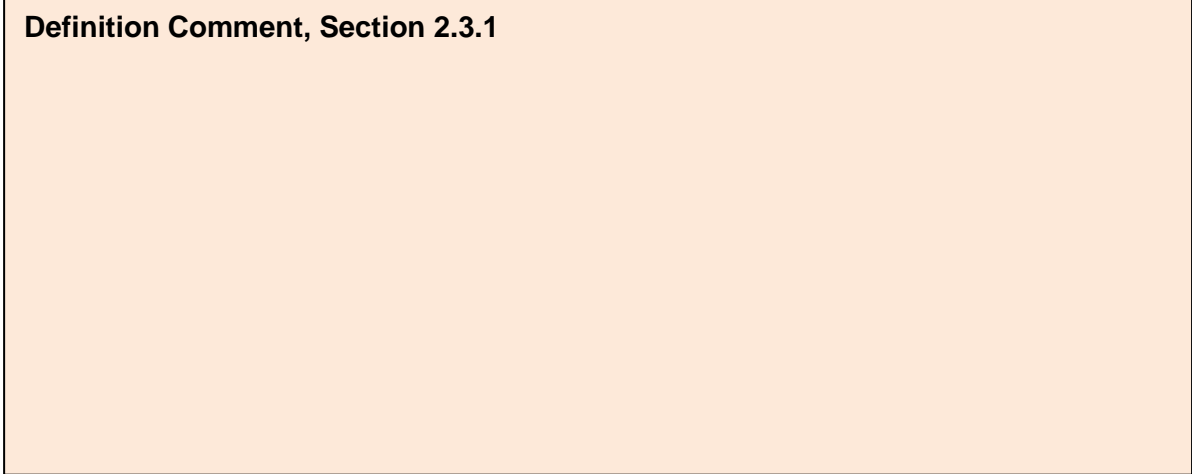
Definition Comment, Section 2.3



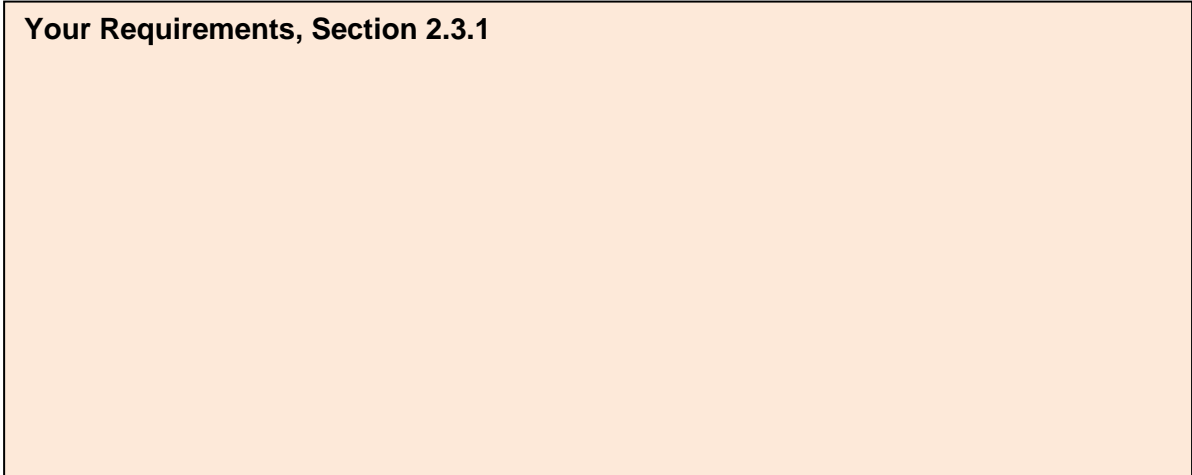
2.3.1 Consultancy

This includes consultancy of the client's use of file and data storage capabilities; producing recommendations over efficiency savings and Digital Continuity risk mitigations, which may include use of technologies such as virtualisation, archiving and storage-level deduplication.

Definition Comment, Section 2.3.1



Your Requirements, Section 2.3.1



2.3.2 Virtualised hardware environment services

This is to deliver a virtualised hardware environment in which legacy applications may continue to be used within a client's existing infrastructure.

Definition Comment, Section 2.3.2

Your Requirements, Section 2.3.2

2.3.3 File archiving services

This is to include integrating an archive within a client's existing infrastructure; planning and performing file migration into the archive including file level deduplication.

Definition Comment, Section 2.3.3

Your Requirements, Section 2.3.3

2.3.4 Email archiving services

This is to include integrating an email archive within a client's existing infrastructure; planning and performing email migration into the archive including file level deduplication.

Definition Comment, Section 2.3.4

Your Requirements, Section 2.3.4

2.3.5 Data management services

We are interested to understand more about services to implement technologies into a client's infrastructure that will enable an organisation to reduce data volumes. The benefits we are seeking in reducing data volumes include:

- 1) A reduction in overall volumes will lead to easier identification and management of digital continuity issues.
- 2) A reduction in overall volumes will lead to more efficient use of physical storage resulting in potential technology cost savings (e.g. fewer physical devices) and potentially reduced running cost (less electricity) which in turn can contribute to an organisations Green targets.

For this lot we are interested to hear about services which can deliver one or more of the above outcomes.

Definition Comment, Section 2.3.5

Your Requirements, Section 2.3.5

2.3.6 Database migration services

This lot covers database migration services including planning and delivery as well as data cleansing and transformation.

Definition Comment, Section 2.3.6

Your Requirements, Section 2.3.6

2.4 Format Conversion/Content Transformation Services

This lot offers services to change file format or content to suit business need or mitigate obsolescence. The types of files in scope may include word processor documents, images, video, presentations, email archive files, CAD files, compressed files, XML or other tagged files, comma or tab delimited files, among others. However it not expected for a single supplier to encompass all of these file formats. The service provided may include consultancy and delivery. Conversion may be of file format, character set, or compression algorithm.

Services to migrate files or email across file stores without conversion should be provided through Lot 0.

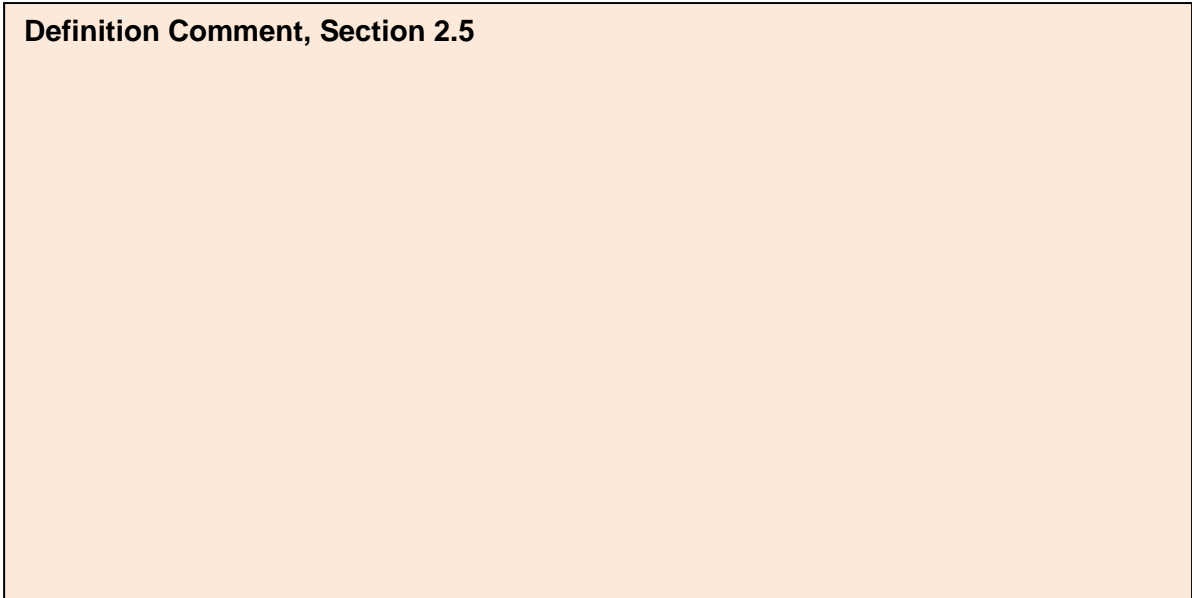
Definition Comment, Section 2.4

Your Requirements, Section 2.4

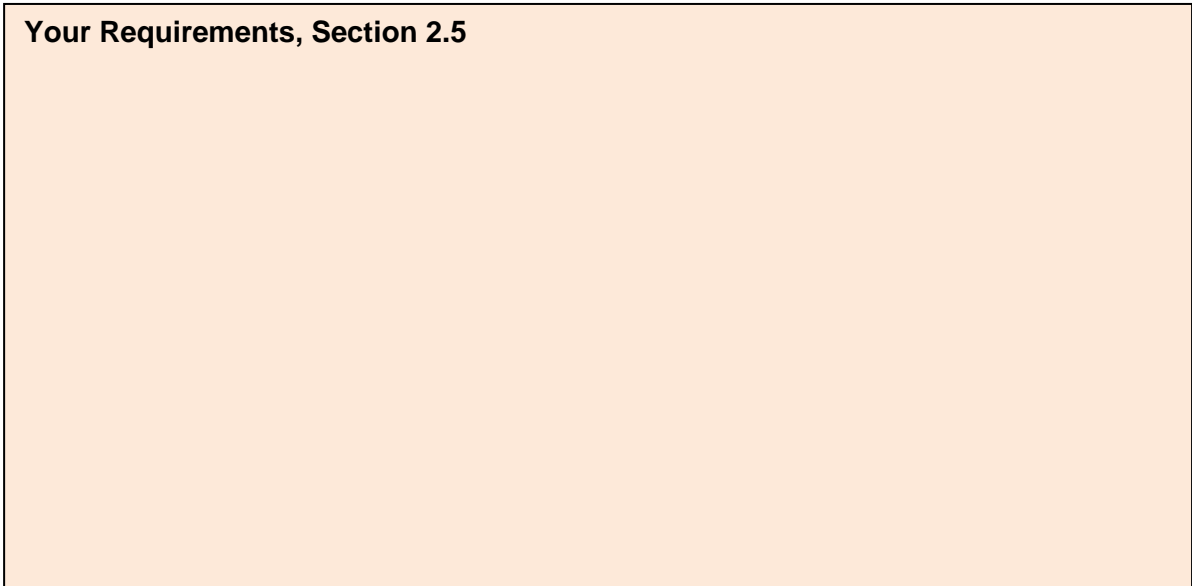
2.5 File/Data Recovery Services

This Lot offers services in consultancy; data forensics and recovery, including restoring data on damaged media; restoring a corrupted file; recovering data from obsolete and unreadable file format.

Definition Comment, Section 2.5



Your Requirements, Section 2.5



3 General Comments

Your General Comments

